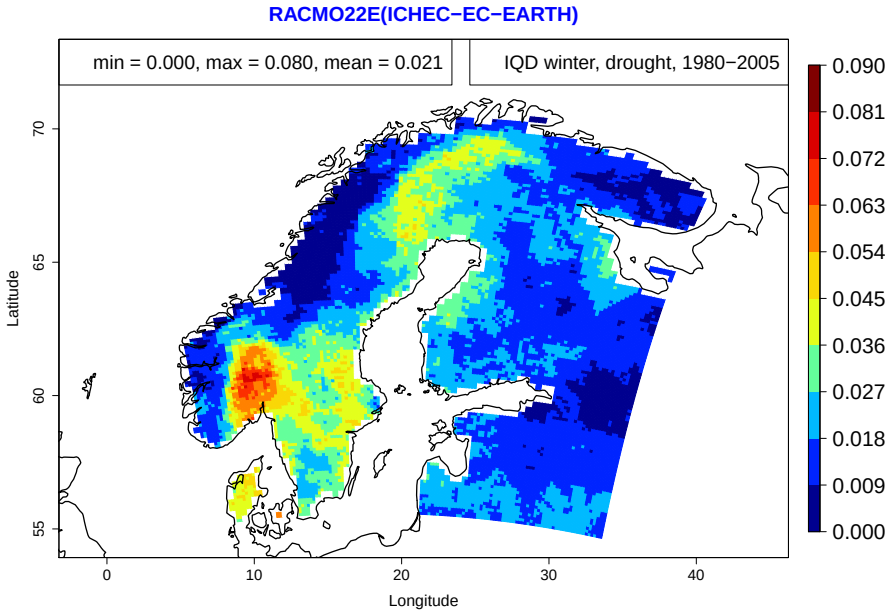
Norsk Regnesentral
NORWEGIAN COMPUTING CENTER

Note

Evaluation of precipitation output from the EURO-CORDEX climate ensemble using E-OBS data



Note no SAMBA/18/17

Authors Silius M. Vandeskog
Marion Haugen
Thordis L. Thorarinsdottir

Date 1st September 2017

Norsk Regnesentral
NORWEGIAN COMPUTING CENTER

The authors

Silius M. Vandeskog is a master's student in industial mathematics at the Norwegian University of Science and Technology, Marion Haugen is Senior Research Scientist and Thordis L. Thorarinsdottir is Chief Research Scientist at Norwegian Computing Center.

Norwegian Computing Center

Norsk Regnesentral (Norwegian Computing Center, NR) is a private, independent, non-profit foundation established in 1952. NR carries out contract research and development projects in information and communication technology and applied statistical-mathematical modelling. The clients include a broad range of industrial, commercial and public service organisations in the national as well as the international market. Our scientific and technical capabilities are further developed in co-operation with The Research Council of Norway and key customers. The results of our projects may take the form of reports, software, prototypes, and short courses. A proof of the confidence and appreciation our clients have in us is given by the fact that most of our new contracts are signed with previous customers.

Norsk Regnesentral
Norwegian Computing Center
Postboks 114, Blindern
NO-0314 Oslo, Norway

Besøksadresse
Office address
Gaustadalléen 23
NO-0373 Oslo, Norway

Telefon · telephone
(+47) 22 85 25 00
Telefaks · telefax
(+47) 22 69 76 60

Internett · internet
www.nr.no
E-post · e-mail
nr@nr.no

Title	**Evaluation of precipitation output from the EURO-CORDEX climate ensemble using E-OBS data**
Authors	**Silius M. Vandeskog** <siliusv@gmail.com> **Marion Haugen** <marionh@nr.no> **Thordis L. Thorarinsdottir** <thordis@nr.no>
Date	1st September 2017
Publication number	SAMBA/18/17

Abstract

Global climate models are used for projecting climate changes on a global scale. However, when we need information for local climate changes, a regional climate model with a finer grid that uses the global climate model as boundary conditions is necessary to gain more precise information. Therefore it is important to make good regional climate models that are unbiased when projecting climate changes. This note investigates the fit of precipitation projections from nine combinations of global and regional climate models from EURO-CORDEX and four bias correction methods applied to some of these. This is performed by comparing the models with data from the E-OBS dataset using integrated quadratic distance (IQD).

All the climate models had difficulties with their projections in the areas of Fennoscandia with high amount of daily precipitation or long drought periods, while the IQD was improved for all climate models after bias correction, the IQD was still highest in the areas with more extreme data after bias correction. The LSCE-IPSL-CDFt-EOBS10-1971-2005 bias correction method obtained the best results for all our tests. However, this method was calibrated using the E-OBS dataset. As the other bias correction methods used different datasets the current results should be compared to an evaluation using alternative observation-based datasets.

Keywords	Climate model evaluation, precipitation, drought, integrated quadratic distance
Target group	Climate scientists, users of climate information
Availability	Open
Project	PostClim
Project number	220778
Research field	Statistics, climate science
Number of pages	58

Contents

1 Introduction

Projecting future precipitation amounts and drought periods are important aspects, for example, when making constructions that should withstand a 100 year flood or when making sure that a city has enough drinking water supplies for the future. To estimate future climate changes, global climate models (GCM) are used to simulate changes on a global scale. However, to project climate changes locally with some precision, regional climate models (RCM) with finer resolution have to be constructed with boundary conditions from a global climate model. These regional models often suffer from various biases, and it is useful to develop bias correction methods that can be applied to the local models.

In this note we will examine properties of the precipitation projections over Fennoscandia for the years 1980 to 2005. This will be examined for nine combinations of global and regional climate models from EURO-CORDEX and four bias correction methods applied to some of these. We will focus on using the integrated quadratic distance (IQD) for evaluating the fit of the models against data from the E-OBS dataset. The two main aspects we will examine are the daily precipitation data and the number of drought periods and their lengths. This will be performed separately for the summer and winter seasons.

The remainder of the note is organized as follows. Section 2 describes the datasets used for all evaluations. Section 3 explains the theory behind IQD and the different bias correction models. Section 4 explains what was done to obtain our results, which are presented and discussed in section 5. Finally, the conclusion is provided in section 6.

Plots displaying the E-OBS precipitation and drought data from 1980 to 2005 are given in Appendix A. The other appendices give plots for some selected climate models. Scatter plots displaying IQD of raw data and bias corrected data are given in Appendix B for precipitation and Appendix F for drought. Plots displaying the mean IQD, with and without bias correction, are given in Appendix C for precipitation and Appendix E for drought. Finally, the difference between the mean bias corrected summer precipitation and the mean observed summer precipitation are illustrated in Appendix D.

2 Data

2.1 Climate models

A total of nine different combinations of global and regional climate models from EURO-CORDEX are chosen for the evaluation. These can be seen in Table 1. Each model contains daily precipitation data from various start and end dates. The precipitation is measured in $\mathrm{kg\,m^{-2}\,s^{-1}}$. There has also been conducted several bias corrections to five of the climate models from Table 1. The correction methods are listed in Table 2. Climate models and bias correction methods will hereby be named by their numbers, found in the two tables.

The observed precipitation data is collected from the E-OBS dataset (Haylock et al., 2008). Here precipitation is measured in mm. All months have 30 days in the models 8 and 9, while the other models follow the standard Gregorian calendar. The E-OBS dataset also uses the Gregorian calendar.

Table 1. Nine GCM/RCM combinations from EURO-CORDEX used in our testing.

Model nr.	Global climate model	Ensemble member	Regional climate model	Institute	Institution name	Adjusted
1	CNRM-CERFACS-CM5	r1i1p1	CCLM4-8-17	CLMcom	Climate Limited-area Modelling Community	Yes
2	CNRM-CERFACS-CM5	r1i1p1	ALADIN53	CNRM	Météo-France / Centre National de Recherches Météorologiques	No
3	ICHEC-EC-EARTH	r1i1p1	RACMO22E	KNMI	Royal Netherlands Meteorological Institute	Yes
4	ICHEC-EC-EARTH	r2i1p1	CCLM4-8-17	CLMcom	Climate Limited-area Modelling Community	Yes
5	MPI-ESM-LR	r1i1p1	CCLM4-8-17	CLMcom	Climate Limited-area Modelling Community	Yes
6	MPI-ESM-LR	r1i1p1	REMO2009	MPI-CSC	Helmholtz-Zentrum Geesthacht, Climate Service Center, Max Planck Institute for Meteorology	Yes
7	MPI-ESM-LR	r2i1p1	REMO2009	MPI-CSC	Helmholtz-Zentrum Geesthacht, Climate Service Center, Max Planck Institute for Meteorology	No
8	MOHC-HadGEM2-ES	r1i1p1	CCLM4-8-17	CLMcom	Climate Limited-area Modelling Community	No
9	MOHC-HadGEM2-ES	r1i1p1	RACMO22E	KNMI	Royal Netherlands Meteorological Institute	No

Table 2. Four bias correction methods applied to EURO-CORDEX simulations.

Method nr.	Bias correction method	Calibration dataset	Years	Institution name	Applied to model nr.
1	LSCE-IPSL-CDFt	EOBS10	1971-2005	Institut Pierre Simon Laplace	1,3,4,5,6
2	METNO-QMAP	MESAN	1989-2010	Norwegian Meteorological Institute	1,3,4,5,6
3	SMHI-DBS45	MESAN	1989-2010	Swedish Meteorological and Hydrological Institute	1,3,5,6
4	IPSL-CDFT21	WFDEI	1979-2005	Institut Pierre Simon Laplace	3,4,6

2.2 Preparation

The data for the climate models is stored in NetCDF-files, and have to be prepared before it can be analysed. First we have to merge several files to retrieve data for the entire period from 1980 to 2005. Then unwanted data is removed. This concerns data before 1980, after 2005 and February 29 to make sure each year has the same number of measurements. February 29 is not removed from the climate models with 30 days in each month. Additionally, we are only interested in the 140x155 grid covering Fennoscandia. The E-OBS data is prepared in the same way. All the data from the climate models is then converted from $\mathrm{kg\,m^{-2}\,s^{-1}}$ to mm. Some models contain negative values, but these have small absolute values and are likely due to rounding errors. The negative values are set equal to zero. Testing gives that the number of zeros after the conversion of negative values has the same magnitude as the number of zeros for models with only non-negative data. To be able to separate between land and sea, NA values are inserted from the E-OBS data wherever there is water in the grid we examine. The data is then sorted by seasons, where summer consists of June, July and August, while winter consists of December, January and February. This gives a total of 92 observations per summer and 90 observations per winter. The models with 30 days in each month have 90 observations for both seasons.

3 Theory

3.1 IQD

We denote a precipitation observation by $y \in \Omega$ where Ω denotes the non-negative real axis $\mathbb{R}_{\geq 0}$. Similarly, for droughts, we have that y is the drought length in days with $\Omega = \{0, 1, 2, \ldots\}$. A probabilistic prediction for y is given by a distribution function with support on Ω denoted by $F \in \mathcal{F}$ for some appropriate class of distributions $\mathcal{F}$, with the density denoted by f if it exists.

Scoring rules assess the accuracy of probabilistic predictions by assigning a numerical penalty to each prediction-observation pair. Specifically, a scoring rule is a mapping

$$S : \mathcal{F} \times \Omega^d \to \mathbb{R} \cup \{\infty\} \tag{1}$$

where, in our notation, a smaller penalty indicates a better prediction. A scoring rule is *proper* relative to the class $\mathcal{F}$ if

$$\mathbb{E}_G S(G, Y) \leq \mathbb{E}_G S(F, Y) \tag{2}$$

for all probability distributions $F, G \in \mathcal{F}$, that is, if the expected score for a random observation Y is optimized if the true distribution of Y is issued as the prediction. The scoring rule is *strictly proper* relative to the class $\mathcal{F}$ if (2) holds with equality only if $F = G$. Propriety will encourage honesty and prevent hedging, which coincides with Murphy's first type of goodness (Murphy, 1993).

In some cases, in particular in climate modelling, it is of interest to compare the predictive distribution F against the true distribution of the observations which is commonly approximated by the *empirical distribution function* of the available observations $y_1, \ldots, y_n$,

$$\widehat{G}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}\{y_i \leq x\}. \tag{3}$$

The two distributions, F and $\widehat{G}_n$, can be compared using a *divergence*

$$D : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_{\geq 0} \tag{4}$$

where $D(F, F) = 0$.

Assume that the observations $y_1, \ldots, y_n$ forming the empirical distribution function $\widehat{G}_n$ are independent with distribution $G \in \mathcal{F}$. A propriety condition for divergences corresponding to that for scoring rules (2) states that the divergence D is *k-proper* for a positive integer k if

$$\mathbb{E}_G D(G, \widehat{G}_k) \leq \mathbb{E}_G D(F, \widehat{G}_k) \tag{5}$$

and *asymptotically proper* if

$$\lim_{k \to \infty} \mathbb{E}_G D(G, \widehat{G}_k) \leq \lim_{k \to \infty} \mathbb{E}_G D(F, \widehat{G}_k) \tag{6}$$

for all probability distributions $F, G \in \mathcal{F}$ (Thorarinsdottir et al., 2013). While the condition in (6) is fulfilled by a large class of divergences, only score divergences have been shown to fulfill (5) for all integers k. A divergence D is a *score divergence* if there exists a proper scoring rule S such that $D(F, G) = \mathbb{E}_G S(F, Y) - \mathbb{E}_G S(G, Y)$. A score divergence that assesses the full distributions is the *integrated quadratic distance* (Thorarinsdottir et al., 2013):

$$\mathrm{IQD}(F, G) = \int_{-\infty}^{+\infty} \big(F(x) - G(x)\big)^2 \mathrm{d}x \tag{7}$$

In the following, we will apply the IQD to compare empirical distributions of climate model output and observations.

3.2 Bias correction methods

All the bias correction methods we have tested work mostly in the same way. They use a data set with precipitation data as calibration, then the distribution from a raw model is transformed to fit better with the calibration data. The calibration datasets and the various transformations however, differ in the correction methods we examine.

3.2.1 QMAP

A popular bias correction approach is based on the quantile-mapping technique (Gudmundsson et al., 2012). This technique maps a model output x with cumulative distribution function (CDF) F_X, to an observation y with CDF F_Y through a function h (Vrac et al., 2016):

$$y = h(x), \text{ such that } F_Y(y) = F_X(x) \tag{8}$$

Bias correction method 2 uses such a quantile-mapping technique. It can be found in the R package `qmap` (R Core Team, 2016).

3.2.2 CDFt

The Cumulative Distribution Function-transform is used in bias correction method 1 and method 4 and can be considered a variant of the empirical quantile-mapping (Gudmundsson et al., 2012)

$$h(x) = F_Y^{-1}(F_X(x)), \tag{9}$$

where F^{-1} is the inverse function of the CDF F. It first estimates the CDF F_{Yp} and F_{Xp} over a projection time period before applying the distribution-derived quantile mapping in (9) (Vrac et al., 2016). It is unknown to us whether there are any differences in the techniques of the LSCE-IPSL-CDFt method and the IPSL-CDFT21 method.

3.2.3 DBS45

Bias correction method 3 uses a distribution-based scaling for correcting regional climate models. First, the number of wet days is adjusted. This is done by identifying a cut-off value that reduces the percentage of wet days in the simulation to that of the MESAN observations. Then all days with less precipitation than the threshold are considered dry days. The remaining precipitation is then transformed to match the observed frequency distribution, using gamma distributions (Yang et al., 2010).

4 Methods

4.1 Precipitation

We want to evaluate the IQD for daily precipitation for the E-OBS data and the chosen climate models, separately for each model grid point. To do this we make 17 comparisons per model. One for each year between 1984 and 2000. We then compare one year from E-OBS with nine years of model data; four years before and four years after in addition to the year we want to examine. This gives a total of 90 or 92 observations from E-OBS and 810 or 828 observations from the climate models for each comparison. Calculations of IQD are done in parallel to save time. After calculating the IQD for each of the 17 comparisons, we find the mean IQD for the entire period in each grid point.

4.2 Drought

To examine how well the models project periods without rain, we make a vector that counts all periods without rain from 1980 to 2005 for each grid point. Data is stored in a way such that the number of periods without rain of length n is stored in entry number n of the vector. This gives an array of size $140 \times 155 \times$ maxLength, where maxLength is equal to 90 or 92, depending on what season we examine. We get two of these arrays per climate model. Since dry periods of small length are not very interesting, we remove the first seven entries of our data and only examine periods without rain of length greater than one week. We then calculate the IQD for our data. This time we only have one array per season for the entire period, so we only get one comparison per grid point. Although the IQD of something with the unit "# drought periods of length $n \in \{8, 9, 10, ..., \text{maxLength}\}$" does not provide much physical understanding, it is still useful for comparing the different models against each other.

5 Results and discussion

5.1 Precipitation

Mean IQD for daily precipitation over all grid points is shown in Figure 1. Almost all bias correction methods give a significantly lower IQD for both winter and summer than the raw data. The mean seems to be smaller for the winter season than for the summer season, but this is not by much. From looking only at the mean IQD, it seems that bias correction method 1 is the most successful method. Method 2 seems to achieve the worst results among the chosen methods. There is however much more variability in the data than one can see from this figure. It is not possible to draw error bars into this scatter plot, as they will make the figure hard to interpret.

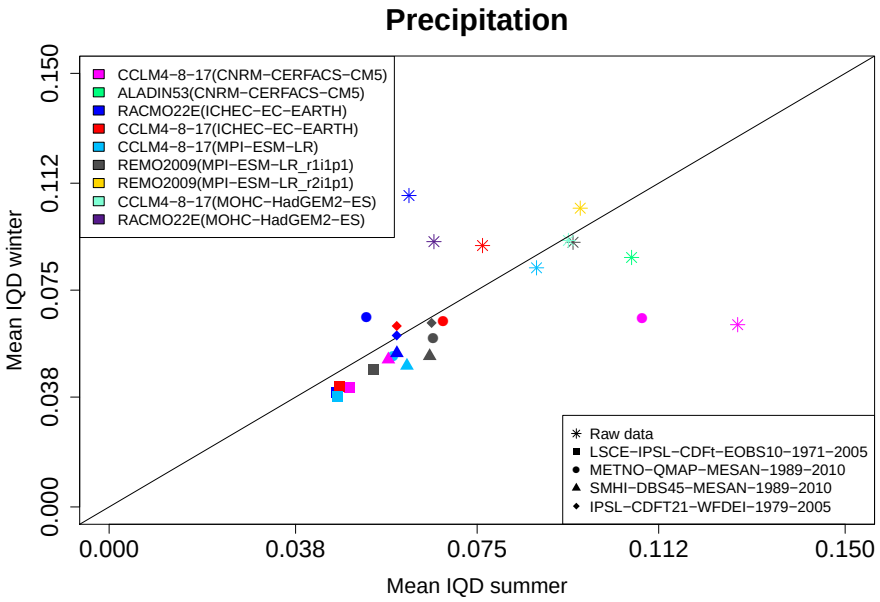


Figure 1. Scatter plot with the mean IQD for daily precipitation for all climate models and bias correction methods. Summer results are plotted along the x-axis and winter results are plotted along the y-axis. Colour indicates climate model, while shape indicates raw data and bias correction method.

Figure 2 shows a ranking of the mean precipitation IQD for summer and winter for all climate models and bias correction methods. 10% and 90% quantiles (error bars) for grid-point-specific values are also plotted. Even though the mean of the IQD is better for the corrected models, we see that the 90% quantiles often are higher than the mean of the raw data. Therefore we may find many grid points where the projections from the corrected models are worse than the projections from the raw models. We see that the variability for the different correction methods is higher in winter than in summer but mostly of the same order of magnitude within one season. This does not hold for the summer IQD, where the quantile bars for two of the CCLM4-8-17 regional climate models are much further apart for bias correction method 2 than for the other methods (see the error bars for the pink and red circles).

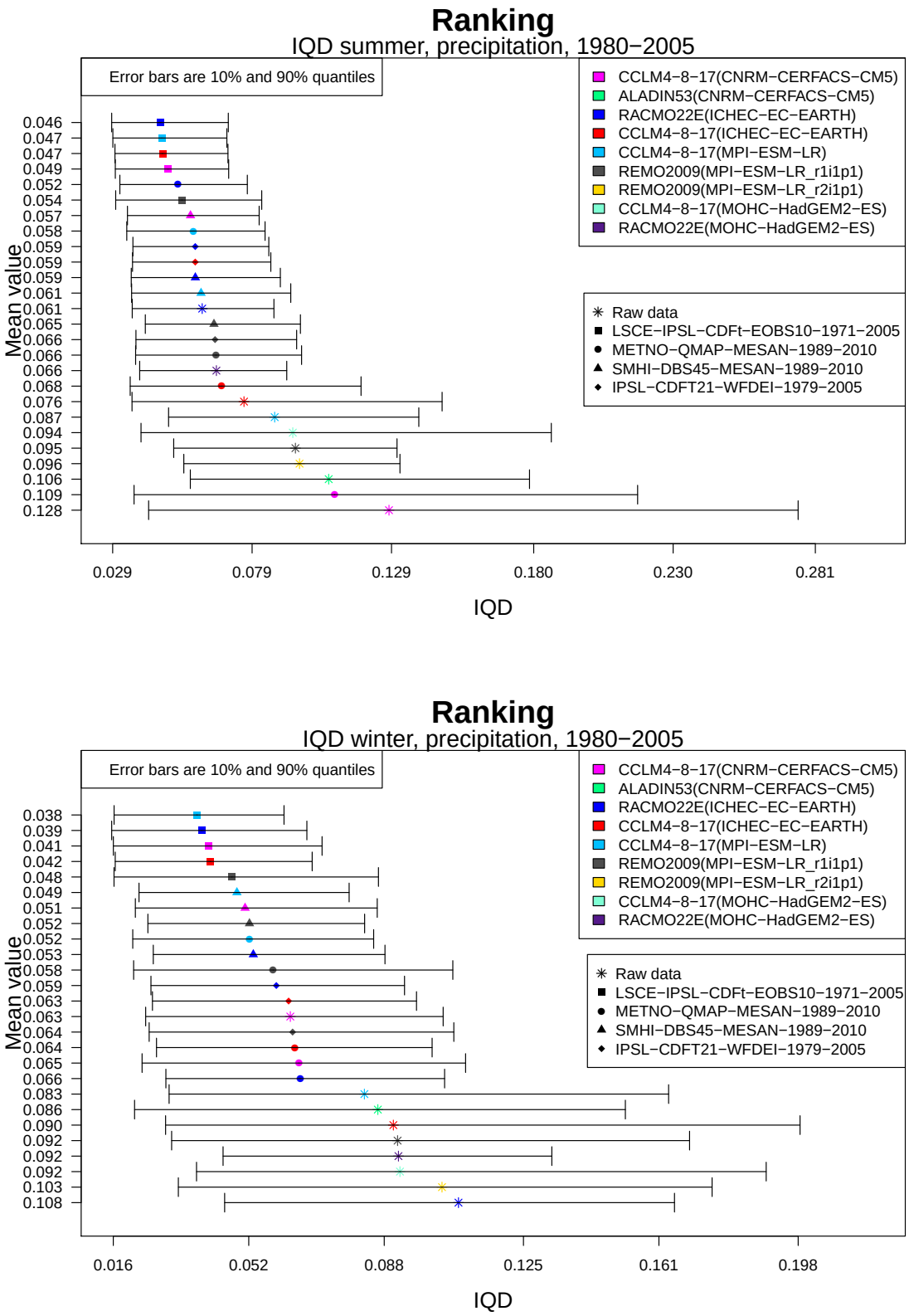


Figure 2. Ranking of the mean precipitation IQD for all climate models and bias correction methods. Summer is given on top and winter at the bottom. The bars indicate the 10% and the 90% quantiles for IQD values at individual grid points.

Since the variability in the data is this large, we make a scatter plot displaying the IQD at every grid point that shows the difference between the climate models before and after bias correction. Some examples for summer can be seen in Figure 3 and for winter in Figure 4. More plots are given in Appendix B. Bias correction method 1 mainly obtains a better IQD than the raw data. However, there are some areas where the projections get worse after correction. The other bias correction methods also lower the local IQD for most of the grid points, but there are more points where the projections get worse than there is for method 1. We also find that all correction models struggle more with improving every single grid point from the CCLM4-8-17 regional climate models than the other climate models.

