https://doi.org/10.5194/gmd-2019-143
Preprint. Discussion started: 28 June 2019

Geoscientific Model Development
Discussions

Uncertainties in climate change projections covered by the ISIMIP and CORDEX model subsets from CMIP5

Ito, Rui[1,2], Hideo Shiogama[3], Tosiyuki Nakaegawa[2], Izuru Takayabu[2]

[1]Japan Meteorological Business Support Center, Tsukuba, 305-0052, Japan
[2]Meteorological Research Institute, Japan Meteorological Agency, Tsukuba, 305-0052, Japan
[3]National Institute for Environmental Studies, Tsukuba, 305-0053, Japan

Correspondence to: Rui Ito (rui.ito@jmbsc.or.jp)

Abstract. Two international projects, ISIMIP (Inter-sectoral Impact Model Inter-comparison Project) and CORDEX (Coordinated Regional Climate Downscaling Experiment), have been established to assess the impacts of global climate change and improve our understanding of regional climate, respectively. Model selection from the GCMs (general circulation models) within CMIP5 (fifth phase of the Coupled Model Inter-comparison Project) was conducted by the different approaches for each project: one is a globally consistent model subset used in ISIMIP and another is a region-specific model subset for each region of interest used in CORDEX. We evaluated the ability to reproduce the regional climatological state by comparing the subsets with the full set of CMIP5 multimodel ensemble. We also investigated how well the subsets captured the uncertainty in the climate change projected by the full set, to provide increased credibility for the scientific outcomes from each project. The spreads of the biases and Taylor's skill scores from the ISIMIP and CORDEX subsets are smaller than that from the full set for the regional means of surface air temperature and precipitation. However, the spreads in ISIMIP and CORDEX extend beyond the spreads from high performed models from full set, despite using a smaller number of models. It was shown that better subsets exist that would have smaller biases and/or higher scores than the current subset. The ISIMIP subset captures the uncertainty range of the regional mean of temperature change projections by the full set better than the CORDEX subsets in 10 of 14 terrestrial regions worldwide. Compared with the randomly selected 10,000 subset samples, CORDEX uses a subset with relatively low coverage of the uncertainty range for the temperature change in some regions, and ISIMIP uses a subset with relatively high coverage in all regions. On the other hand, for the regional mean of precipitation change projections, although the coverage from the CORDEX subsets is lower among the 10,000 subset samples in half of the regions, the CORDEX subsets indicate a tendency for better coverage of the uncertainty range than the ISIMIP subsets. In the regions where CORDEX used nine models or more, good coverage (>50%) is evident for the projections of both temperature and precipitation. The globally consistent model subset used in ISIMIP could have difficulty in capturing uncertainties in the regional precipitation change projections, whereas it widely covers uncertainties in the temperature change projections. The region-specific model subset, like CORDEX, can cover the relative-widely uncertainties in both temperature and precipitation changes, but it depends on the number of models used.

https://doi.org/10.5194/gmd-2019-143
Preprint. Discussion started: 28 June 2019

Geoscientific Model Development Discussions

1 Introduction

A global dataset of climate change projections has been generated by the Coupled Model Inter-comparison Projects (CMIP). Using this dataset, numerous climatological studies have been in progress to advance our understanding of the increasingly severe problems associated with climate change. Regarding regional climate change, dynamical and statistical downscaling experiments have been conducted to create high-resolution climate products derived from the global CMIP dataset via a regional climate model. In addition, impact studies and examinations of adaptation planning have progressed in close parallel with the climate studies, using those climate products at both global and regional scales.

When we conduct an impact assessment of climate change and consider possible adaptation or mitigation measures, the information regarding the largest potential change in the climate is required to consider the most severe states of climate change, in addition to information regarding how the climate changes on average. Although the CMIP multiple global climate model (GCM) ensemble is the ensemble of opportunity and do not necessarily represent the full uncertainty in the climate projections (Knutti 2010), they are useful for investigating the uncertainty in the future projections. By using the climate projections from the CMIP ensemble, it is at least possible to examine the maximum-minimum climate change scenarios within the ensemble.

Although it is desirable to use GCMs as much as possible, due to limitations in computing resources, relatively small subsets of the models are generally used in regional downscaling studies and impact assessments. The subset is selected under the conditions that the simulation accuracy is better for the climatological state of interest or the data required for the study is readily available. Methods of specifying the best subset, based on the accuracy of the historical climate simulations and/or capturing the possible maximum range in the variation of projections among the models (hereafter uncertainty), have been proposed (Reichler and Kim 2008; Cannon 2015; Mendlik and Gobiet 2016), but the optimum approach remains to be determined. When the sample size of a subset is limited, appropriate strategies are necessary to select subsets of GCMs that have smaller biases in the historical climate simulations and cover the widest possible uncertainty range of future projections. Without such a strategy, we might erroneously interpret the information regarding climate change and impact assessment obtained from the subsets.

The inter-sectoral impact model inter-comparison project (ISIMIP; https://www.isimip.org) was designed as a framework to assess the impacts of climate change in different sectors and at different scales (Schellnhuber et al. 2014). This project used consistent climate and socio-economic input data to multiple impact models. Five GCMs were selected in the fast track of ISIMIP: HadGEM2-ES, GFDL-ESM2, IPSL-CM5A-LR, MIROC-ESM-CHEM, and NorESM1-M. The main selection condition was that the climate data generated by the models was available at the relevant stage of the project, with the attempt of broadly capturing the global change in surface air temperature (hereafter referred to as 'temperature' for simplicity) and precipitation (Warszawski et al. 2014; ISIMIP protocol 2018). A feature of the uncertainty range identified from these five GCMs was investigated in detail by McSweeney and Jones (2016), who indicated that the subset of these five GCMs covers the uncertainty in the projected changes in the temperature and precipitation expressed from 36 CMIP5 GCMs wider than the other five-GCM subsets which were randomly sampled. In addition, a higher coverage of the uncertainty range had been shown

https://doi.org/10.5194/gmd-2019-143
Preprint. Discussion started: 28 June 2019

to appear as an average of 26 global regions, in region-specific subsets than in globally consistent subsets which are consistent with the aim of ISIMIP.

One subset of GCMs was globally used in ISIMIP, but in the coordinated regional climate downscaling experiment (CORDEX; http://www.cordex.org) project, a GCM subset was selected for each defined region to generate a regional climate dataset for climate studies and impact assessments (Giorgi et al. 2009; Giorgi and Gutowski 2015). Fourteen regions of interest were defined and subsets of between 3 and 15 GCMs were used for each region. The conditions required here were that input data to a regional climate model (RCM) were available and easily acquired, and they also tended to select GCMs that were developed at the institute located in the region of interest. The advantage of CORDEX is that it enables a regional climate assessment using a high-resolution climate dataset from 'optimal' multi-GCMs and multi-RCMs for the region of interest. Meanwhile, in the next generation of CORDEX to be included in the sixth phase of CMIP, they have an intention to downscale projections from a core set of GCMs as a minimum model set that is common across the regions, similar to the approach in ISIMIP (Gutowski et al. 2016).

A globally consistent GCM subset will facilitate discussion of climate change and its impacts beyond regional divisions. However, it is unclear whether the globally consistent subset adequately represents the phenomena that characterize the climate in the region of interest. In particular, the spatial pattern of a projected change in precipitation is strongly dependent on the GCMs selected (Giorgi and Gutowski 2015; McSweeney et al. 2015). Therefore, the possibility of insufficiently capturing the regional climate change and its valid uncertainty could be increased, as noted by McSweeney and Jones (2016). In contrast, a region-specific GCM subset can include GCMs which more precisely reproduce the target regional climate (McSweeney et al. 2015). However, it does not enable discussions about the difference among regions and the interaction of impacts across the regions. Although there are advantages to both approaches to select a subset, it is necessary that we understand the characteristics of the current subsets selected using the approaches of the ongoing projects if we are to improve the process in the next generations of the projects.

In this study, we assessed the current subsets of CMIP5 multi-GCM ensemble being used in ISIMIP and CORDEX by clarifying the climatological characteristics expressed by each subset, which is an important aspect for increasing the credibility of the scientific outcomes from each project. By comparing the simulations of the subsets and also the full set of the multi-GCM ensemble with observed data, we evaluated their ability to reproduce the historical climate (i.e., model performance). We also compared the projected change of climate between the subsets and the full set, and clarified how extent the uncertainty in the projections obtained from the subsets covers the uncertainty from the full set. In addition, with reference to McSweeney and Jones (2016), we also explored whether the subset used was able to capture the uncertainty from the full set more widely than the other model subsets when using the same sample size. Although the five GCMs analysed by McSweeney and Jones (2016) were selected in the fast track of ISIMIP, this has been changed to four GCMs in the next round of the ISIMIP simulations (ISIMIP2b; Frieler et al. 2017). Frieler et al. (2017) explained that NorESM1-M was removed from the five GCMs because of a lack of near-surface wind data, and MIROC-ESM-CHEM was changed to MIROC5 because of the horizontal resolution and improvements in the representation of various fields (e.g., El Niño–Southern Oscillation and the monsoon) in the historical experiments. Therefore, we used the four GCMs from ISIMIP2b here, and ISIMIP refers to ISIMIP2b hereafter unless specified otherwise.