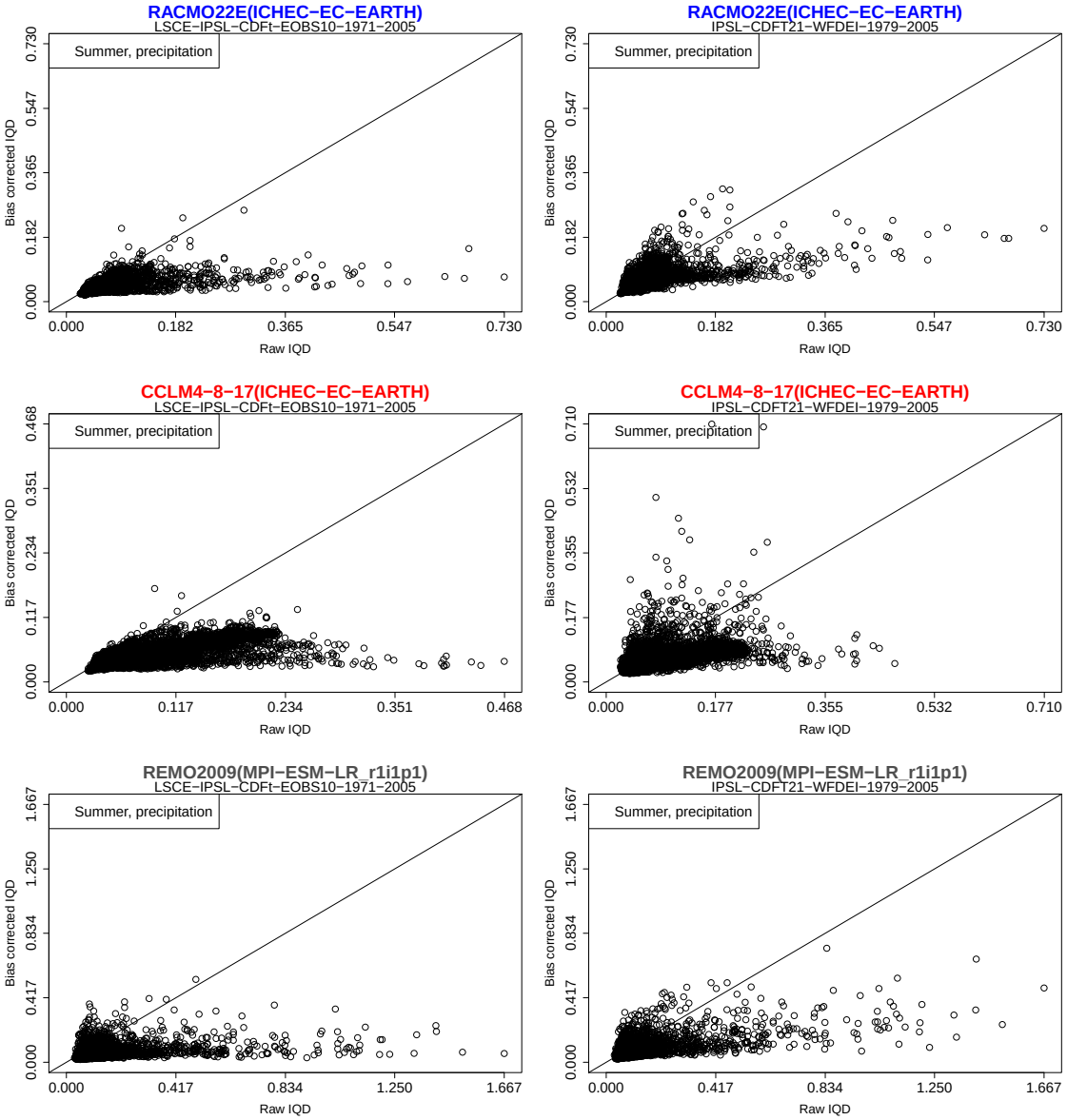


Figure 3. Scatter plots displaying precipitation IQD of raw data and bias corrected data for the models 3, 4 and 6 during summer. The figures on the left show bias correction method 1, while the figures on the right show bias correction method 4.

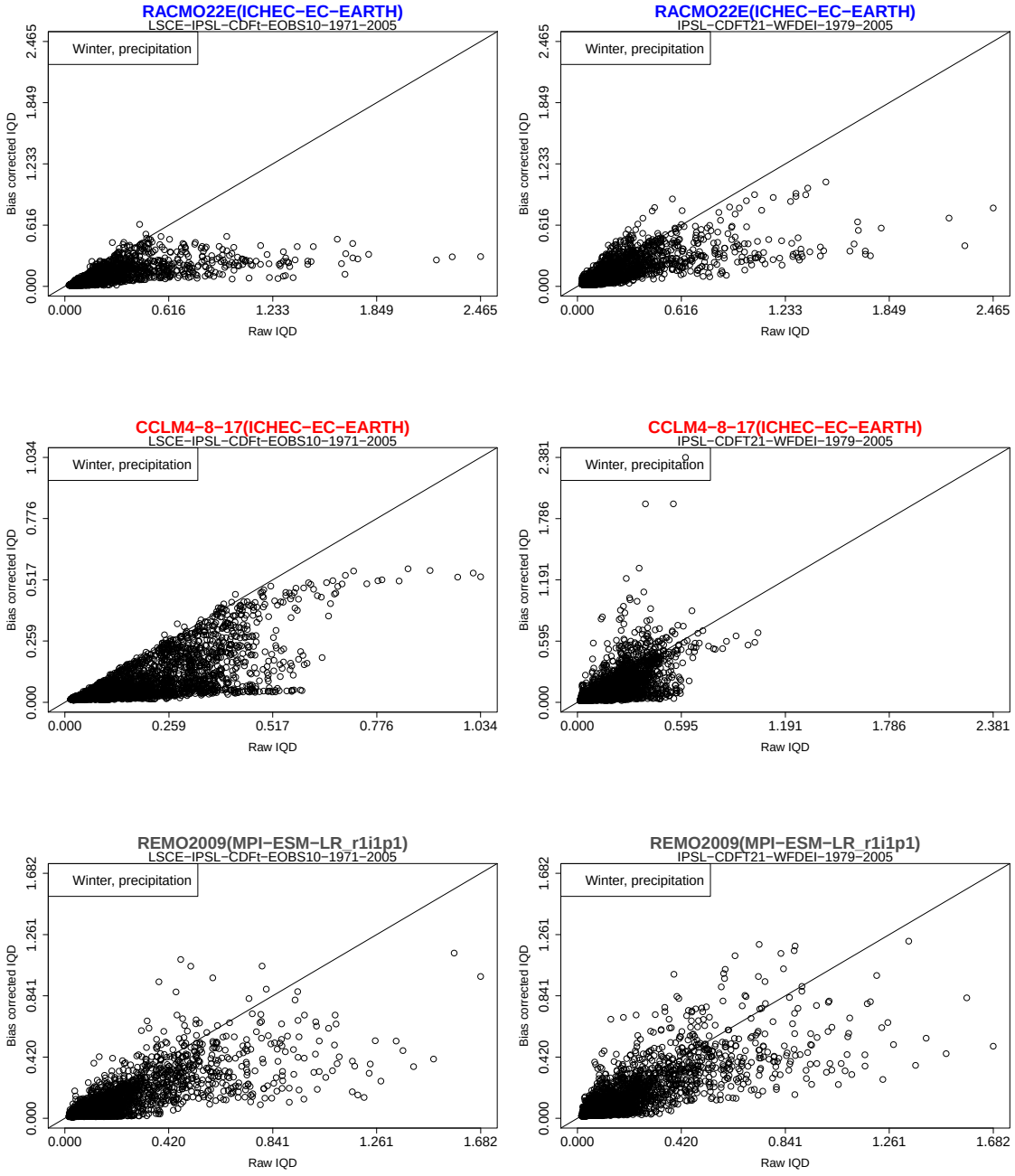


Figure 4. Scatter plots displaying precipitation IQD of raw data and bias corrected data for the models 3, 4 and 6 during winter. The figures on the left show bias correction method 1, while the figures on the right show bias correction method 4.

To examine the spatial properties of the errors for different climate models, the mean IQD at each grid point is plotted on top of a map of Fennoscandia. Figures 5 and 6 show results for model 3 during summer and winter, before and after applying bias correction method 4. More plots are given in Appendix C. These plots revealed that the IQD is highest along the coast of Norway and sometimes also south-east in the Baltic states. For the other parts of the map, the IQD is quite small indicating a good correspondence between modelled and observed precipitation distributions. For the raw climate models, the IQD along the coast of Norway is much higher during winter than during summer, with maximum values almost twice that of the summer IQD. However, we see from Figure 1 that the mean IQD is mostly the same for the two seasons. This means that the IQD during winter in turn must be slightly better than the summer IQD for most of the other parts of Fennoscandia.

The same patterns that we find in the IQD for the raw models can also be found in the mean measured precipitation for E-OBS during summer and winter (see Appendix A). This implies that the raw climate models have difficulties with their projections whenever the amount of precipitation becomes too high. The bias correction methods manage to remove much of the errors along the coast, but the IQD is still larger there than for the rest of Fennoscandia. The mean for the winter IQD is still slightly lower than the mean summer IQD, and the errors along the coast of Norway are still larger during winter than during summer.

In order to search for systematic errors, we make plots displaying the mean precipitation data for summer for the observed data together with data from one model and one bias correction method applied to that model. We then investigate the 10 grid points that have the largest difference in IQD between the bias corrected models and the uncorrected models. We also look at the 10 grid points where the IQD is highest for the bias corrected models. This is done for every combination of climate model and bias correction method. We observe that the corrected models sometimes seem to keep the same distance from the raw models, independent of the error in the raw model. One example can be seen in Figure 7. There, the raw data from model 4 is too high in 9 of the 10 plots we investigate. Every time, the corrected model gives less precipitation than the raw model, and it is therefore closer to the observed data. In the last grid point, as seen to the right in Figure 7, the raw model correctly predicts the mean precipitation, but the corrected model still gives less precipitation than the raw model. Therefore, the adjusted output is less correct than the old one at this location.

Because of this apparent dependence between the corrected and the raw models, we make some further investigations of how the corrected models are affected by the raw models. This is done by comparing the difference between the mean of the raw models and the mean of the observed data with the difference between the mean of the bias corrected models and the mean of the observed data. Figures 8 and 9 show that in the south-west of Norway, the bias corrected model obtains two blue areas where it gives too little precipitation. The raw model however, better matches the observed data at these locations. To the west of these areas, the raw model gives too much precipitation. This

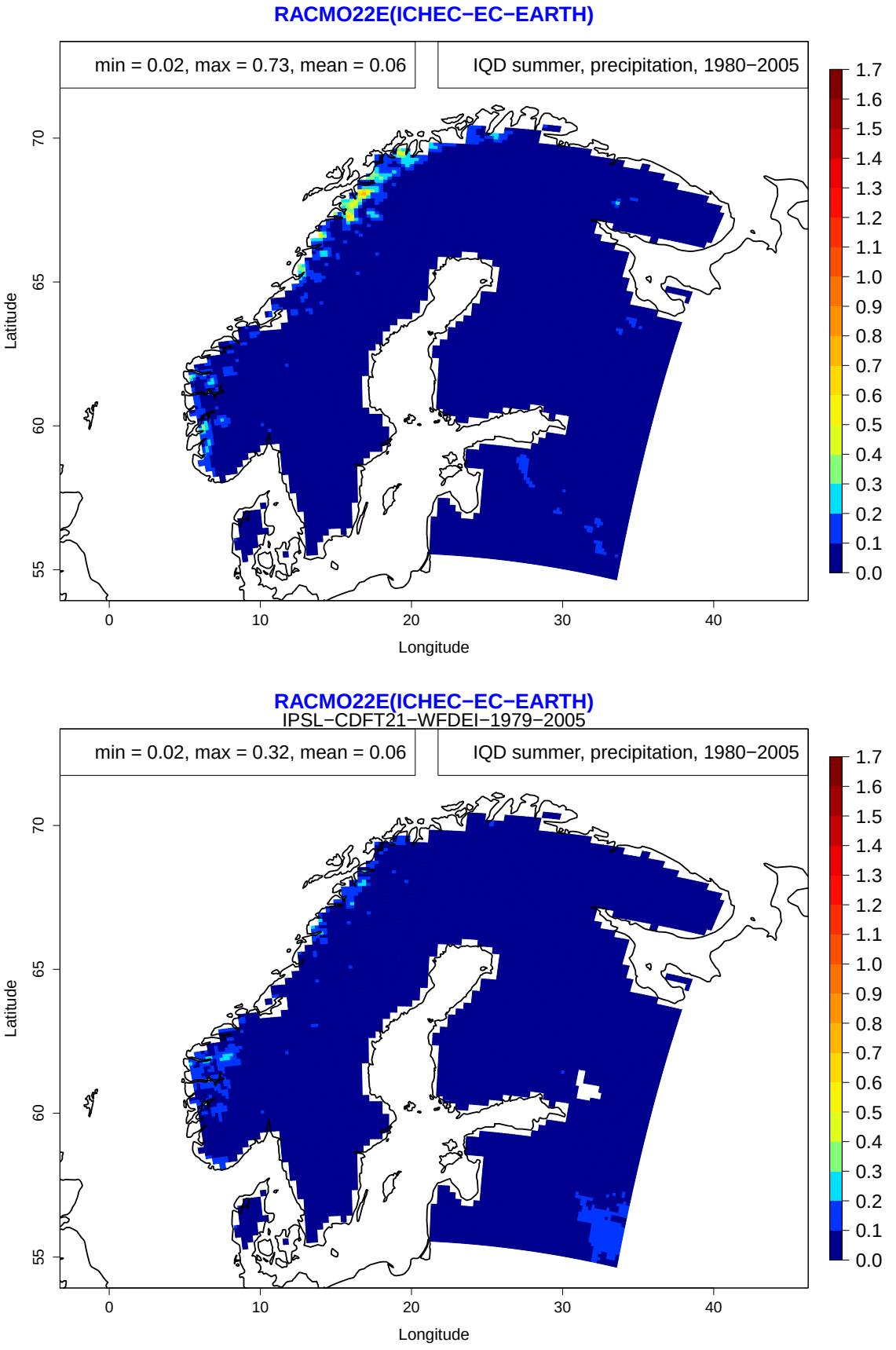


Figure 5. The mean precipitation IQD from 1980 to 2005 for Fennoscandia during summer. The upper figure shows the IQD for model 3 with no bias correction. The lower figure shows the IQD for model 3 after applying bias correction method 4.

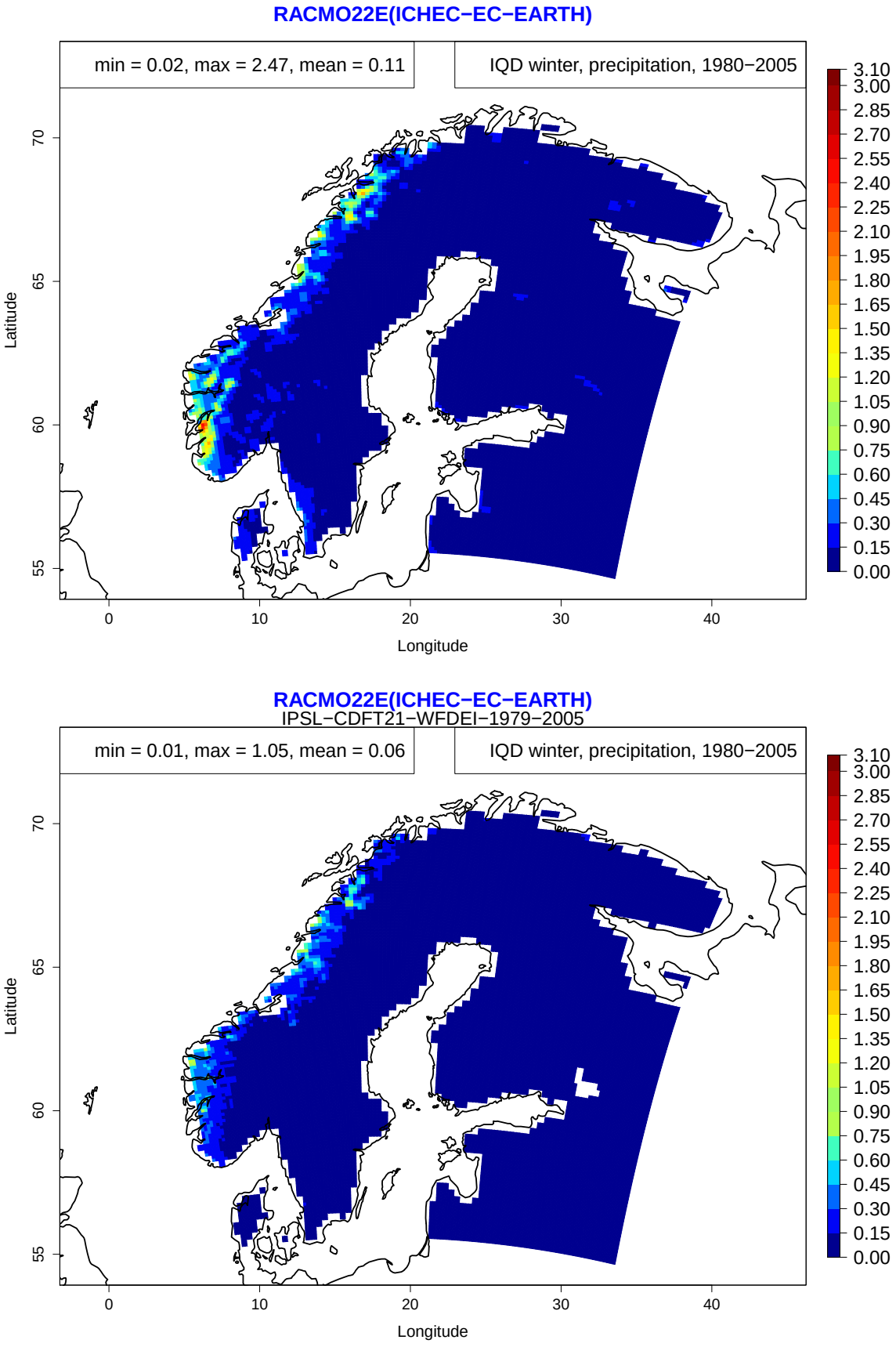


Figure 6. The mean precipitation IQD from 1980 to 2005 for Fennoscandia during winter. The upper figure shows the IQD for model 3 with no bias correction. The lower figure shows the IQD for model 3 after applying bias correction method 4.

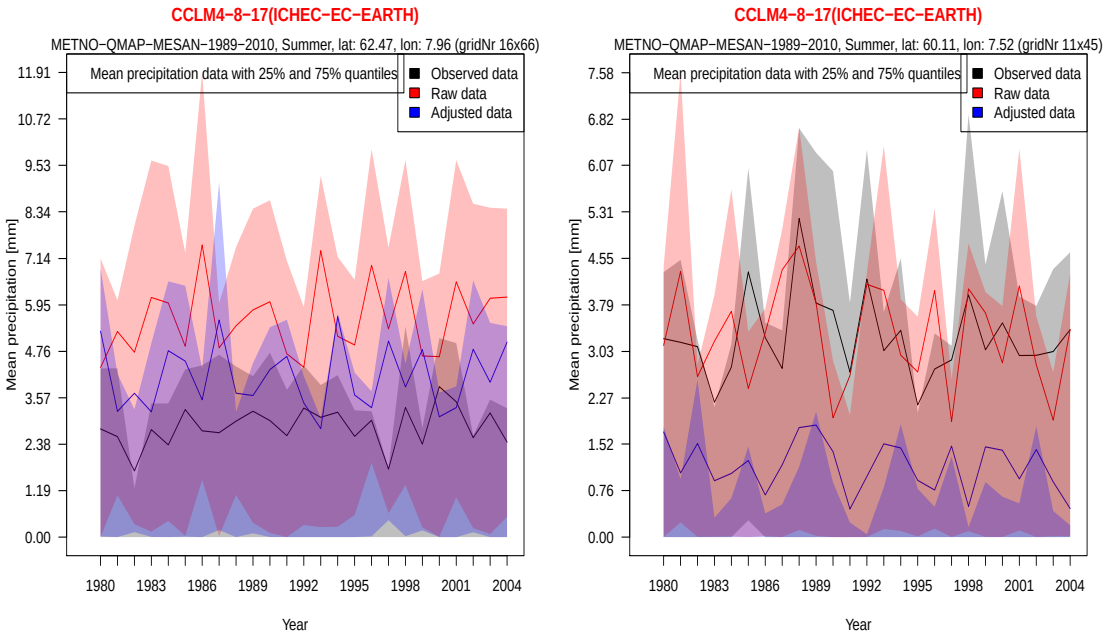


Figure 7. Yearly mean precipitation during summer for the observed data, model 4 ("Raw data") and bias correction method 2 applied to model 4 ("Adjusted data"). The coloured areas are bounded by the 25% and the 75% quantiles for summer.

has been corrected by the new model. From Figure 10 we see that the areas from the raw model with too much precipitation have been corrected.

Several plots showing the difference between the mean bias corrected summer precipitation from the models 3, 4 and 6 and the mean observed summer precipitation are found in Appendix D. The blue areas appear in every correction model that has not been calibrated on the E-OBS dataset. This seems to be an indicator that the choice of the calibration dataset holds more importance than the differences between the raw models and the observed data. However, the magnitude of the blue areas also does seem to change depending on how large the differences are for the raw models, so it might be that both the difference between the raw model and the E-OBS dataset and the choice of calibration dataset affect the corrected model.

By examining differences between raw models and observed data we also find that during summer almost all raw climate models project too much precipitation along the coast of Norway. This means that whenever there is a high amount of summer precipitation, the climate models give even more extreme amounts. When correcting this, method 2 seems to be correcting too much, and therefore gives too little precipitation compared to E-OBS. The other methods give more mixed results. During winter this is not as clear as during summer. For the correction methods, it seems that the projections of methods 1, 3 and 4 are mostly too low, while method 2 projects too much precipitation.

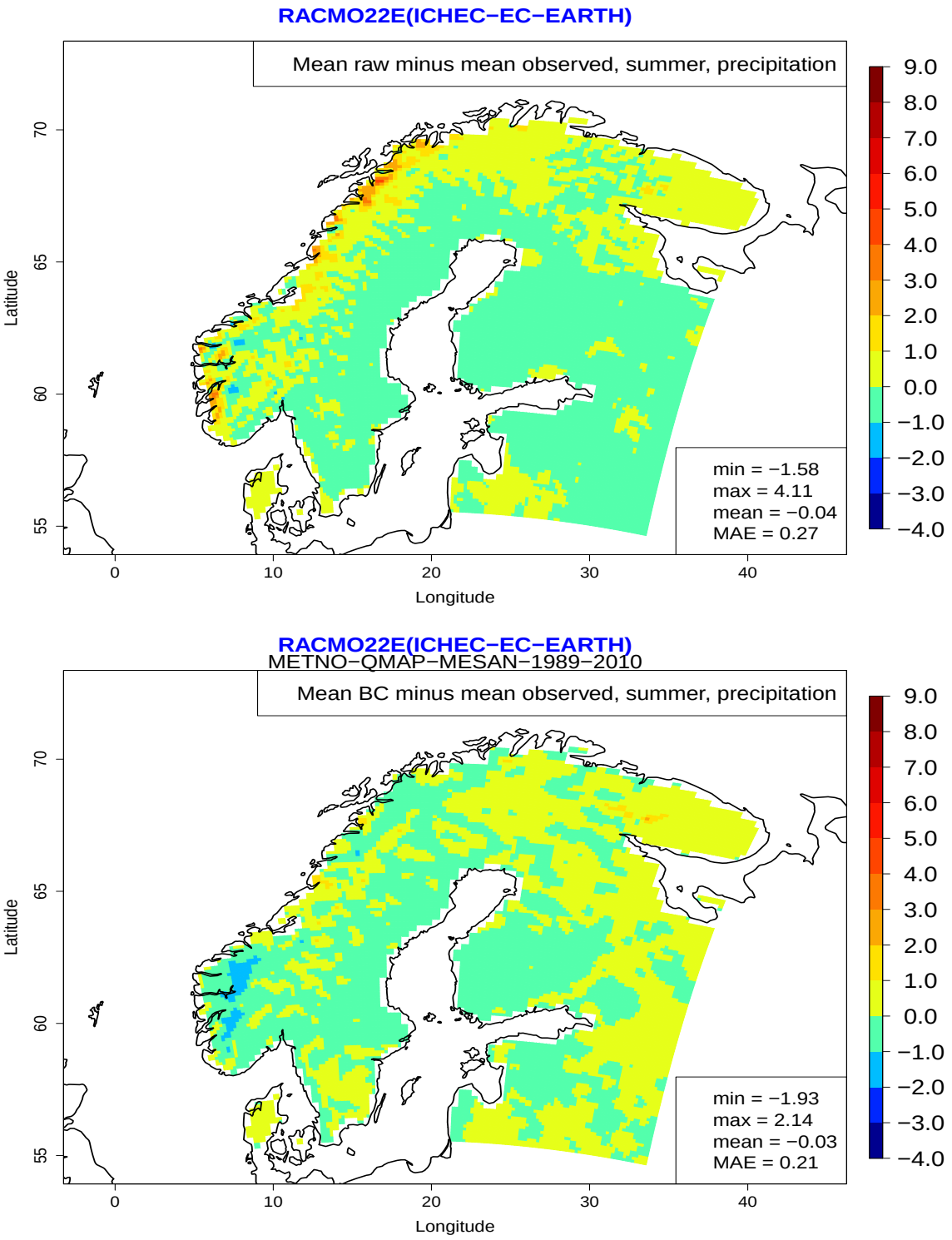


Figure 8. Difference between mean climate model precipitation and mean observed precipitation during summer from 1980 to 2005. The difference for the raw data from model 3 on top and the difference for bias correction ("BC") method 2 applied to model 3 at the bottom. MAE is the mean absolute error.