2 Data and Methods

2.1 Dataset

We analysed the historical runs of 49 atmosphere–ocean GCMs (AOGCMs) and the Representative Concentration Pathways (RCP) 8.5 scenario runs of 42 AOGCMs participating in CMIP5 (Taylor et al. 2012). A single ensemble member, r1i1p1, was selected for each model, except for CESM1-WACCM (r2i1p1) and EC-EARTH (r8i1p1). It is because the member, r1i1p1, of CESM1-WACCM was not available and temperature change from r1i1p1 of EC-EARTH was over two-standard deviation of the changes from the 42 models in more than 60% of our target regions. In the followings, the full set of the multi-GCM ensemble indicates the 49 historical runs when we assessed the ability to reproduce the historical climate ($\mathrm{CMIP}_{\mathrm{Full\_Hist}}$), while does the 42 future projections which are estimated from both historical and rcp85 runs when we discussed the future projections ($\mathrm{CMIP}_{\mathrm{Full\_Future}}$).

We compared the simulations of the subsets of GCMs used in ISIMIP and CORDEX with the full ensemble. ISIMIP used four GCMs for their various impact assessments: GFDL-ESM2M, HadGEM2-ES, IPSL-CM5A-LR and MIROC5 (Frieler et al. 2017). On the other hand, CORDEX used the subset in which the combination of GCMs were altered for each defined region. The number of GCMs used in each of the defined regions is listed in Table 1, and the GCMs used in CORDEX are listed in Supplement 1. The regional classification used to investigate the regional performance and the projection was based on the classification in CORDEX shown in Supplement 2. In this study, we focused on global land area only.

The analysis periods were the year 1986–2005 (but 1985–2004 for HadGEM2-CC and HadGEM2-ES) for the historical runs and the year 2081–2100 (but 2080–2099 for MRI-AGCM60 and CESM1-WACCM) for the RCP8.5 runs. Monthly mean temperature and precipitation data over these periods were interpolated onto a 2.5° × 2.5° grid for each model. 'Too dry' grids (then mean precipitation are < 0.1 mm/day in each member) were excluded from the analyses using precipitation.

To validate the model representations, we compared the simulated estimates with the observed datasets. With respect to precipitation, Sun et al. (2018) highlighted differences among the observational datasets. Consequently, to avoid a misreading of the model performance due to such discrepancies, we used multi-precipitation products that covered the global land area over the period of interest. The observation products were the Climatic Research Unit Timeseries (CRU) v.4.01 (Harris et al. 2014) for temperature and precipitation, and the following for precipitation only: the CPC v.1.0 (Xie et al. 2010), the Global Precipitation Climatology Centre (GPCC) full data reanalysis v.7.0 (Schneider et al. 2016), NOAA's Precipitation reconstruction over Land (PRECL) v.1.0 (Chen et al. 2002), the CPC Merged Analysis of Precipitation (CMAP; Xie and Arkin, 1997), the Global Precipitation Climatology Project (GPCP) v.2.2 (Huffman et al. 2015), and the Multi-Source Weighted-Ensemble Precipitation (MSWEP) v2.1 (Beck et al. 2019). To quantify the ability to reproduce spatial patterns of the observations, we used Taylor's skill score (Taylor 2001) (hereafter referred to as skill score).

https://doi.org/10.5194/gmd-2019-143
Preprint. Discussion started: 28 June 2019

Geoscientific Model Development Discussions

2.2 Coverage of uncertainty and random selection

Coverage was estimated from a comparison between the full uncertainty range of the projections made by two model sets, which was defined by McSweeney et al. (2015) as a fractional range coverage, FRC. In this study, we computed the regionally averaged projections for each model, and then the FRC were estimated using the regional averages for each model. The FRC from the regional averages (FRA) was defined as the fraction of the uncertainty range of the regionally averaged projections obtained from each model subset (R_{Sub}) to the range from $\mathrm{CMIP}_{\mathrm{Full\_Future}}$ (R_{Full}), as follows:

$$FRA = \frac{R_{Sub}}{R_{Full}}. \tag{1}$$

To investigate how well the model subsets used in ISIMIP and CORDEX captured the uncertainty range of projections compared with the other arbitrary subsets, which McSweeney and Jones (2016) presented as ‘representation’, we randomly selected n models without repetition from $\mathrm{CMIP}_{\mathrm{Full\_Future}}$ 10,000 times, where n is the sample size of subsets in ISIMIP ($n = 4$) or CORDEX (n depends on the regions; see Table 1). Then, the 10,000 of R_{Sub} values and the spread from the 10,000 of the FRA were estimated from the samples.

3 Results and discussion

3.1 Performance in reproducing historical climate

Using model biases and skill scores, we evaluated the historical climate reproduced by the GCM subsets used in ISIMIP and CORDEX. The GCM subsets used in ISIMIP and CORDEX are hereafter referred to as the ISIMIP subsets and CORDEX subsets, respectively. For the evaluations, we also used two high performed subsets: one is composed of models with lower bias than the 50th percentile (median) of the $\mathrm{CMIP}_{\mathrm{Full\_Hist}}$ biases: the other is models with higher skill score than the median of the $\mathrm{CMIP}_{\mathrm{Full\_Hist}}$ scores (referred to $\mathrm{CMIP}_{\mathrm{lowB}}$ and $\mathrm{CMIP}_{\mathrm{highS}}$, respectively). B($v$($E$)) and S($v$($E$)) indicate the regional mean biases and skill scores for variable v and ensemble subset E, respectively.

Figure 1 shows the model bias associated with the annual mean precipitation in the 14 CORDEX regions over a 20-year period. Compared with the maximum values of B(P($\mathrm{CMIP}_{\mathrm{Full\_Hist}}$)) for the precipitation (v=P), the maximum values of B(P(ISIMIP)) and B(P(CORDEX)) are clearly small, especially in the Mediterranean (MED), Southeast Asia (SEA), and the polar regions. The spreads of B(P(ISIMIP)) and B(P(CORDEX)) in MED are within the spread of the discrepancy among the observations, which suggests that the model selection works effectively to select models with high ability to reproduce the observed regional mean precipitation. However, compared with the high performed subsets, some models in the ISIMIP and CORDEX subsets have a bias exceeding the maximum values of B(P($\mathrm{CMIP}_{\mathrm{lowB}}$)), or B(P($\mathrm{CMIP}_{\mathrm{highS}}$)) in some regions, despite the small number of models used in ISIMIP and CORDEX. Therefore, our results indicate that less bias models could be selected than those currently being used. The difference in the spread between the ISIMIP and CORDEX subsets is large in the high- and mid-latitude Northern Hemisphere regions, regardless of the number of models.

https://doi.org/10.5194/gmd-2019-143
Preprint. Discussion started: 28 June 2019

Geoscientific Model Development Discussions

With respect to the spatial pattern of the annual mean precipitation, ISIMIP and CORDEX incorporate some models with a worse score than the minimum value of S(P($CMIP_{highS}$)) (Supplement 3). S(P(ISIMIP)) and S(P(CORDEX)) fall within the observational spread only in the Arctic.

We also assessed model performance for the annual mean temperature (v=T). The spread of B(T($CMIP_{Full\_Hist}$)) is greater in the high- and mid-latitude Northern Hemisphere regions than in the low-latitude Northern and Southern hemisphere regions, which would be related to the magnitude of seasonal variability (Supplement 4). The same spatial pattern of spread is also evident in B(T(ISIMIP)) and B(T(CORDEX)). The maximum biases of B(T(ISIMIP)) and B(T(CORDEX)) are smaller, or equal to the maximum value of B(T($CMIP_{highS}$)) (except for the CORDEX subsets in East Asia and North America), but are larger than the maximum value of B(T($CMIP_{lowB}$)). The spreads of B(T(ISIMIP)) and B(T(CORDEX)) are similar, indicating that CORDEX used models with a similar performance to ISIMIP, despite using more models than ISIMIP (except for Central Asia). Both subsets included models with a worse score than the minimum value of S(T($CMIP_{highS}$)) in 85% of the regions (Supplement 5).