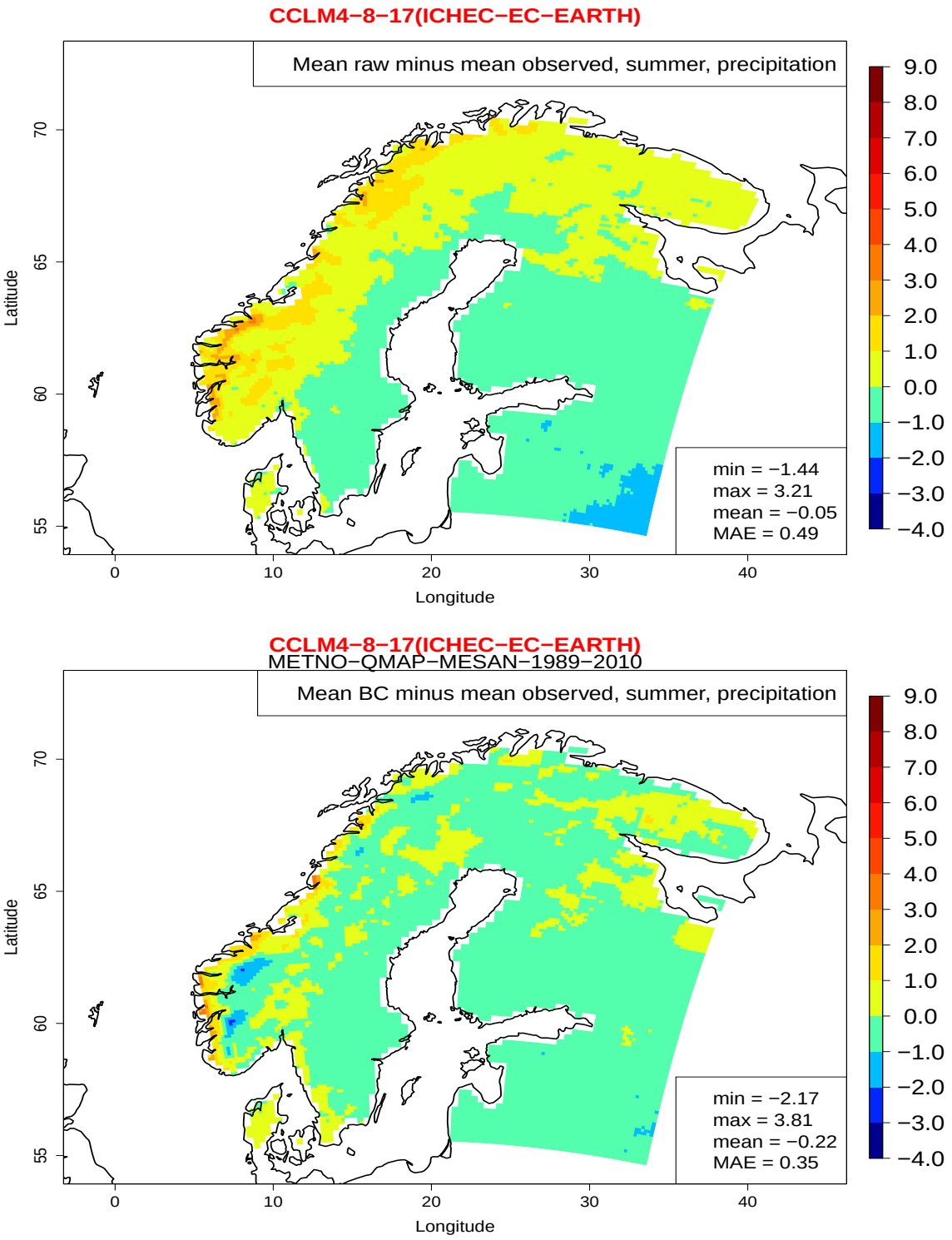


Figure 9. Difference between mean climate model precipitation and mean observed precipitation during summer from 1980 to 2005. The difference for the raw data from model 4 on top and the difference for bias correction ("BC") method 2 applied to model 4 at the bottom. MAE is the mean absolute error.

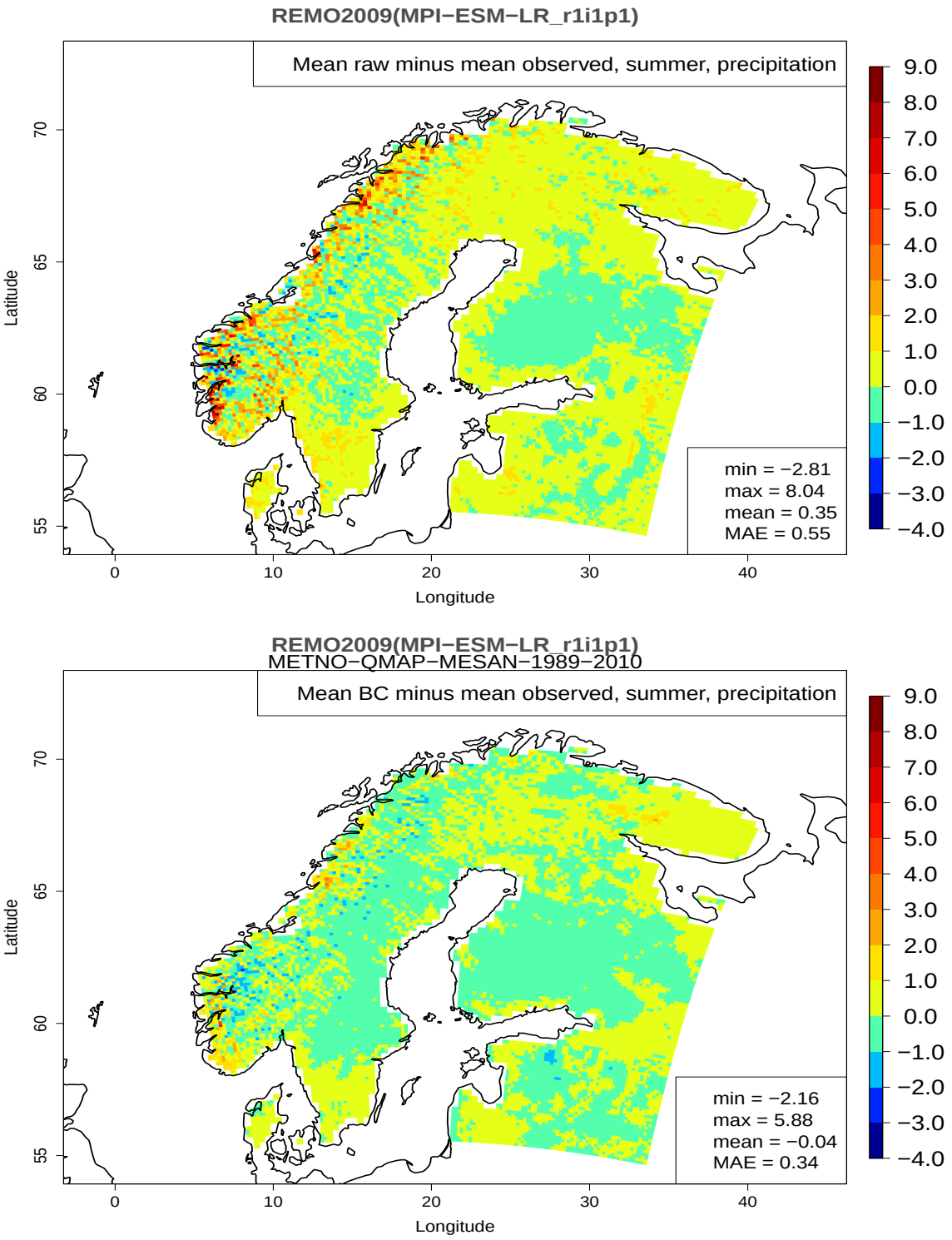


Figure 10. Difference between mean climate model precipitation and mean observed precipitation during summer from 1980 to 2005. The difference for the raw data from model 6 on top and the difference for bias correction ("BC") method 2 applied to model 6 at the bottom. MAE is the mean absolute error.

5.2 Drought

Plots of the number of droughts longer than seven days for all grid points reveal that all the raw models systematically project fewer long drought periods than we find in the E-OBS dataset. The exceptions are the areas where the models had problems with daily precipitation. This means that some of the models project too many drought periods along the coast of Norway during winter and some models project too many drought periods in the Balkans during summer.

Figures 11 and 12 show the mean drought IQD during summer and winter for model 3, the raw model on top and the bias corrected model at the bottom. More plots are given in Appendix E. In general, the areas with high IQD during summer are distributed all over Fennoscandia, and there does not seem to be any clear pattern to its size. During winter, the IQD is significantly smaller along the western coast of Norway (due to the lack of dry periods) and in the most eastern areas of Fennoscandia. From Appendix A we see that the same pattern holds for the observed droughts during summer and winter. This implies once again that the raw climate models have trouble projecting the precipitation in areas with more "extreme" weather conditions. The same patterns can be found in the IQD plots for the bias corrected models, although the IQD is much lower.

Mean drought IQD over all grid points is shown in Figure 13. Once again, method 1 clearly works best. The other three correction methods seem to achieve much the same results. However method 4 performs slightly better than methods 2 and 3. The difference between the mean IQD for the raw and the corrected climate models is slightly smaller for the drought periods than for the precipitation data with respect to the size of the 10% and 90% quantiles (see Figure 14). Therefore we might expect that the change in IQD per grid point for the corrected climate models is not as good as for precipitation.

However, the results are much the same as for precipitation, except for a few examples where a lot more of the grid points obtain a higher IQD after correction (see Figures 15 and 16, and Appendix F). Especially for model 4, the results after correction during summer are really bad when the goal is to improve every single grid point. This might happen because the raw summer IQD for model 4 is quite low. We also see that a corrected version of model 6 obtains better results than those of model 4, even though the raw IQD is much higher for model 6. Therefore it does not seem like the IQD of the raw models matters very much for the corrected results.

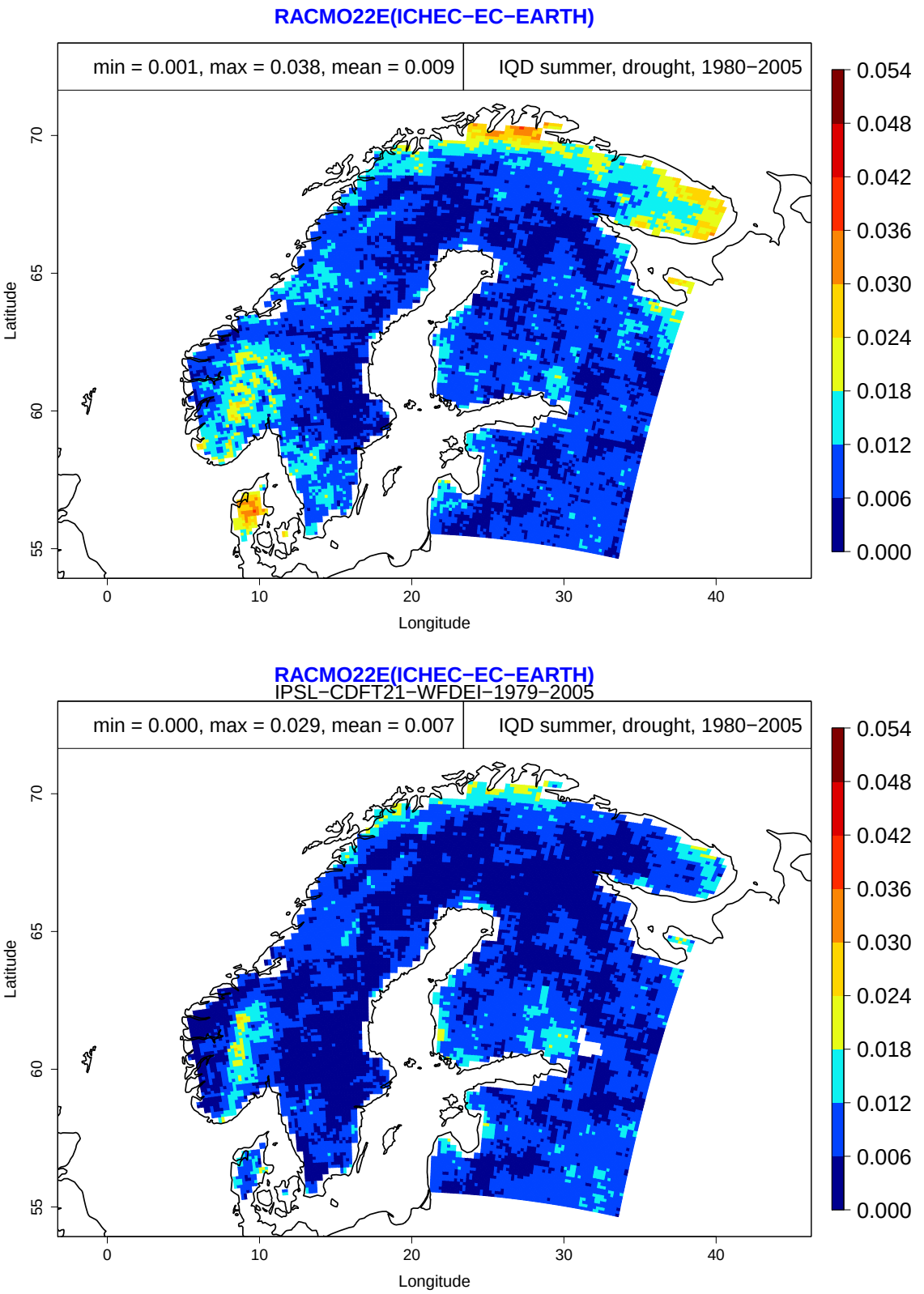


Figure 11. The mean drought IQD from 1980 to 2005 for Fennoscandia during summer. The upper figure shows the IQD for model 3 with no bias correction. The lower figure shows the IQD for model 3 after applying bias correction method 4.

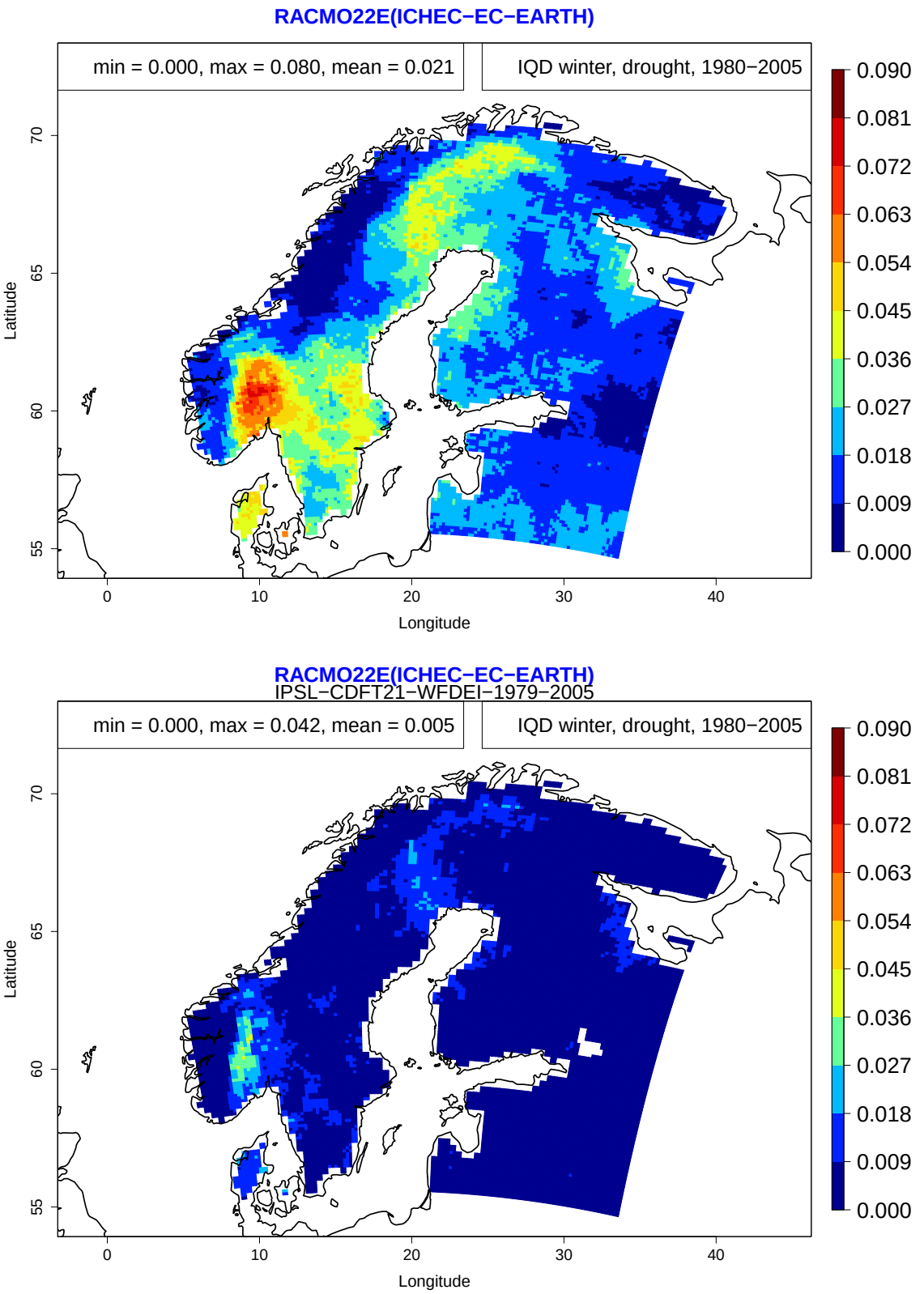


Figure 12. The mean drought IQD from 1980 to 2005 for Fennoscandia during winter. The upper figure shows the IQD for model 3 with no bias correction. The lower figure shows the IQD for model 3 after applying bias correction method 4.

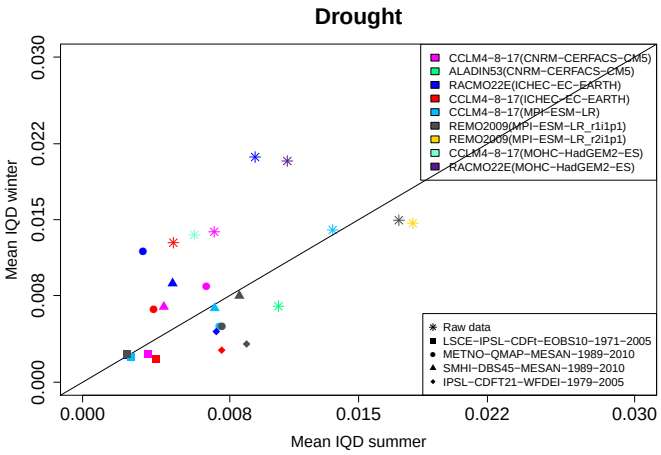


Figure 13. Scatter plot with the mean drought IQD for all climate models and bias correction methods. Summer results are plotted along the x-axis and winter results are plotted along the y-axis. Colour indicates climate model, while shape indicates raw data and bias correction method.

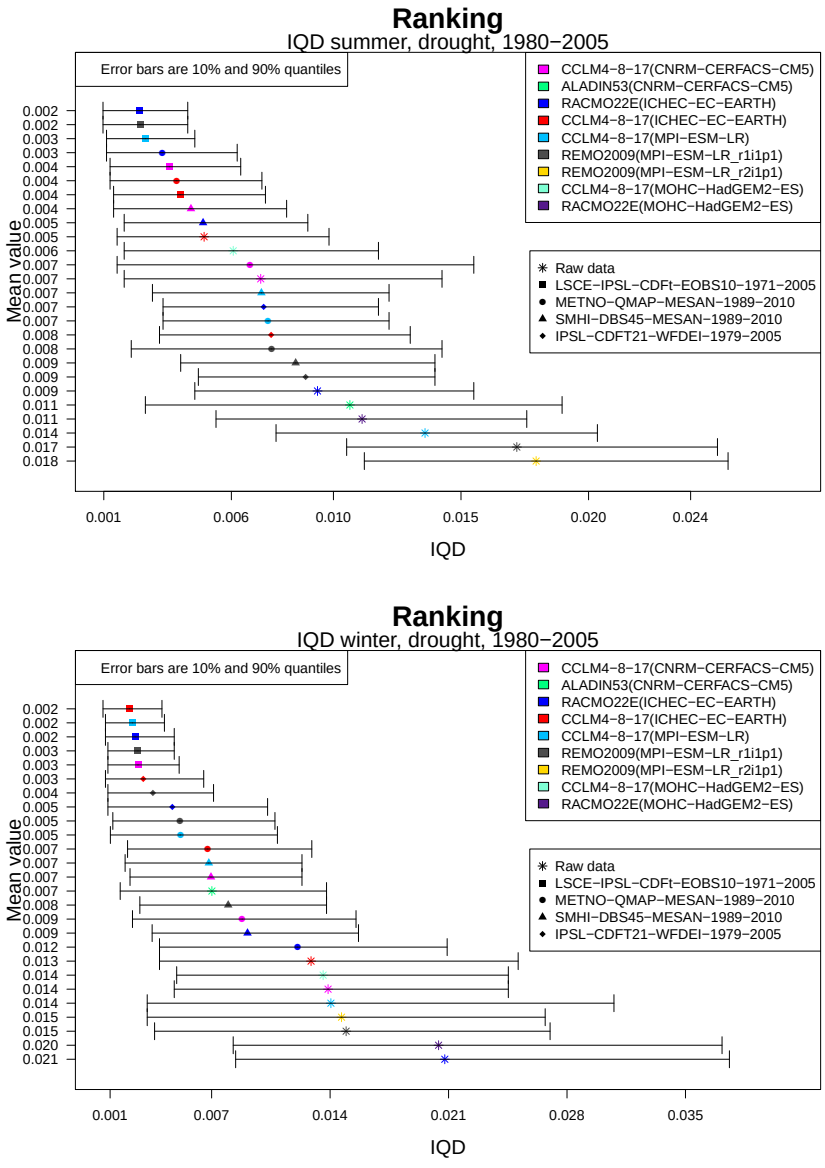


Figure 14. Ranking of the mean drought IQD for all climate models and bias correction methods. Summer is given on top and winter at the bottom. The bars indicate the 10% and the 90% quantiles for IQD values at individual grid points.

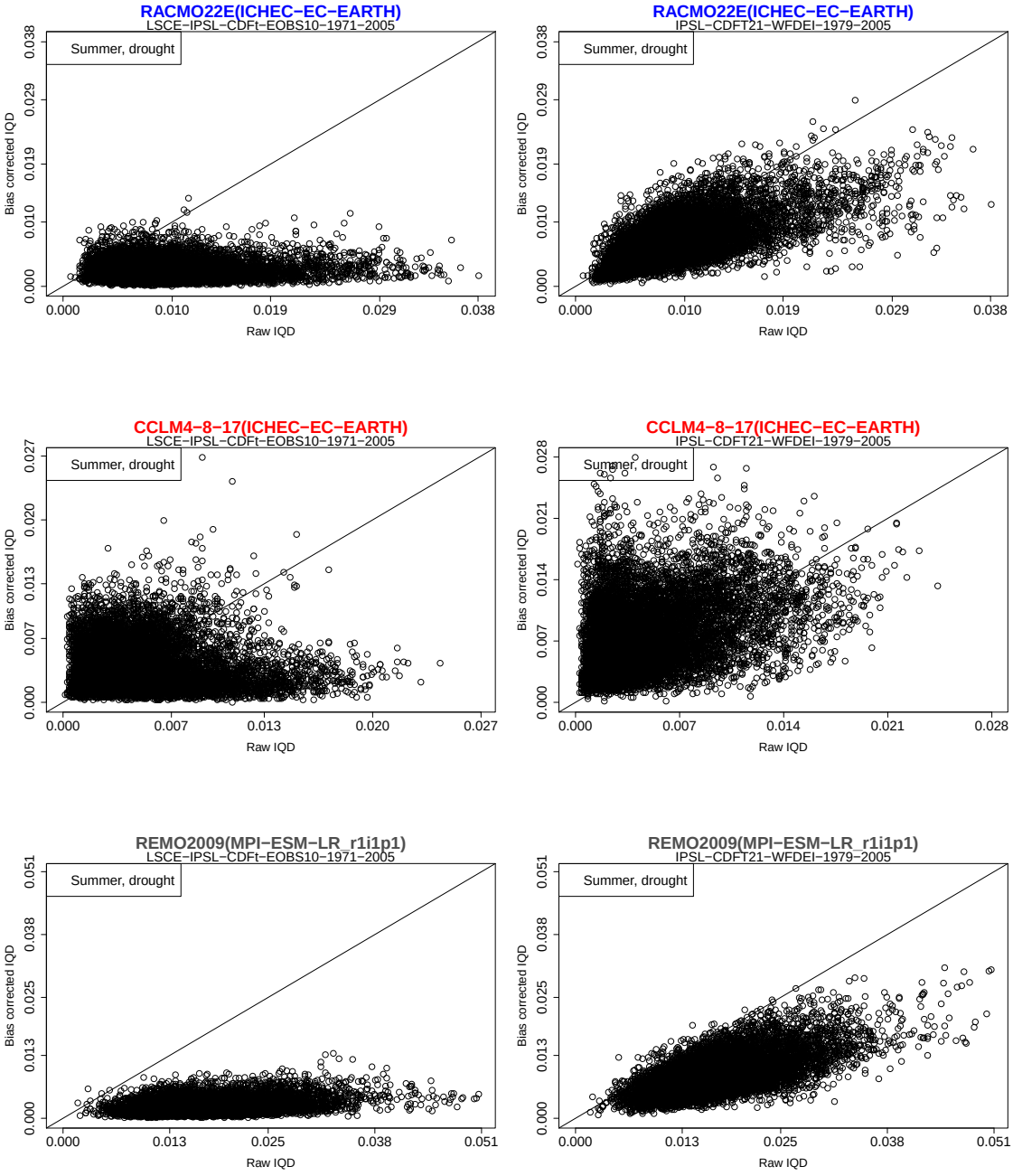


Figure 15. Scatter plots displaying drought IQD of raw data and bias corrected data for the models 3, 4 and 6 during summer. The figures on the left show bias correction method 1, while the figures on the right show bias correction method 4.

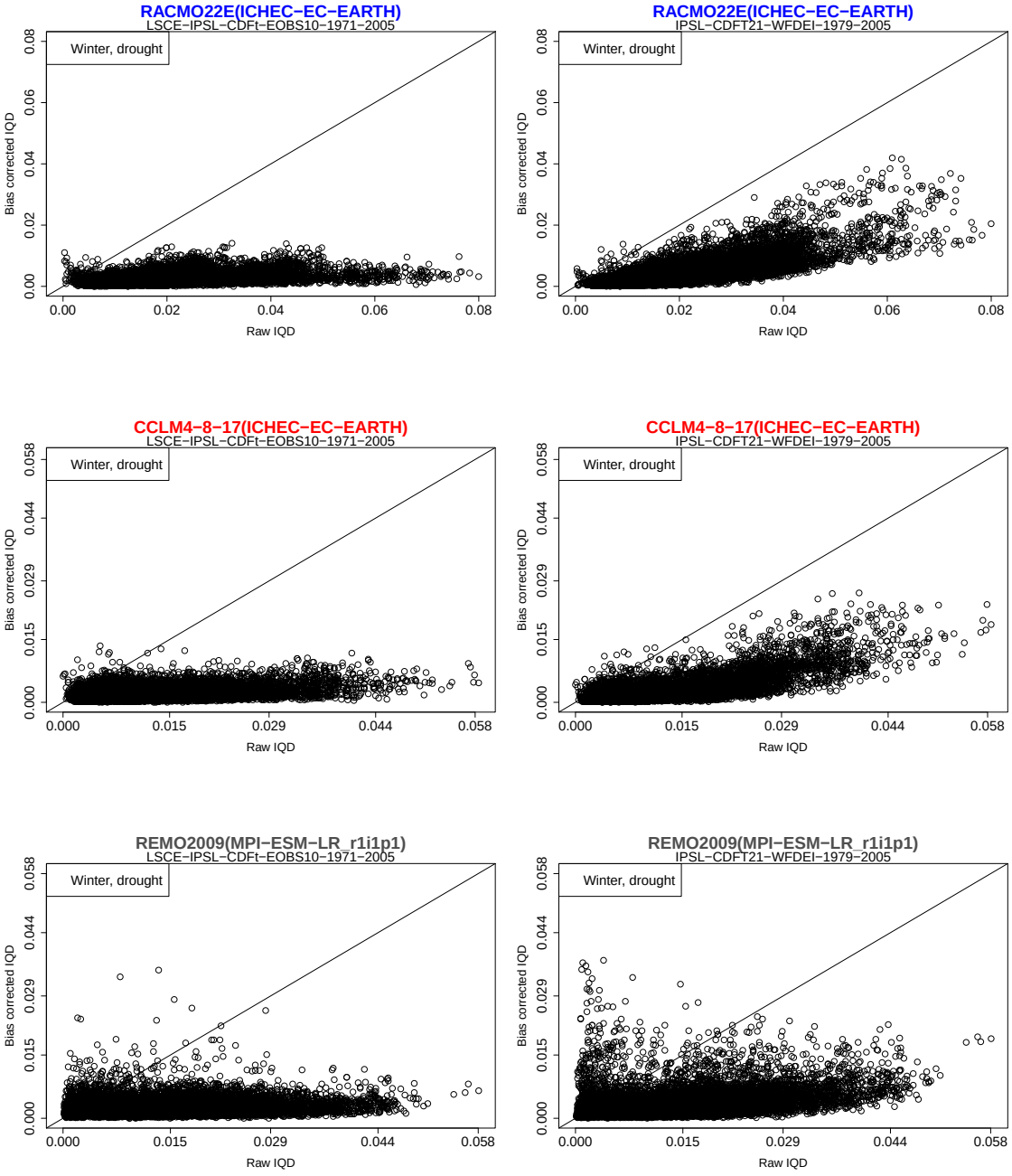


Figure 16. Scatter plots displaying drought IQD of raw data and bias corrected data for the models 3, 4 and 6 during winter. The figures on the left show bias correction method 1, while the figures on the right show bias correction method 4.

6 Conclusion

We have seen that the patterns in IQD for the raw models for both daily precipitation distributions and drought periods are quite similar to the patterns in the E-OBS dataset in areas with larger amounts of precipitation and many drought periods. This indicates that the raw models are having problems with projections whenever they become "more extreme". In turn, the most successful bias correction methods seem to be the ones that correct these extreme areas best.

Bias correction method 1 obtains the best results for all our testing. It has the lowest mean IQD, and the 10% and 90% quantiles are not too far away from the mean. Almost every single coordinate obtains a lower IQD score after correcting the raw models. However, we saw that the corrected models are highly dependent on what dataset that is used for calibration. Therefore one would expect method 1 to achieve good results, as it was based on the E-OBS dataset, which we are testing our data against. We have also seen that there was a difference between method 1 and method 4, which achieved worse results. The technique for these two methods is almost or exactly the same, but the calibration datasets are different. This indicates that the techniques behind method 1 might not work as well as it looks from our data. It could, however, also mean that the small changes which may have been done from method 1 to method 4 made a big difference. Further testing with another dataset of observations is therefore recommended.

Bias correction method 3 also performs quite well in our tests. It has some grid points where it performs significantly worse after correction during summer for the CCLM4-8-17 models, but otherwise there is much improvement. All the quantiles for the IQD are also close to the mean. Except for the coast of Norway, there are three areas where the IQD is higher than what is usual. These three areas are also found for method 2 and therefore most likely come from the MESAN dataset. It is most likely that method 3 performs better than method 2. Both are calibrated using the same dataset, thus making it easy to compare the two, and in every test method 3 obtained better results than method 2.

The differences between method 3 and method 4 are very small. Method 3 does slightly better for precipitation while method 4 does slightly better for the number of and lengths of drought periods longer than seven days. Therefore it is hard to conclude which of these bias correction methods is the most promising. Further testing of the correction models will hopefully reveal more.

References

Gudmundsson, L., Bremnes, J. B., Haugen, J. E., and Engen-Skaugen, T. (2012). Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations - a comparison of methods. *Hydrology and Earth System Sciences*, 16(9):3383–3390. 9

Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., and New, M. (2008). A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006. *Journal of Geophysical Research: Atmospheres*, 113:D20119. doi:10.1029/2008JD010201. 6

Murphy, A. H. (1993). What is a good forecast? An essay on the nature of goodness in weather forecasting. *Weather and Forecasting*, 8(2):281–293. 8

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Available from: https://www.R-project.org/. 9

Thorarinsdottir, T. L., Gneiting, T., and Gissibl, N. (2013). Using proper divergence functions to evaluate climate models. *SIAM/ASA Journal on Uncertainty Quantification*, 1(1):522–534. 9

Vrac, M., Noël, T., and Vautard, R. (2016). Bias correction of precipitation through singularity stochastic removal: Because occurrences matter. *Journal of Geophysical Research: Atmospheres*, 121(10):5237–5258. 9

Yang, W., Andréasson, J., Graham, L. P., Olsson, J., Rosberg, J., and Wetterhall, F. (2010). Distribution-based scaling to improve usability of regional climate model projections for hydrological climate change impacts studies. *Hydrology Research*, 41(3-4):211–229. 9

A Data from E-OBS

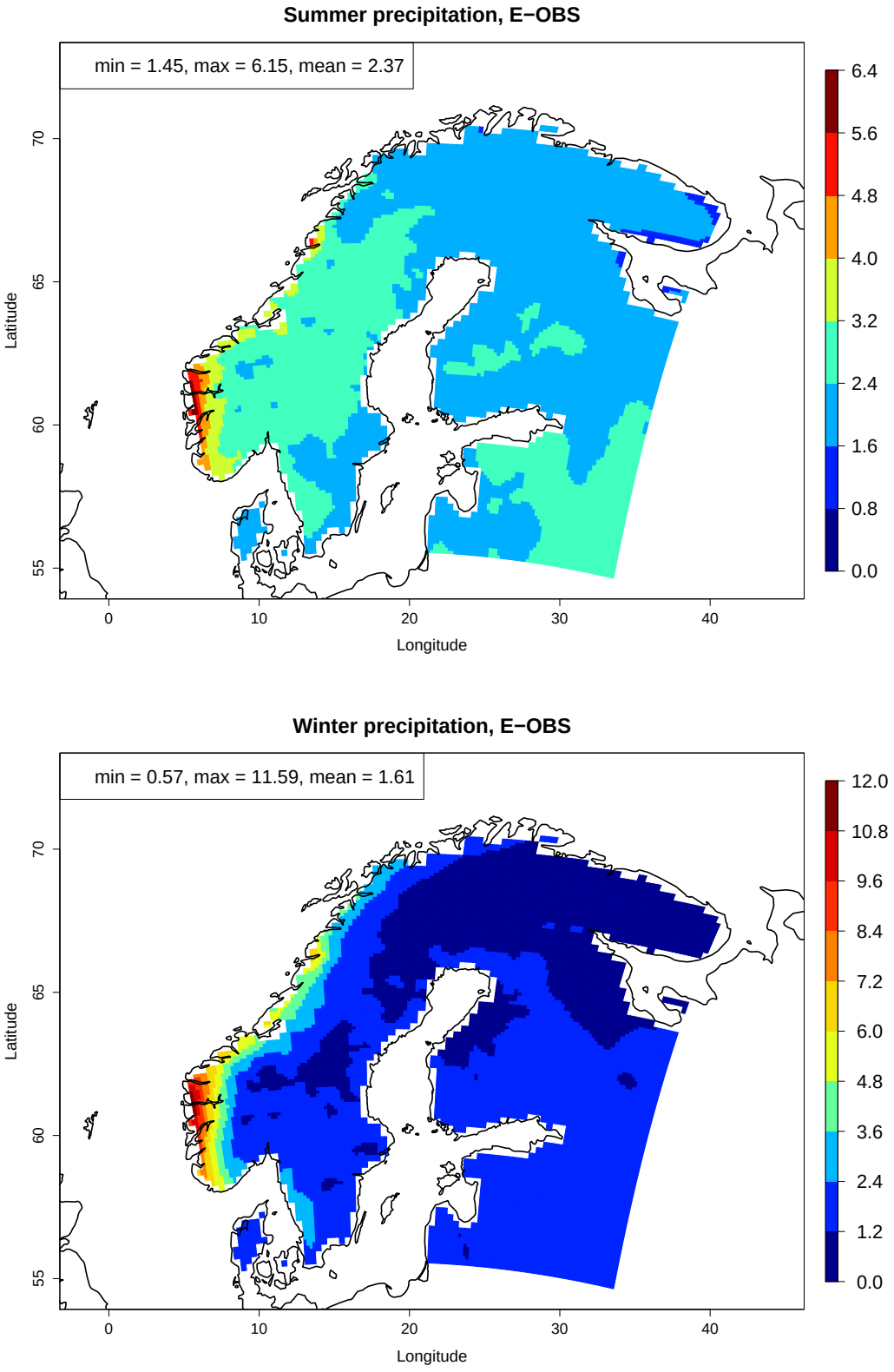
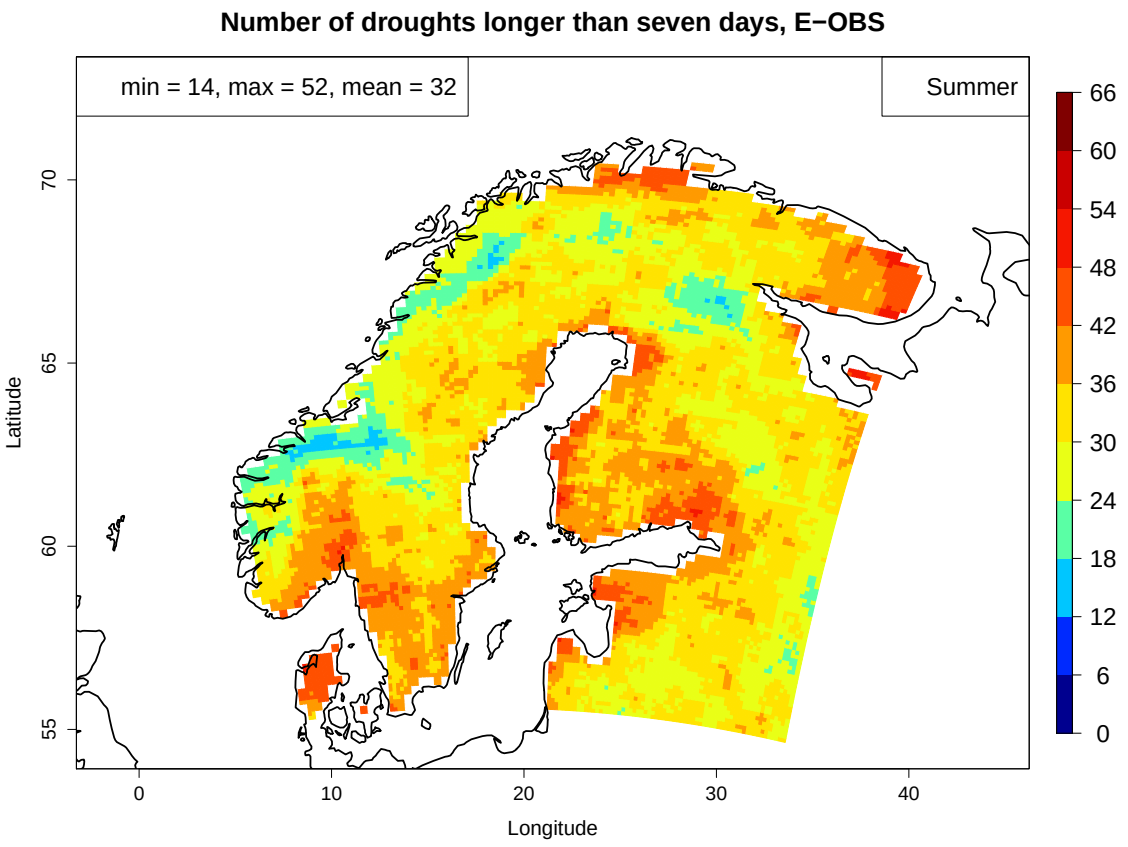


Figure A.1. Mean observed precipitation from E-OBS, 1980-2005, for summer (on top) and winter (at the bottom).



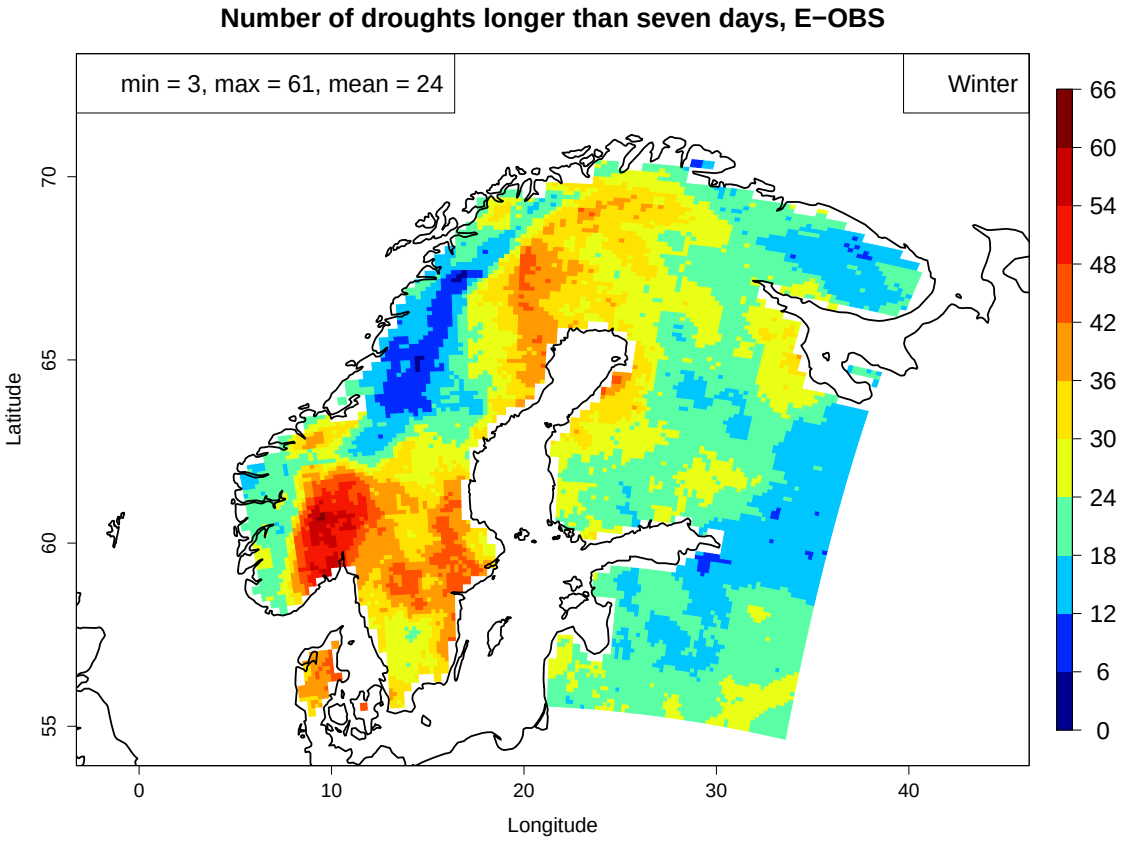


Figure A.2. Mean observed number of drought periods longer than seven days from E-OBS, 1980-2005, for summer (on top) and winter (at the bottom).

B Raw vs. bias corrected precipitation IQD

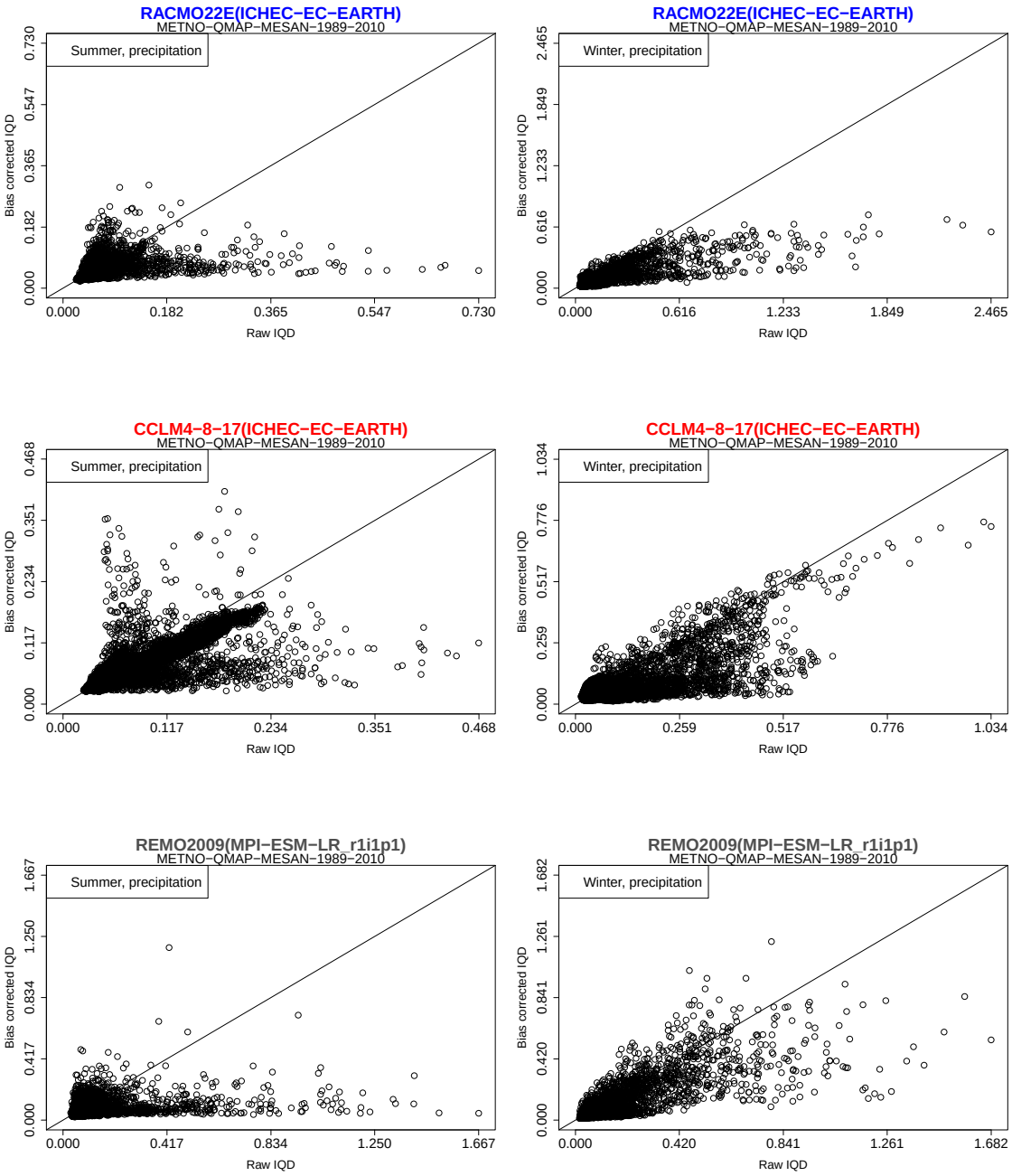


Figure B.3. Scatter plots displaying precipitation IQD of raw data and bias corrected data from method 2 for the models 3, 4 and 6 during summer (on the left) and winter (on the right).