Even though the model selections conducted in ISIMIP and CORDEX narrow the spreads of model bias and the score from $CMIP_{Full\_Hist}$, the largest bias and the worst score from the ISIMIP and CORDEX subsets distribute beyond the biases and the scores from high performed models in the full set. Therefore, a much better model subset, regarding to biases and skill scores, can be selected with making use of the advantage of the small number of models.

3.2 Uncertainty range of the projected changes in annual mean temperature and precipitation

Future projections obtained from the ISIMIP and CORDEX subsets were compared with those from the full set, and also from high performed models, as with the evaluations in Section 3.1. Because the small biases or high skill scores models used in this section are composed of the models included in $CMIP_{Full\_Future}$, we refer as $CMIP'_{lowB}$ and $CMIP'_{highS}$ instead of $CMIP_{lowB}$ and $CMIP_{highS}$. Projected change of annual mean temperature and precipitation are designated by ΔT(E) and ΔP(E), respectively.

Figure 2 shows the uncertainty range of the projected increments of the temperature for each GCM subset. Although ISIMIP used fewer models than CORDEX, the uncertainty range of ΔT(ISIMIP) exceeds that of ΔT(CORDEX) except for South Asia, Australasia, South America, and Central America. The uncertainty ranges of ΔT($CMIP'_{lowB}$) and ΔT($CMIP'_{highS}$) broadly cover the range of ΔT($CMIP_{Full\_Future}$), suggesting that the bias and skill score are not good emergent constraints to reduce the uncertainty of ΔT.

The uncertainty range associated with the projected change in annual precipitation is shown in Fig. 3. Compared with ΔT in Fig. 2, model selection has a large impact on the reduction of the uncertainty in ΔP, as was also found by McSweeney and Jones (2016) using five GCMs used in the fast track of ISIMIP. The interquartile range of ΔP(CORDEX), IQR(ΔP(CORDEX)), shows a high coincidence with the IQR(ΔP($CMIP_{Full\_Future}$)), as well as with the IQR(ΔP($CMIP'_{lowB}$)) and IQR(ΔP($CMIP'_{highS}$)) (yellow and orange boxes in Fig. 3). Therefore, the CORDEX subsets can capture the average tendency of the change projected by the 25th to 75th percentile of $CMIP_{Full\_Future}$. In addition, the median of the uncertainty range is similar between the

https://doi.org/10.5194/gmd-2019-143
Preprint. Discussion started: 28 June 2019

CORDEX subset and $\mathrm{CMIP}_{\mathrm{Full\_Future}}$. Only in Central Asia does the maximum−minimum range of ΔP(CORDEX) extend below the 25th percentile of ΔP($\mathrm{CMIP}_{\mathrm{Full\_Future}}$) and, in contrast, the maximum−minimum range of ΔP(ISIMIP) covers the IQR(ΔP($\mathrm{CMIP}_{\mathrm{Full\_Future}}$)). Thus, three models of the CORDEX subset in Central Asia cannot capture the average tendency of the change projected by $\mathrm{CMIP}_{\mathrm{Full\_Future}}$, despite being able to select suitable models to discuss the climate change in Central Asia, differing from ISIMIP.

With regard to the full range of uncertainty from $\mathrm{CMIP}_{\mathrm{Full\_Future}}$, the CORDEX subsets capture more than 50% of the full range in eight regions (Europe, MED, Africa, SEA, Australasia, Central America, South America and the Antarctica). On the other hand, the ISIMIP subsets capture the full range less than 60% in all regions. In 11 regions, the CORDEX subsets capture the wider range than the ISIMIP subsets, differing from broad coverage by the ISIMIP subset for ΔT as seen in Fig. 2. Therefore, global consistent four models used in ISIMIP2b, which are taken into consideration of the ability of reproduction, still remains difficult to capture the uncertainties in regional precipitation change, as in McSweeney and Jones (2016) which analysed for five models in the fast track.

3.3 Comparison of uncertainty of the projected changes using randomly sampled models

We investigated whether the ISIMIP or CORDEX subsets were more suitable for capturing the uncertainty range obtained from $\mathrm{CMIP}_{\mathrm{Full\_Future}}$ by comparing the fractional coverage of uncertainty, FRA, of each subset with those of the 10,000 randomly sampled subsets. As the result, the ISIMIP subset (four models) shows high coverage for the temperature change in all regions compared with the random samples. By contrast, the CORDEX subset yields relatively wide coverage for the temperature and precipitation changes, but this depends on the number of models used.