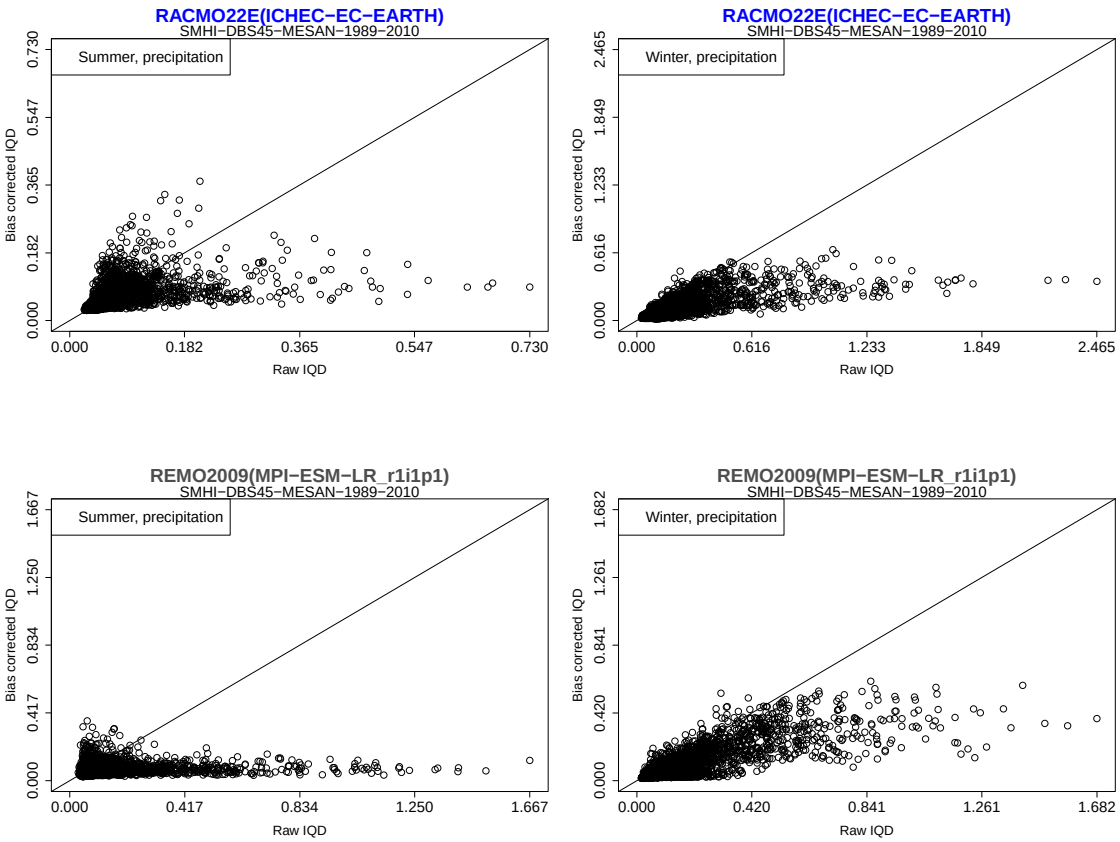


Figure B.4. Scatter plots displaying precipitation IQD of raw data and bias corrected data from method 3 for the models 3 and 6 during summer (on the left) and winter (on the right).

C Mean precipitation IQD

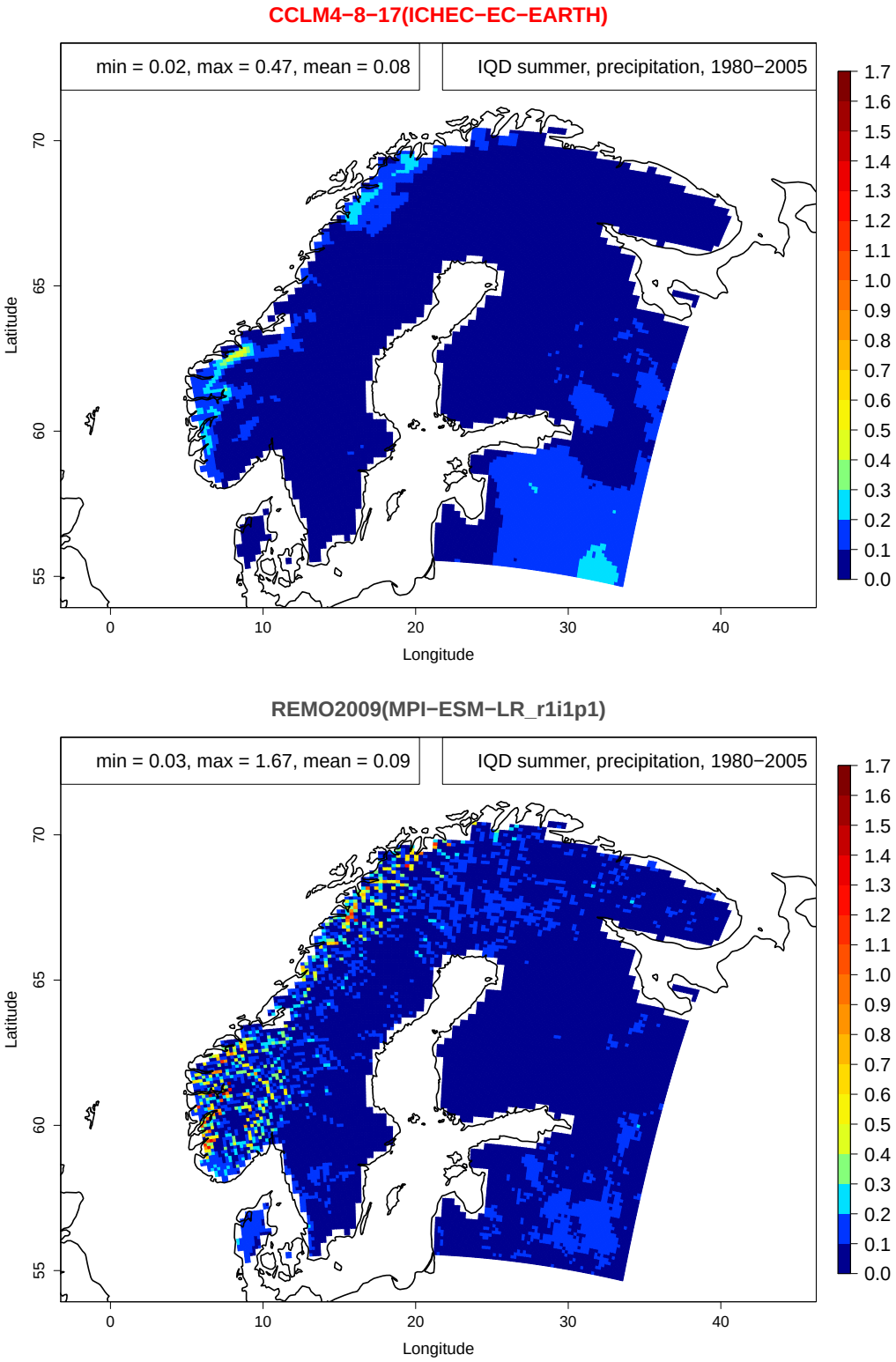


Figure C.5. Mean precipitation IQD for the models 4 and 6 during summer.

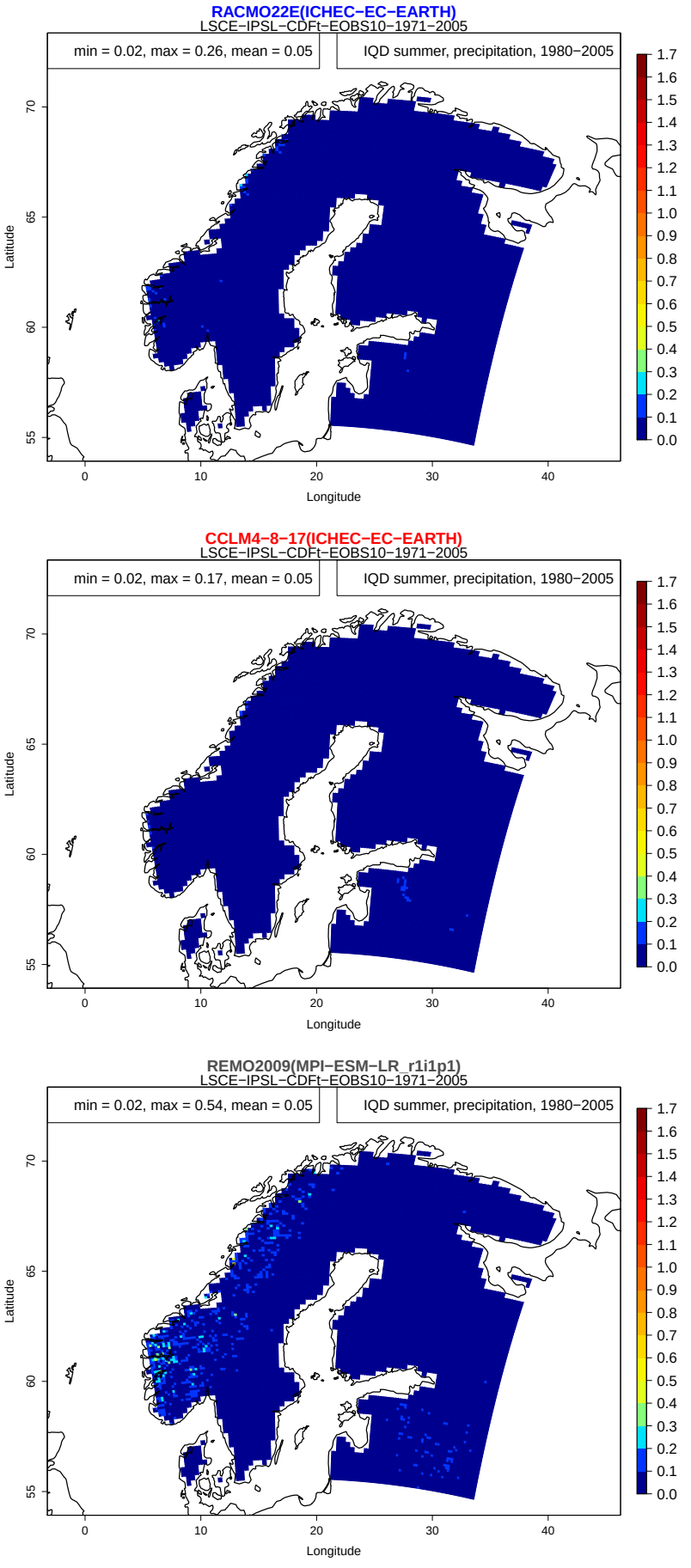


Figure C.6. Mean precipitation IQD for the models 3, 4 and 6 during summer after applying bias correction method 1.

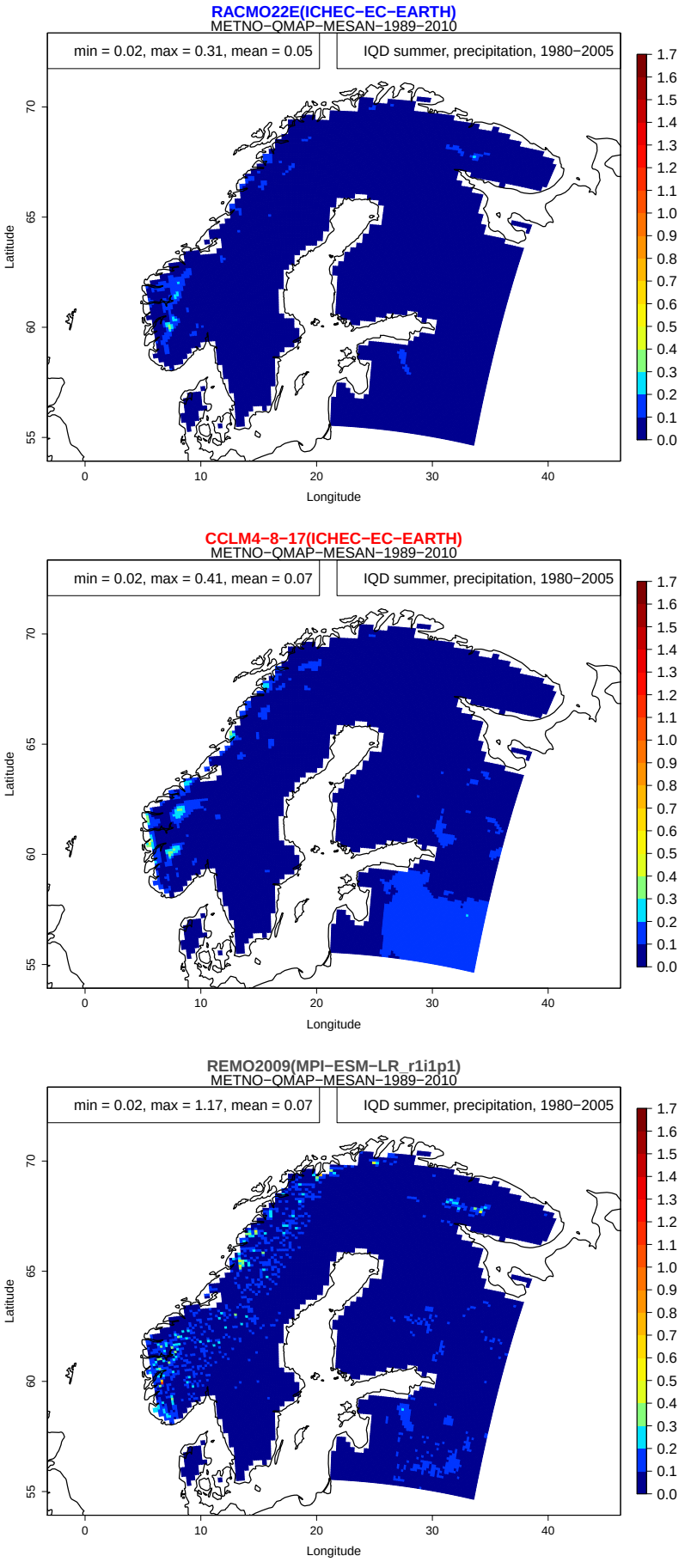


Figure C.7. Mean precipitation IQD for the models 3, 4 and 6 during summer after applying bias correction method 2.

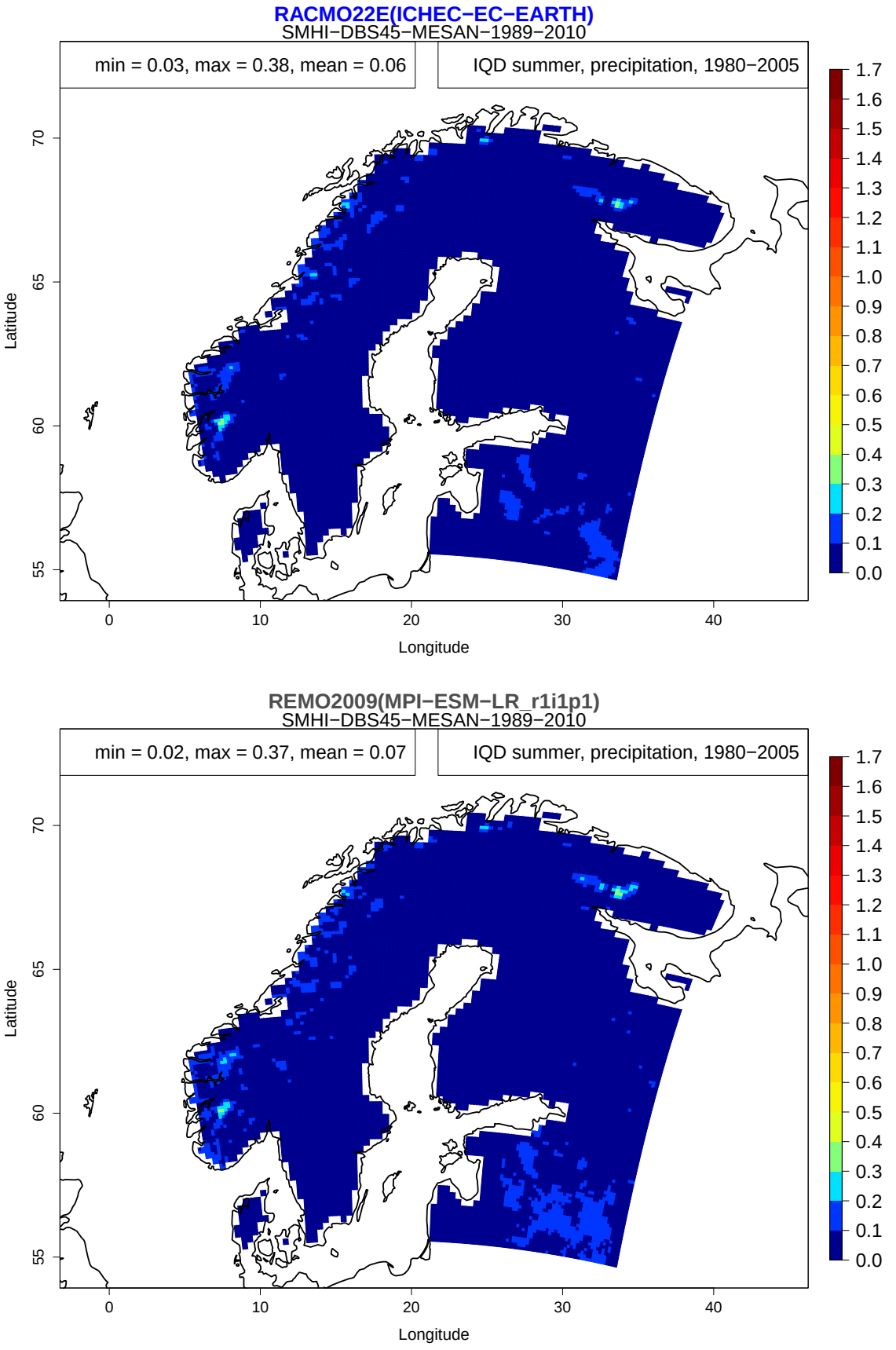


Figure C.8. Mean precipitation IQD for the models 3 and 6 during summer after applying bias correction method 3.

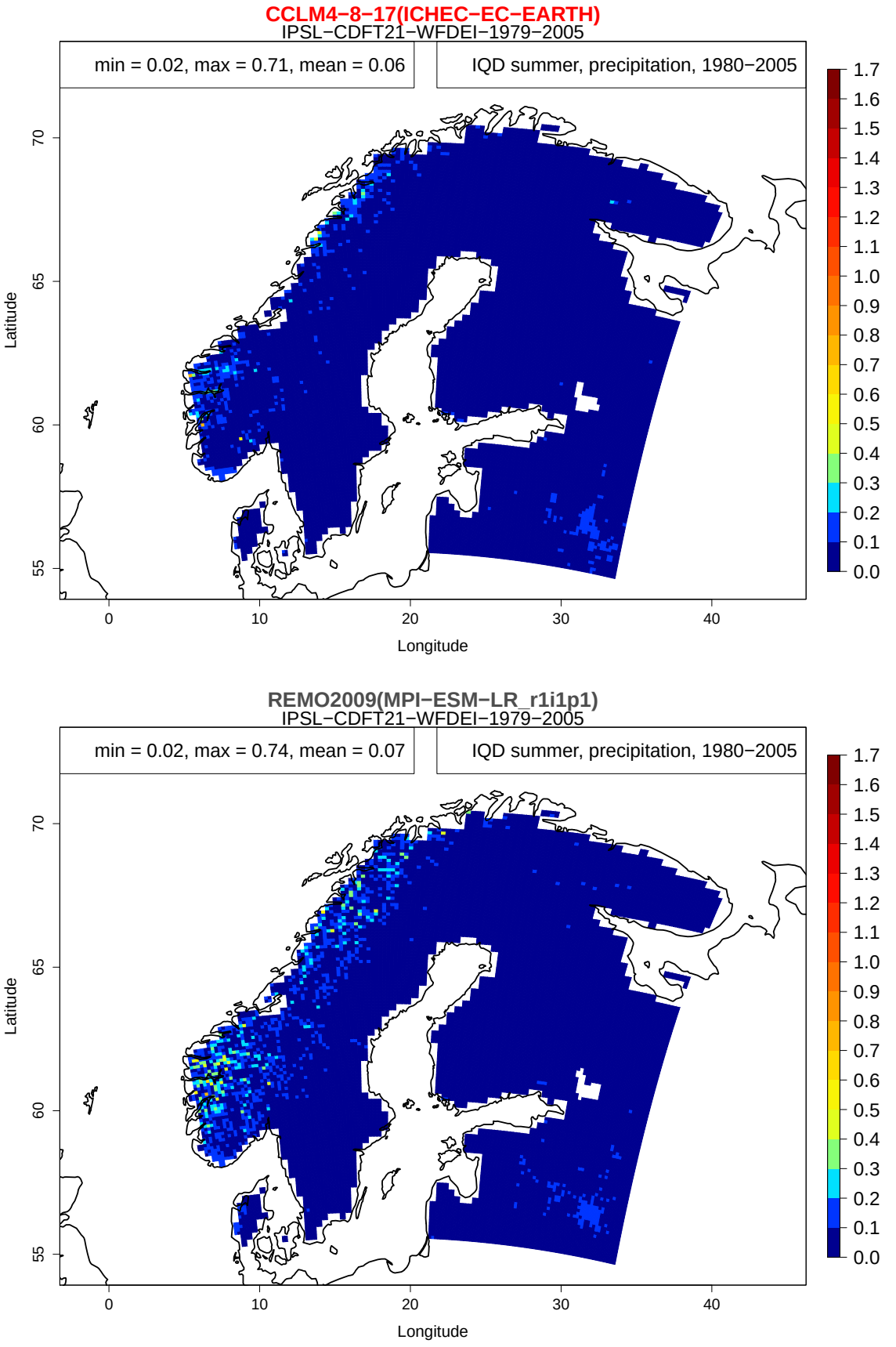


Figure C.9. Mean precipitation IQD for the models 4 and 6 during summer after applying bias correction method 4.

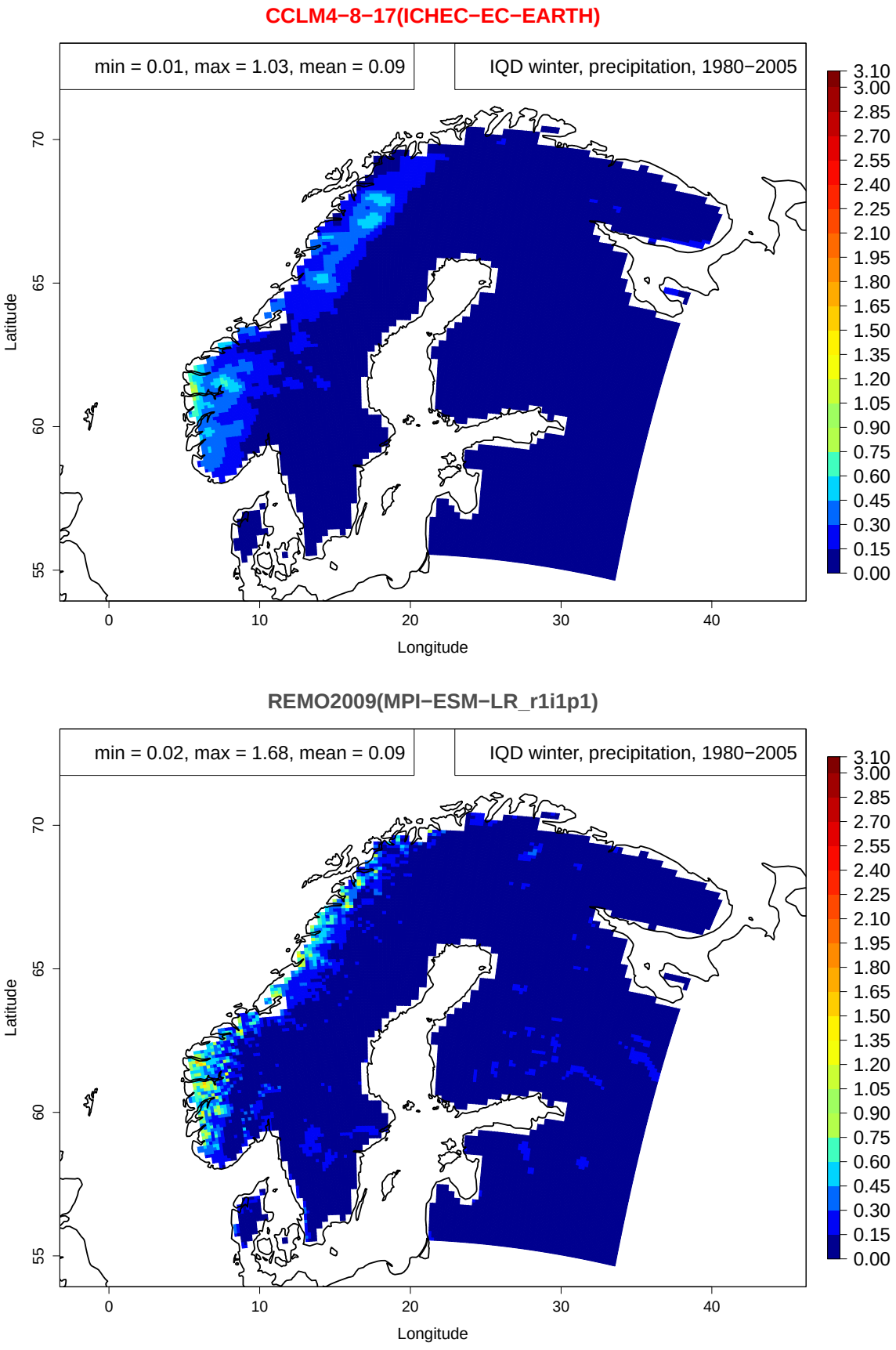


Figure C.10. Mean precipitation IQD for the models 4 and 6 during winter.

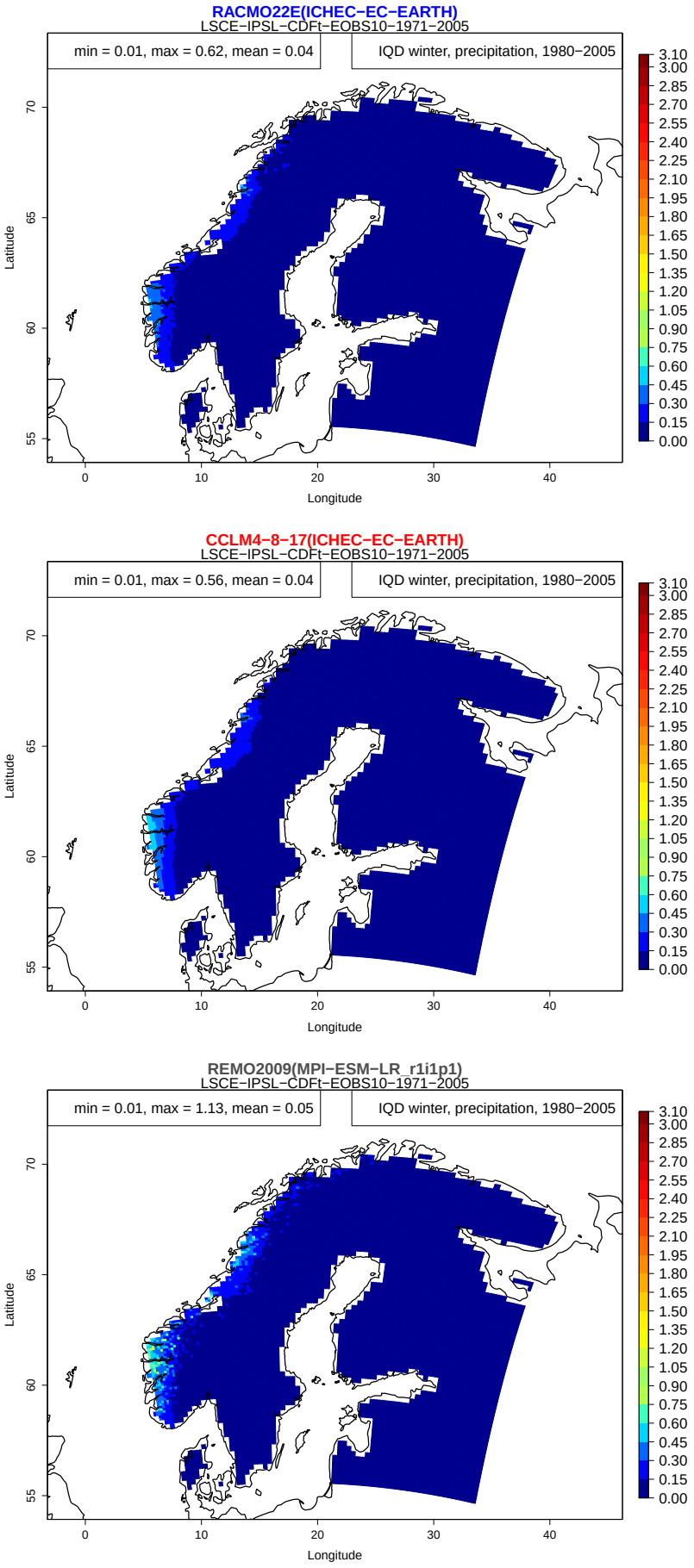


Figure C.11. Mean precipitation IQD for the models 3, 4 and 6 during winter after applying bias correction method 1.

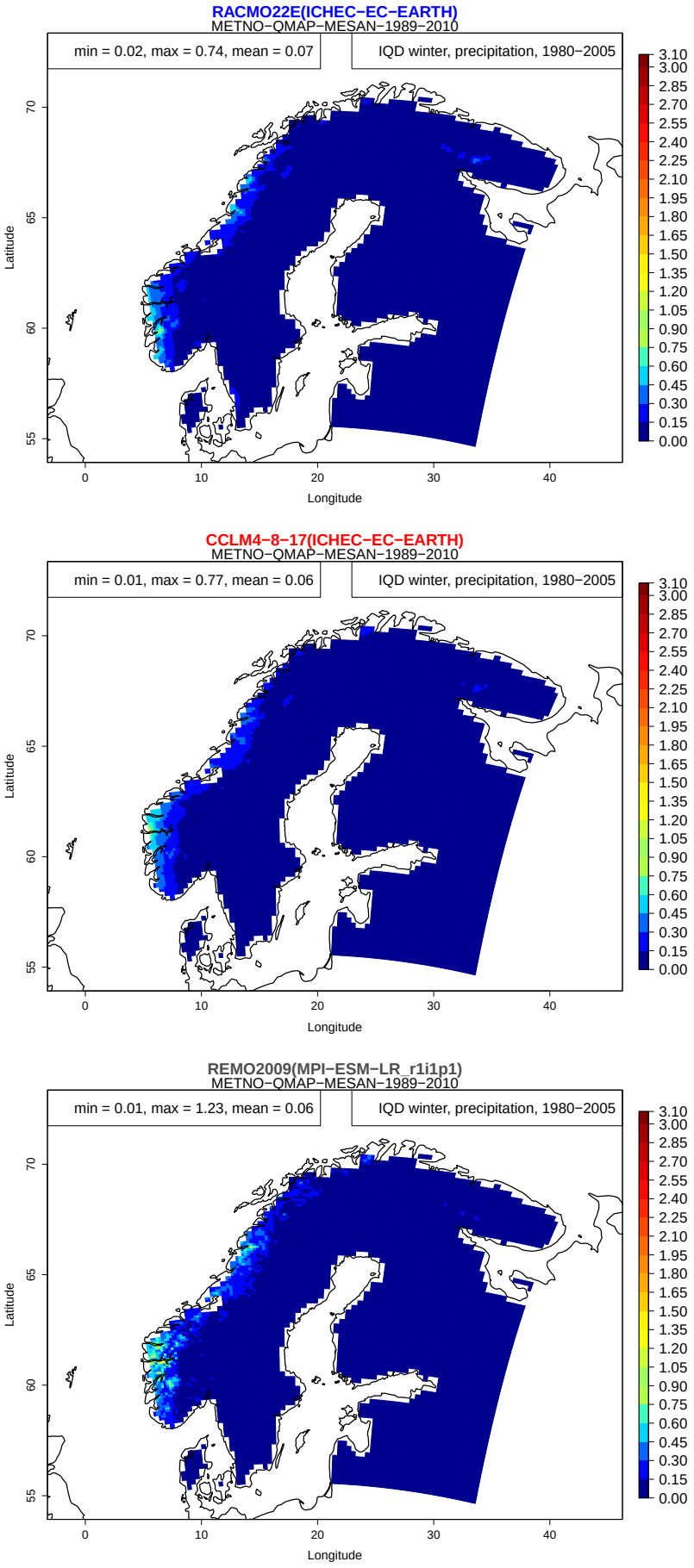


Figure C.12. Mean precipitation IQD for the models 3, 4 and 6 during winter after applying bias correction method 2.

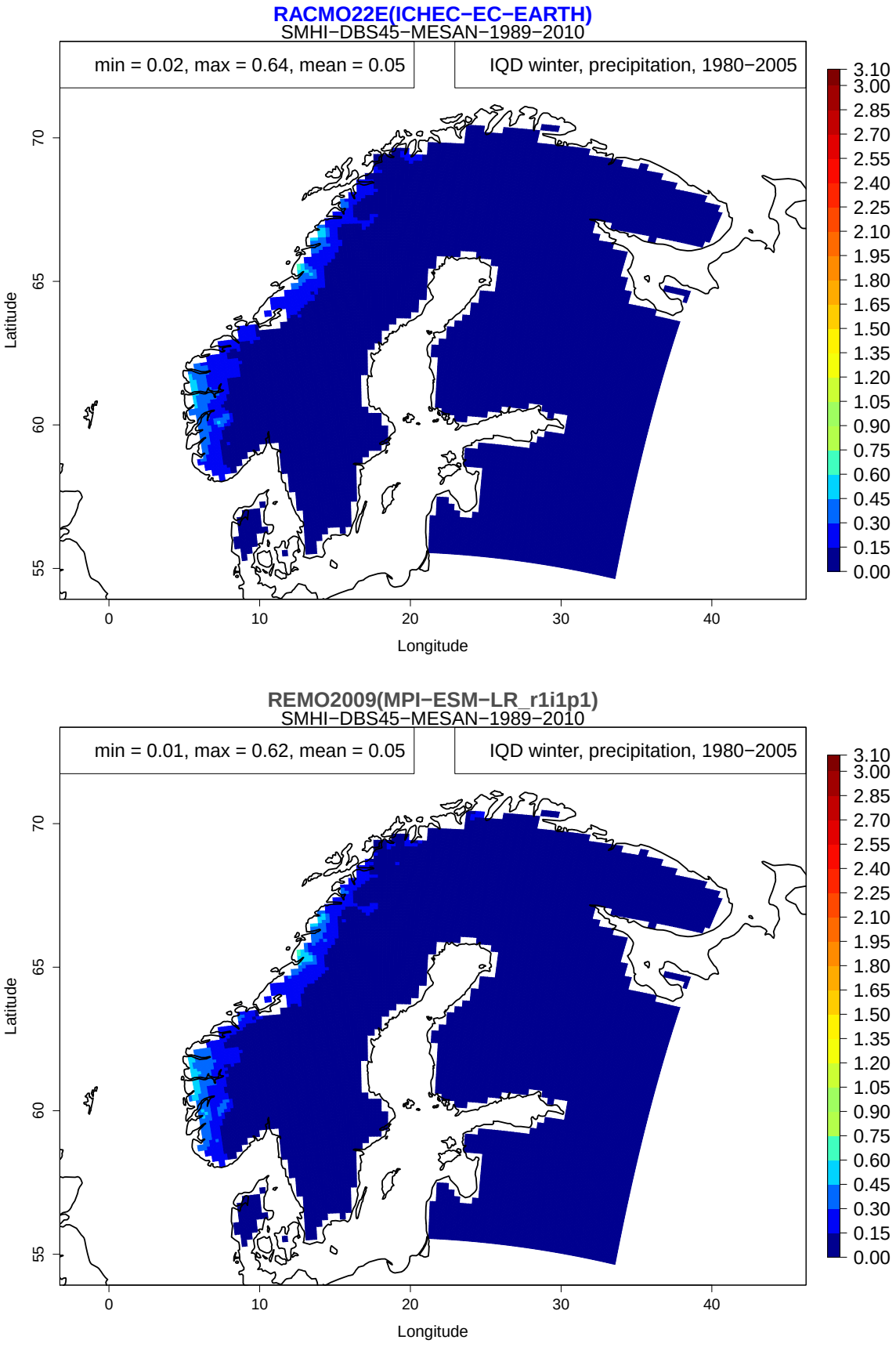


Figure C.13. Mean precipitation IQD for the models 3 and 6 during winter after applying bias correction method 3.

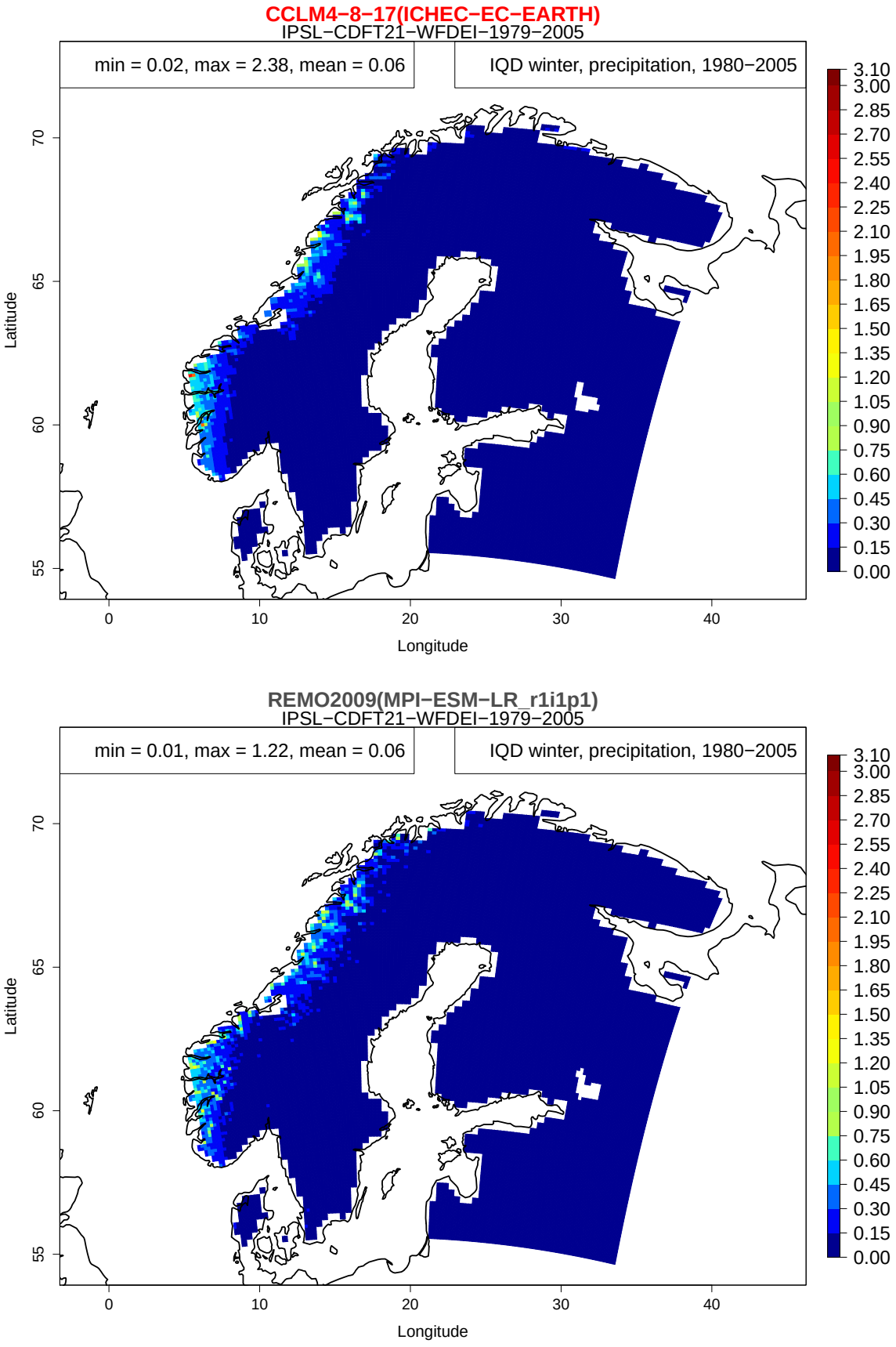


Figure C.14. Mean precipitation IQD for the models 4 and 6 during winter after applying bias correction method 4.

D Bias corrected minus observed precipitation

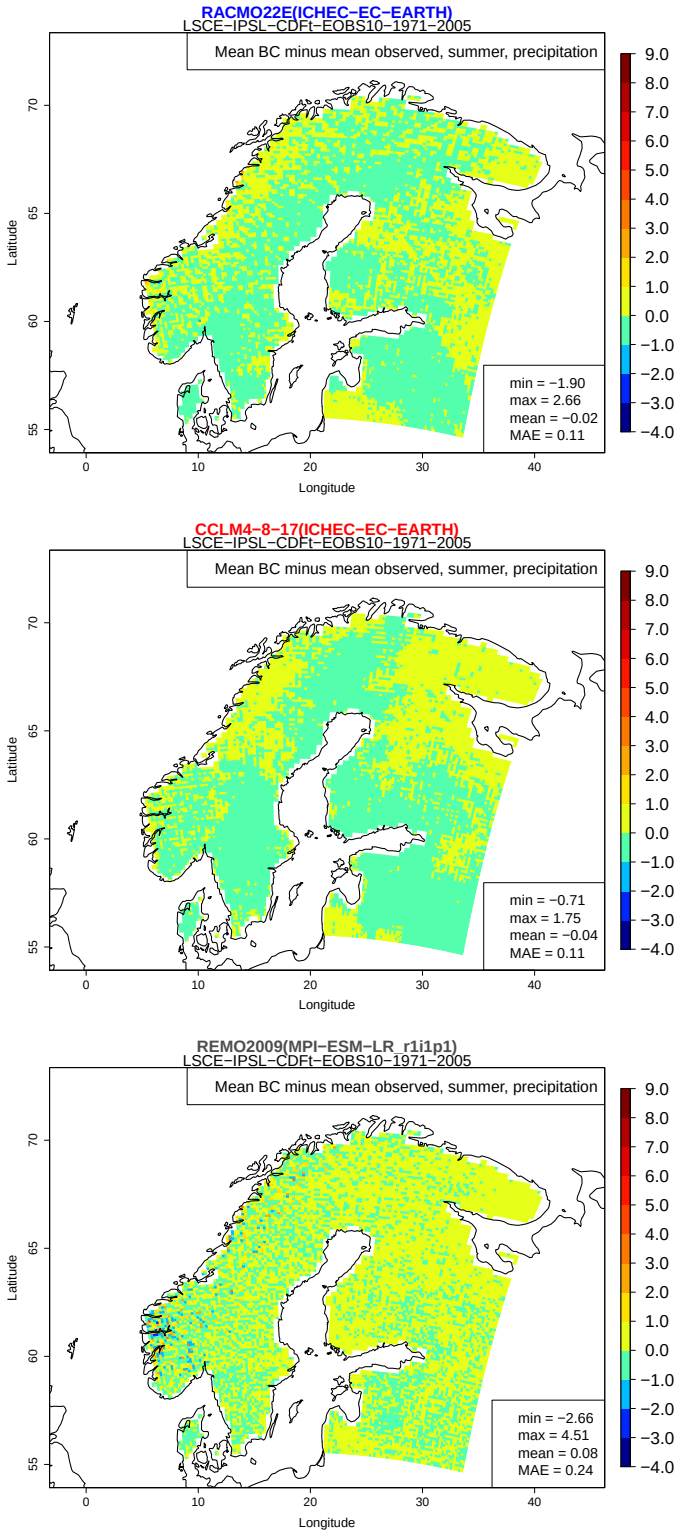


Figure D.15. Difference between mean bias corrected ("BC") precipitation from the models 3, 4 and 6 and mean observed precipitation during summer. Bias correction method 1. MAE is the mean absolute error.

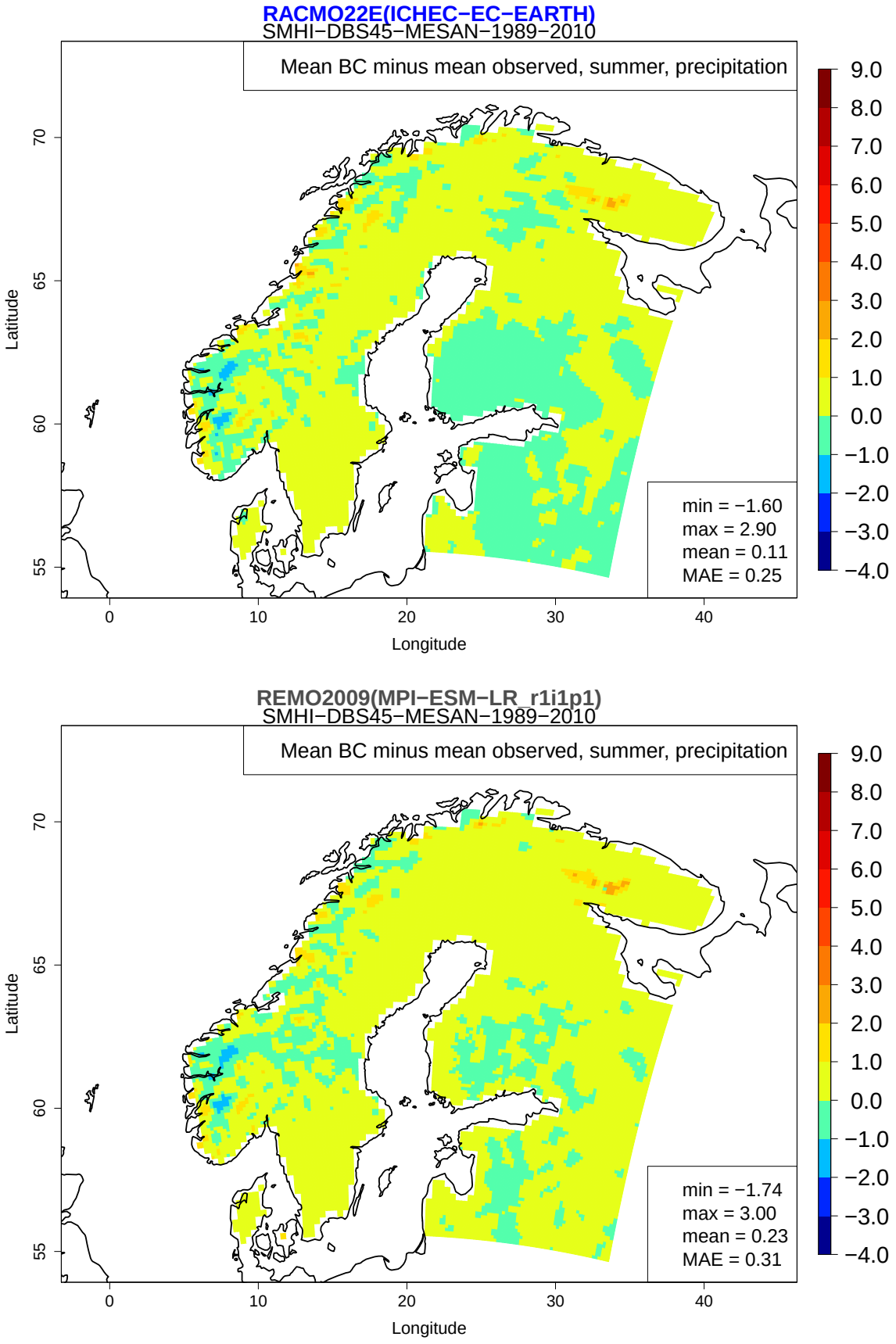


Figure D.16. Difference between mean bias corrected ("BC") precipitation from the models 3 and 6 and mean observed precipitation during summer. Bias correction method 3. MAE is the mean absolute error.

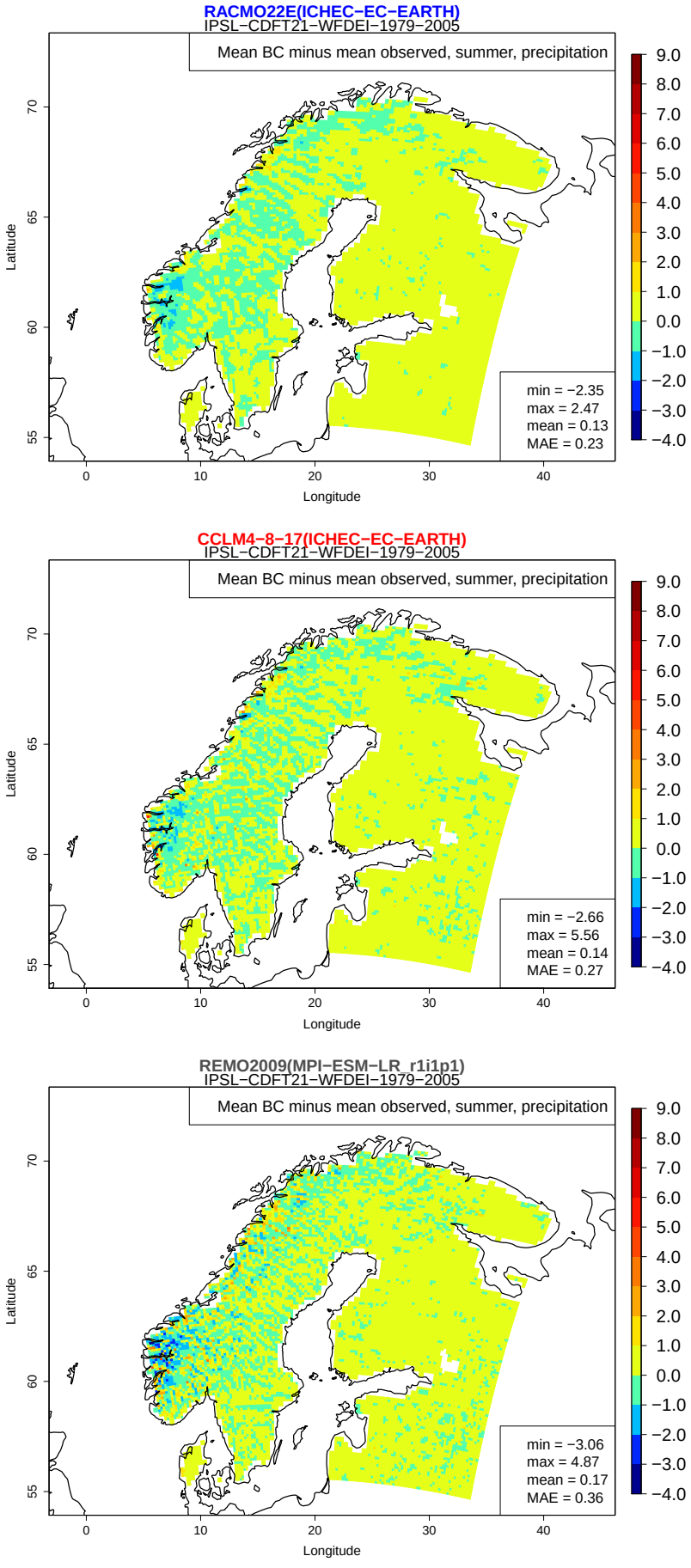


Figure D.17. Difference between mean bias corrected ("BC") precipitation from the models 3, 4 and 6 and mean observed precipitation during summer. Bias correction method 4. MAE is the mean absolute error.

E Mean drought IQD

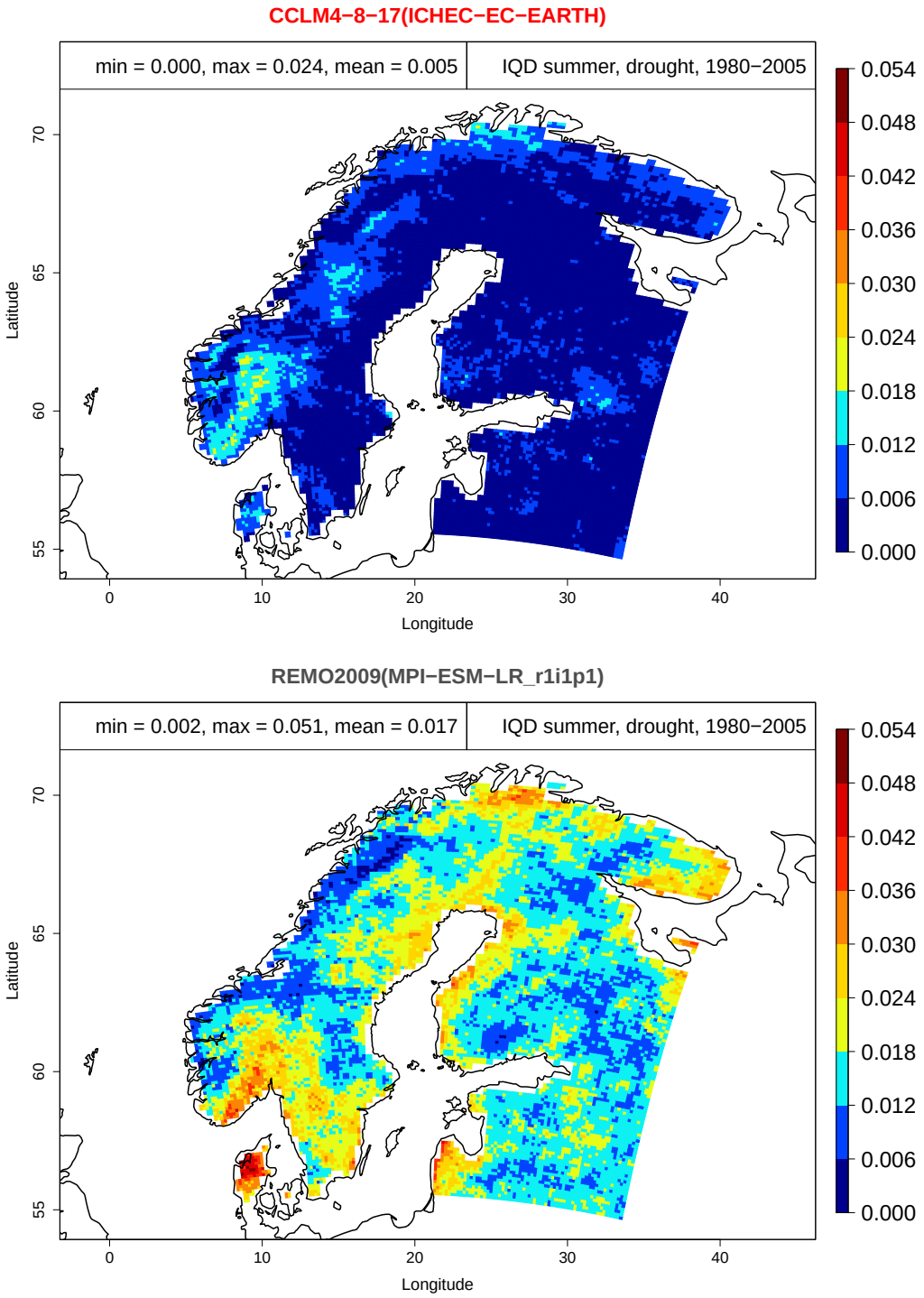


Figure E.18. Mean drought IQD for the models 4 and 6 during summer.

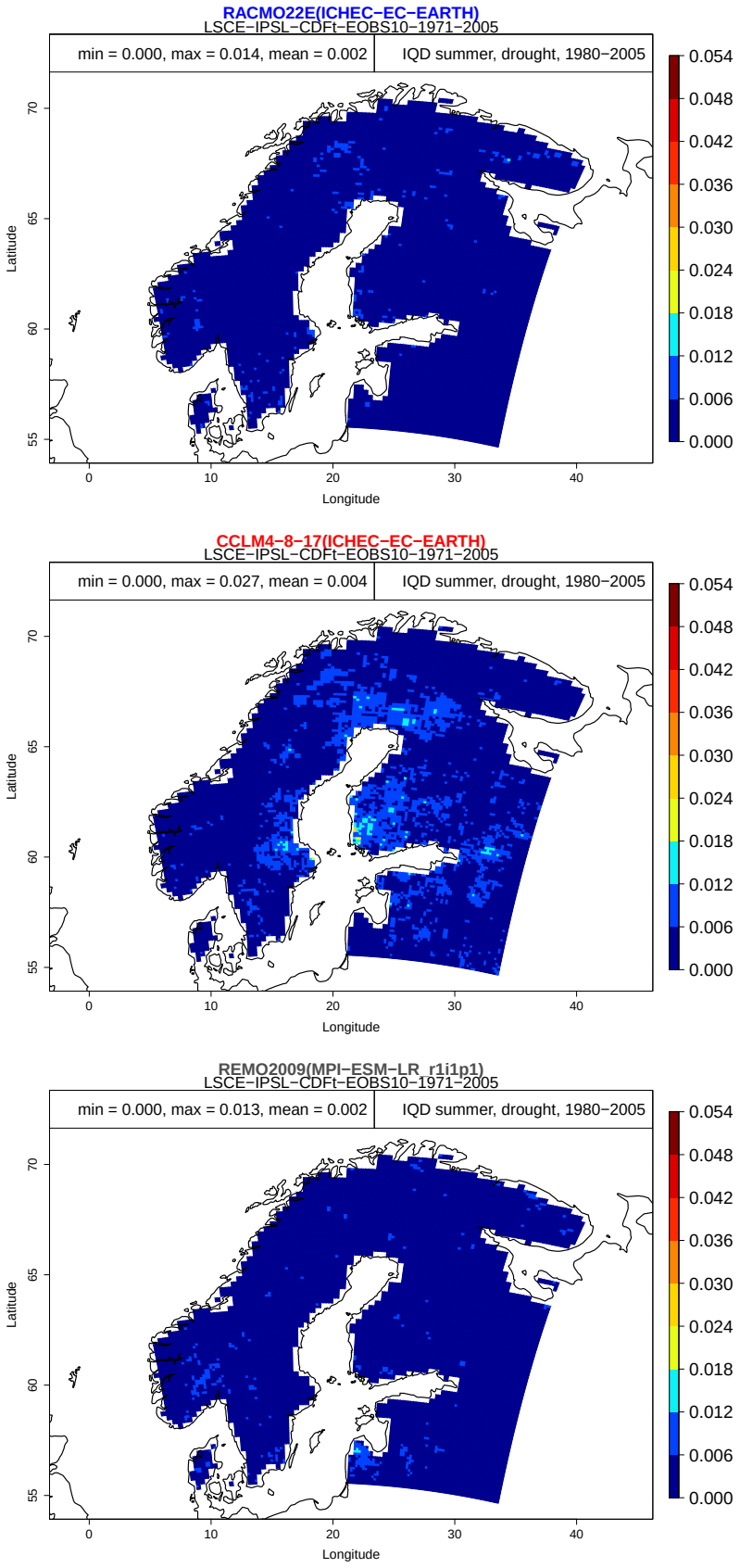


Figure E.19. Mean drought IQD for the models 3, 4 and 6 during summer after applying bias correction method 1.

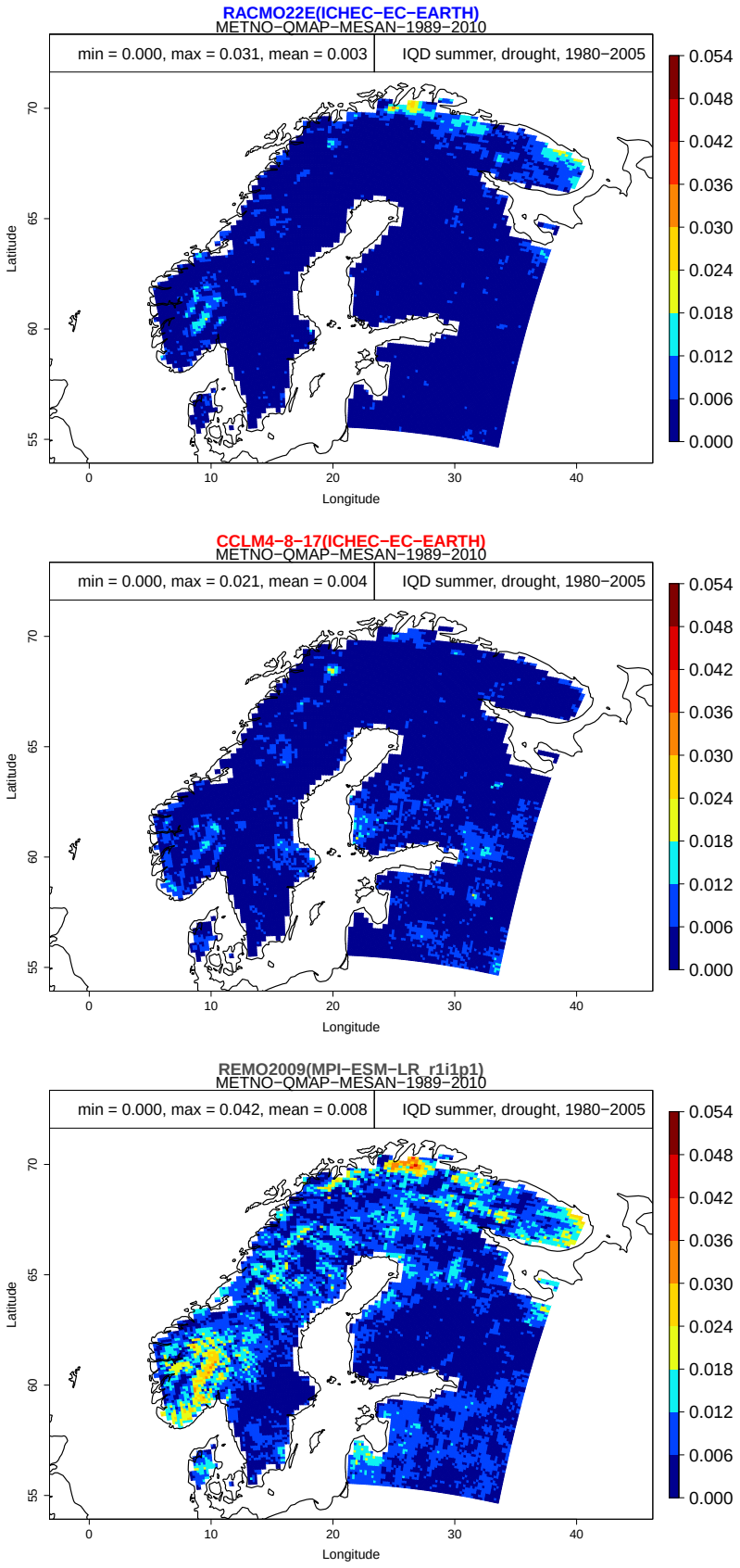


Figure E.20. Mean drought IQD for the models 3, 4 and 6 during summer after applying bias correction method 2.

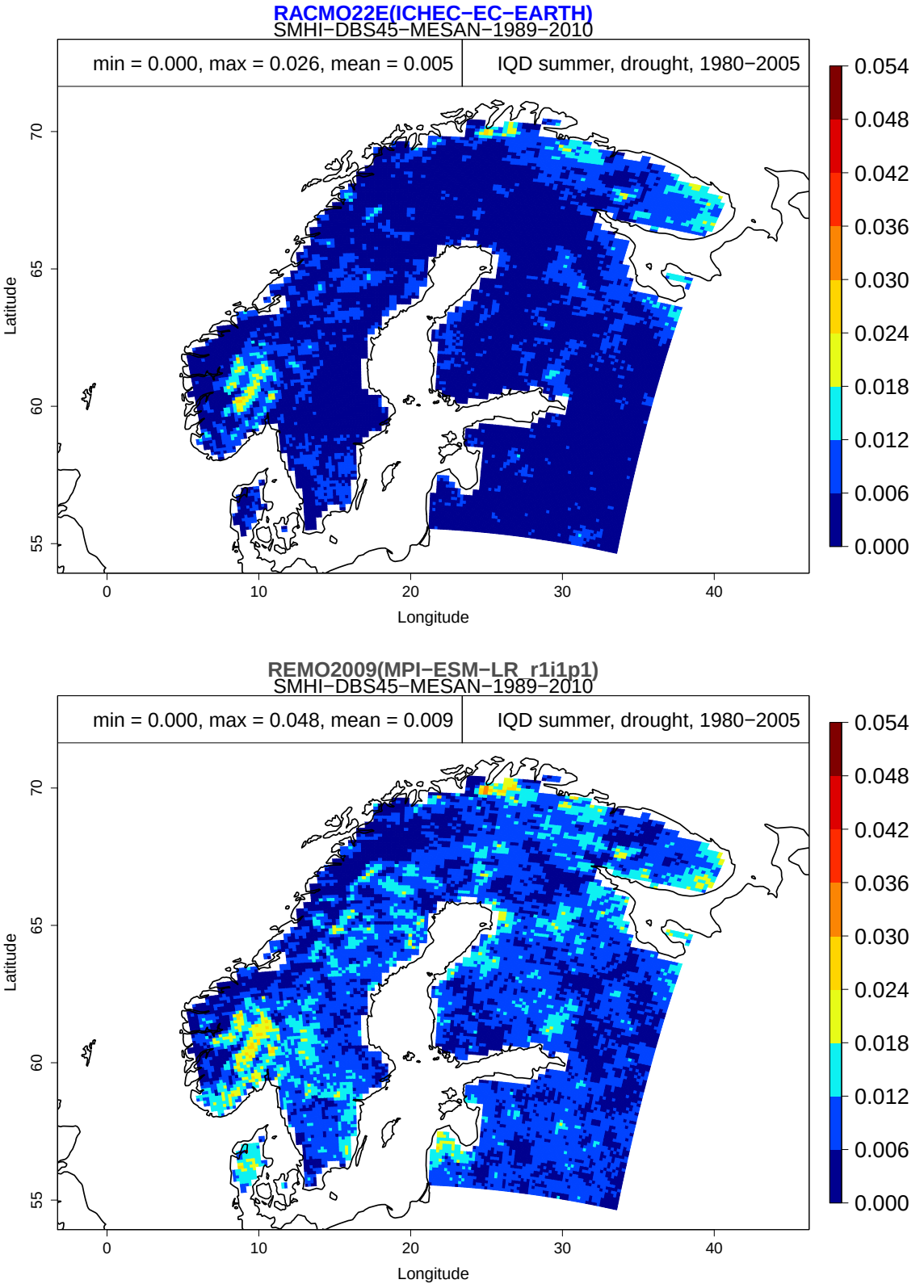


Figure E.21. Mean drought IQD for the models 3 and 6 during summer after applying bias correction method 3.

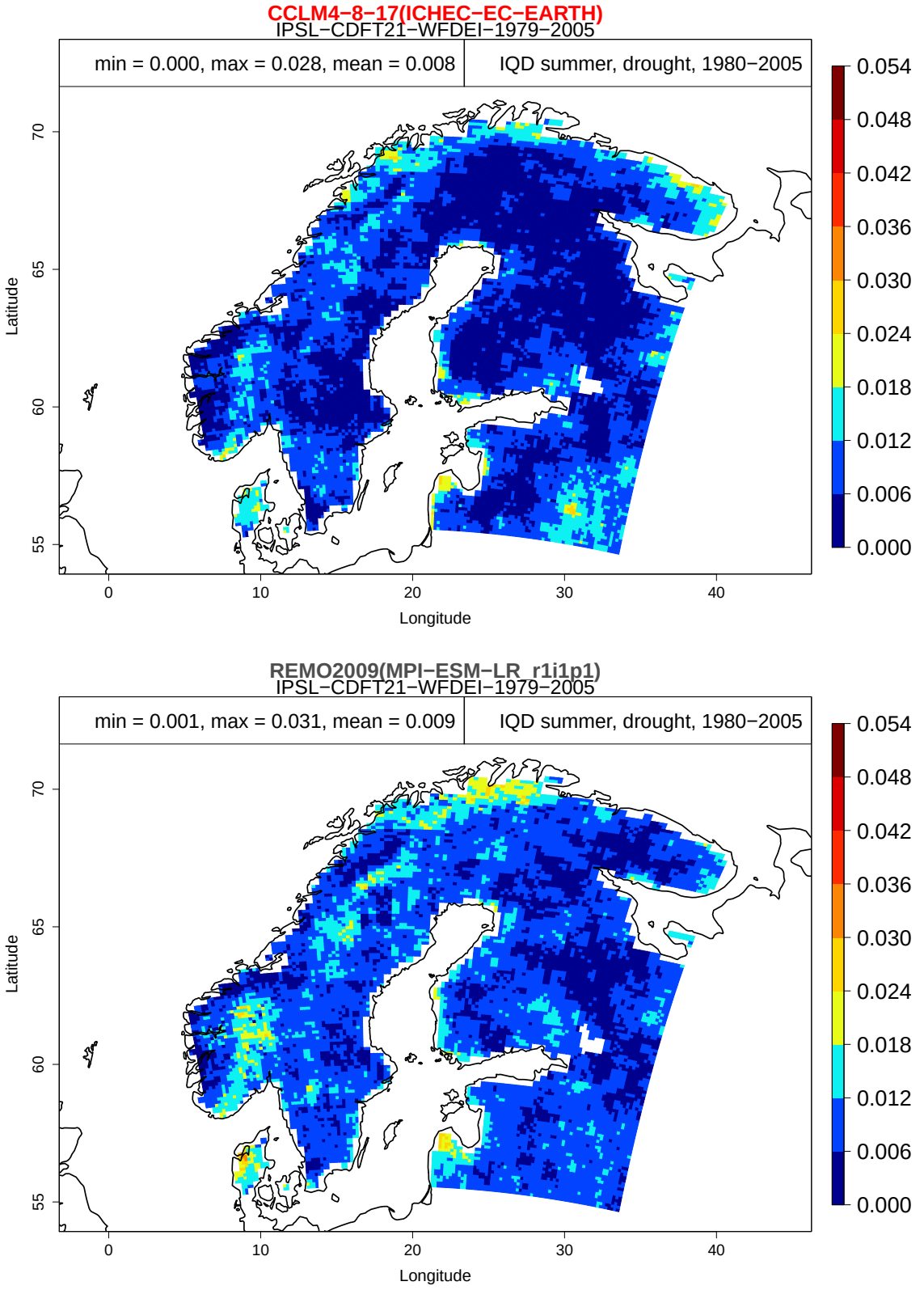


Figure E.22. Mean drought IQD for the models 4 and 6 during summer after applying bias correction method 4.

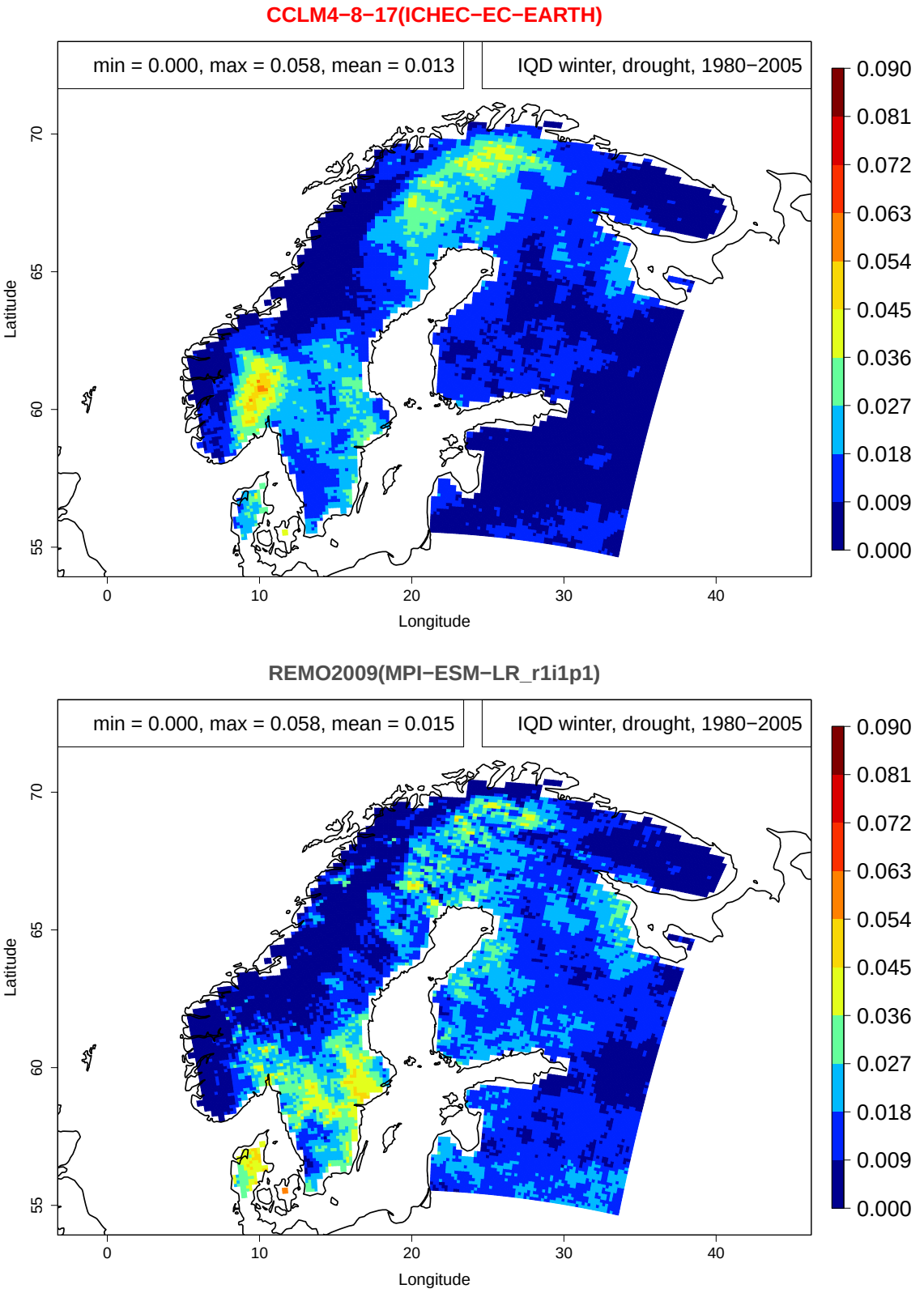


Figure E.23. Mean drought IQD for the models 4 and 6 during winter.

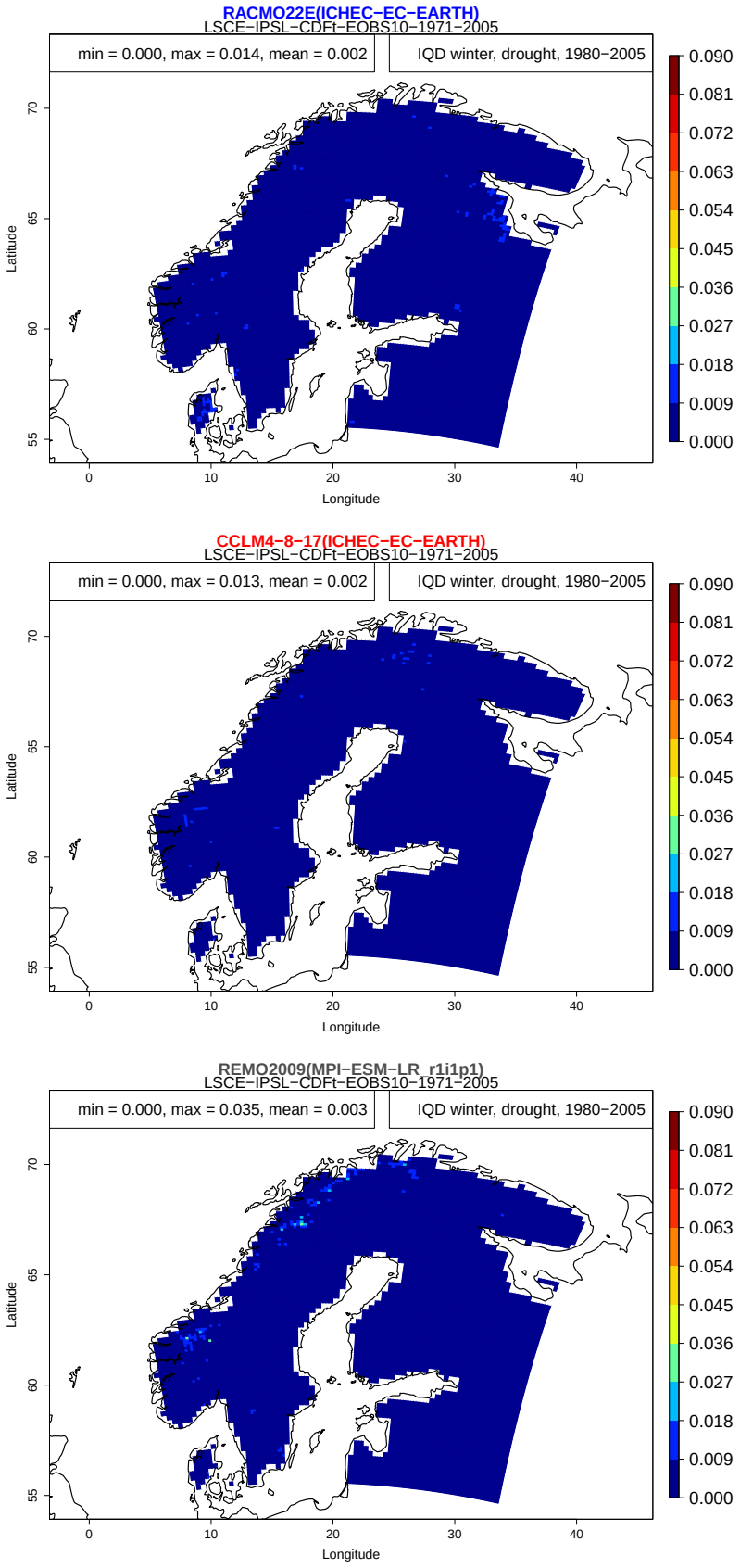


Figure E.24. Mean drought IQD for the models 3, 4 and 6 during winter after applying bias correction method 1.

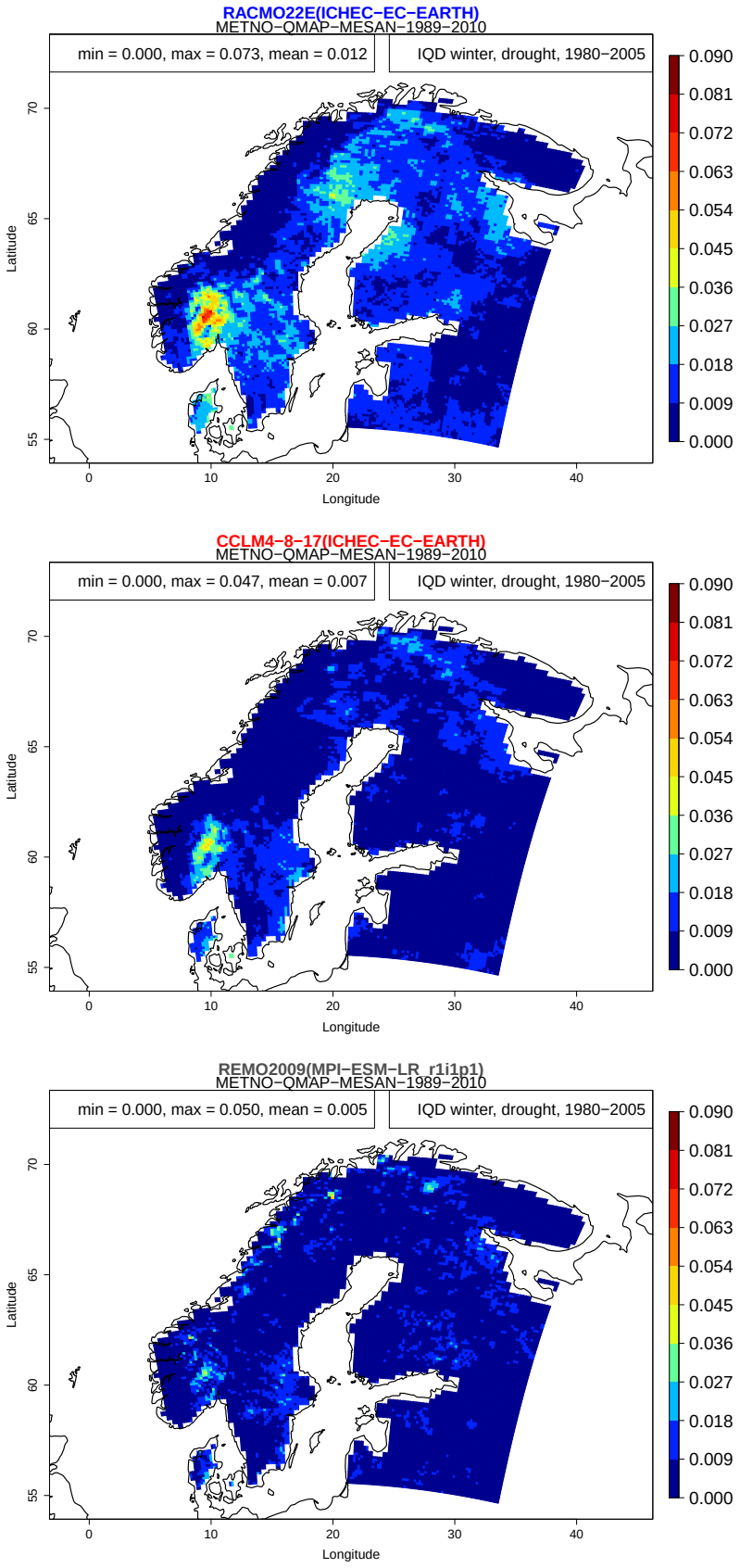


Figure E.25. Mean drought IQD for the models 3, 4 and 6 during winter after applying bias correction method 2.

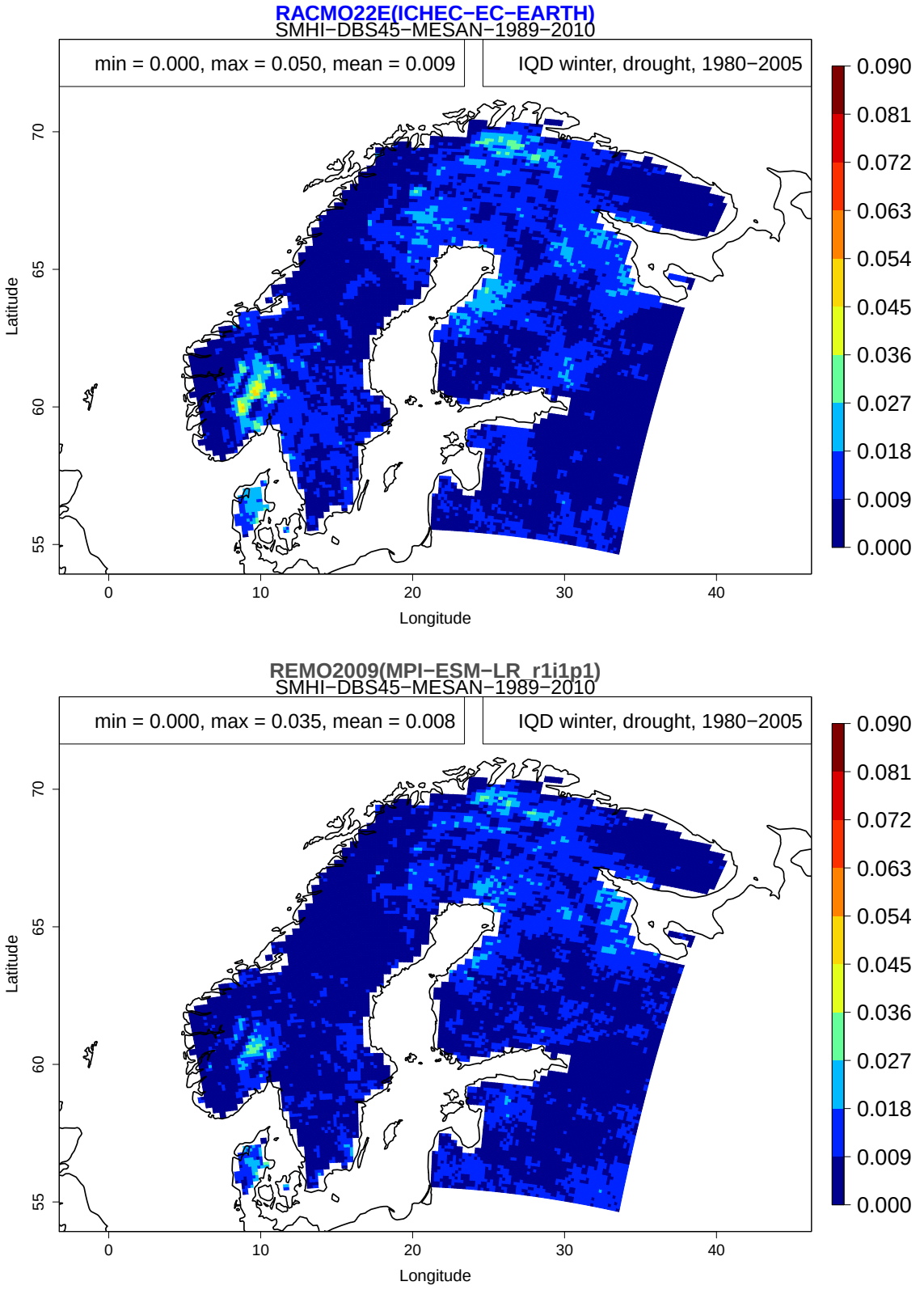


Figure E.26. Mean drought IQD for the models 3 and 6 during winter after applying bias correction method 3.

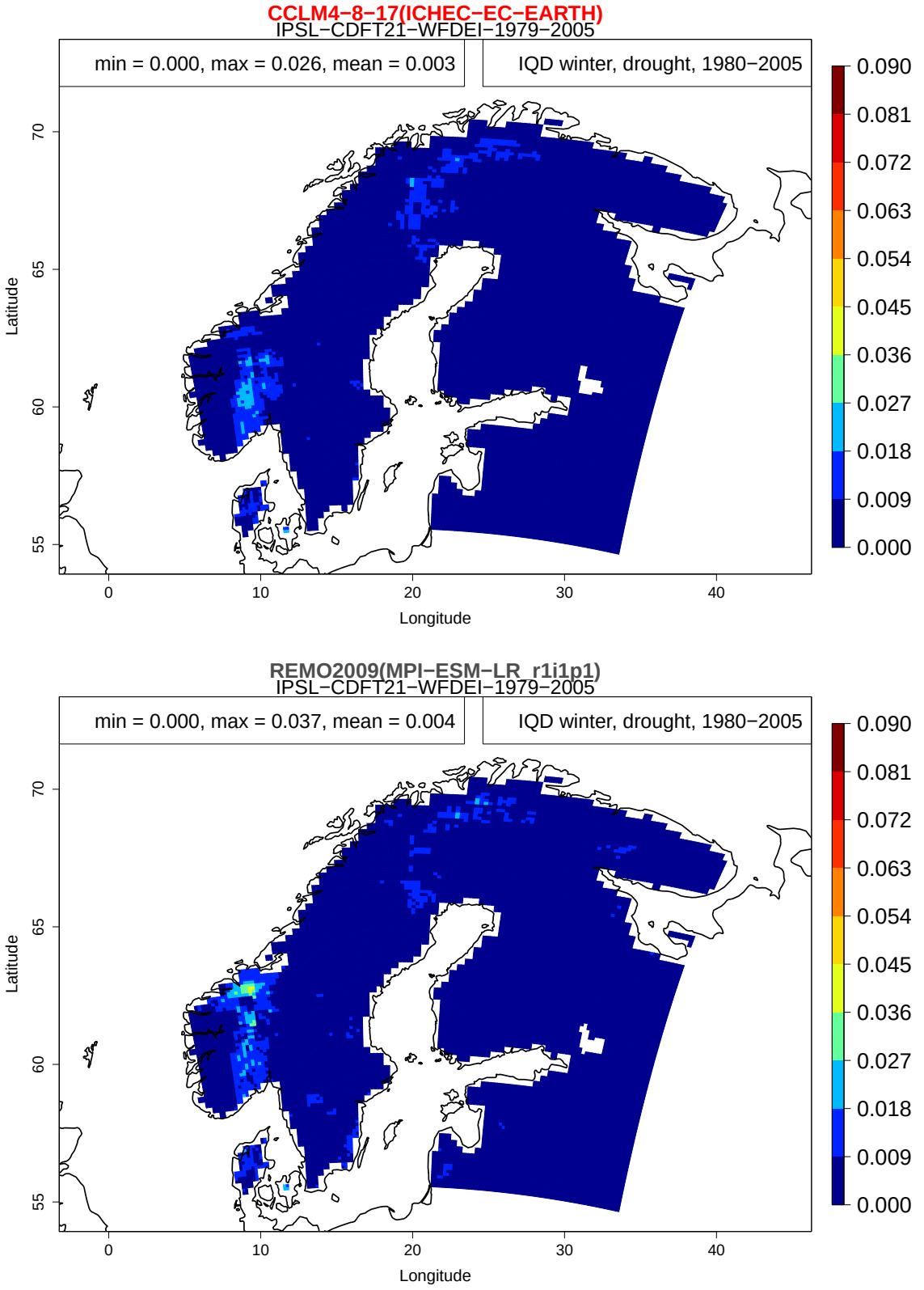


Figure E.27. Mean drought IQD for the models 4 and 6 during winter after applying bias correction method 4.

F Raw vs. bias corrected drought IQD

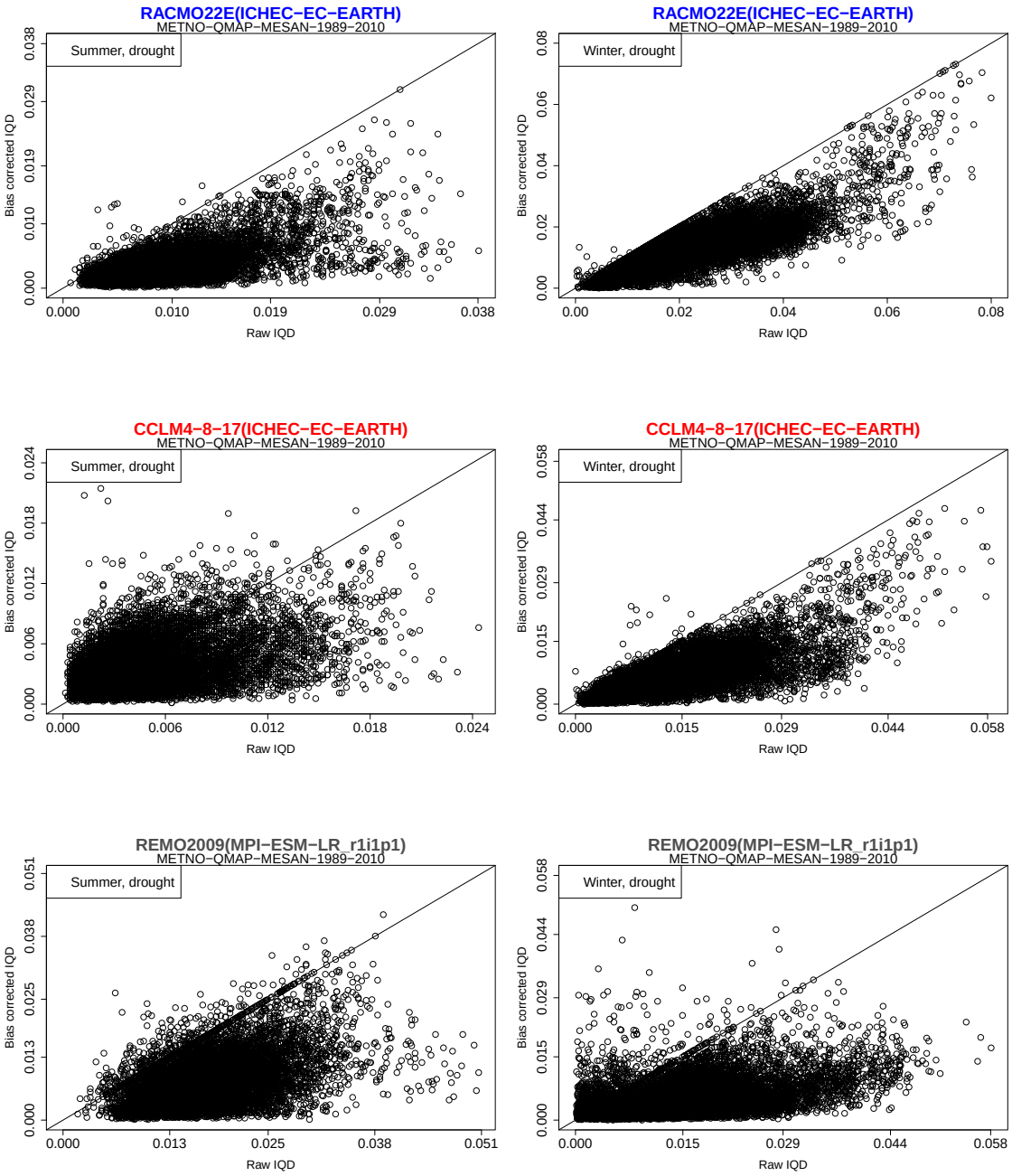


Figure F.28. Scatter plots displaying drought IQD of raw data and bias corrected data from method 2 for the models 3, 4 and 6 during summer (on the left) and winter (on the right).

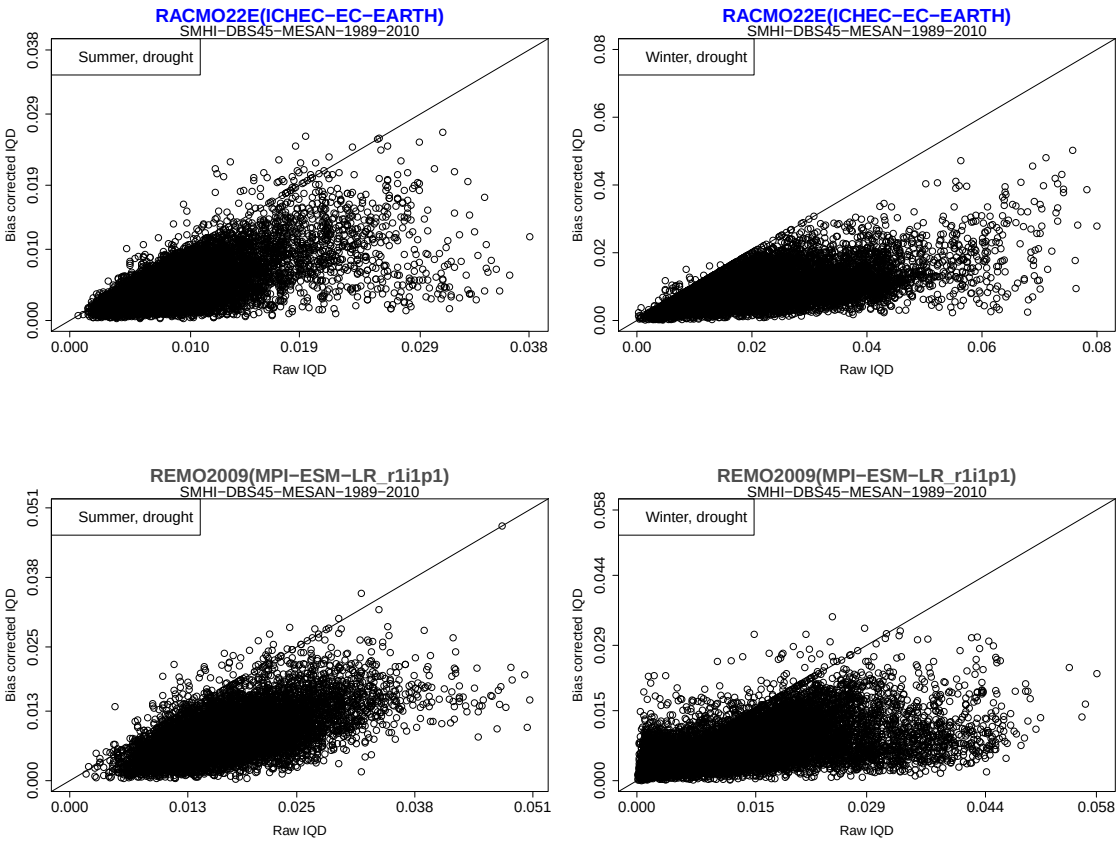


Figure F.29. Scatter plots displaying drought IQD of raw data and bias corrected data from method 3 for the models 3 and 6 during summer (on the left) and winter (on the right).