Figure 4 illustrates FRA of the ISIMIP and CORDEX subsets (referred to $\mathrm{FRA}_{\mathrm{ISIMIP}}$ and $\mathrm{FRA}_{\mathrm{CORDEX}}$, respectively) in each region. Along the x-axis, the name of regions is arranged in ascending order of the number of models used in CORDEX. The number of models used in CORDEX is indicated in each parenthesis after the name, and by contrast, the number in ISIMIP is four in all regions. The y-axis indicates FRA of the uncertainty from each subset relative to that from the full set. The bar presents distribution of the FRA values obtained from the possible 10,000 random samples ($\mathrm{FRA}_{\mathrm{Random}}$). The blue bar means the distribution using the subsets with four models ($\mathrm{FRA}_{\mathrm{Random\_I}}$), as large as the ISIMIP subset, and the red bar means that with the same number of models used in CORDEX ($\mathrm{FRA}_{\mathrm{Random\_C}}$). Both ends of the bar indicate the lowest and highest values of FRA, and both ends of the bar with a dark color and horizontal line in the bar denotes the 25th and 75th percentiles and the median, respectively.

For the temperature change, ΔT, $\mathrm{FRA}_{\mathrm{ISIMIP}}$ and $\mathrm{FRA}_{\mathrm{CORDEX}}$ (blue and red dots, respectively) exceed 60% in 13 and 10 regions, respectively (Fig. 4a). However, $\mathrm{FRA}_{\mathrm{CORDEX}}$ locates in the range of around the 25th percentile or less of $\mathrm{FRA}_{\mathrm{Random\_C}}$ (red bar) in MED, East Asia, SEA, Europe, and the polar regions where $\mathrm{FRA}_{\mathrm{CORDEX}}$ is lower than $\mathrm{FRA}_{\mathrm{ISIMIP}}$. In the region with a larger number of models in CORDEX, $\mathrm{FRA}_{\mathrm{CORDEX}}$ tends to be less than the median of $\mathrm{FRA}_{\mathrm{Random\_C}}$. On the other hand, $\mathrm{FRA}_{\mathrm{ISIMIP}}$ is typically around the 75th percentile or higher than the median of $\mathrm{FRA}_{\mathrm{Random\_I}}$ (blue bar) for all regions.

https://doi.org/10.5194/gmd-2019-143
Preprint. Discussion started: 28 June 2019

Geoscientific Model Development Discussions

A relatively high coverage, above ~50%, is obtained for both changes of temperature and precipitation in eight regions when using nine models or more, except for temperature in Antarctica (Fig. 4b). The value of FRA_{CORDEX} for ΔP is lower than those for ΔT. A high coverage of more than 70%, however, can be gained in MED, South America, Europe, Australasia and Africa, which indicates a wide range compared with the median of FRA_{Random_C} (except for Europe). In half of the regions, FRA_{CORDEX} are in the range of the 25th percentile or less of FRA_{Random_C} (four regions of Asia, MENA, the Arctic, and North America). In Central and East Asia, and North America of these regions, FRA_{CORDEX} is smaller than FRA_{ISIMIP}, even though CORDEX has the advantages of selecting suitable models for the region and also more models can be used, especially in East Asia and North America. The ISIMIP subsets in Antarctica and Australasia show a larger coverage than the 75th percentile of FRA_{Random_I}, but the FRA_{ISIMIP} of 60% is less than that for ΔT. In more than 60% of all regions, FRA_{ISIMIP} is less than the median of FRA_{Random_I}; the averaged FRA_{ISIMIP} over all regions is 33%.

From the FRA distributions estimated from the possible random samples regarding to both changes, ΔT and ΔP, the IQR of FRA_{Random_C} itself rises toward a FRA of 100% as a larger number of models is used. When random samples are composed of a subset with 15 models as large as subsets in CORDEX-Africa and -South Asia, the 75th percentile of FRA_{Random_C} is more than 90% in ΔT (Fig. 4a). In addition, the width of the IQR for ΔT is narrowed with increasing number of models used. The relationship between the number of models and FRA is clearly evident in ΔT because there is a small difference in R_{Full} among regions for ΔT compared with ΔP (Fig. 2), and thus the larger number of models results in an increase in the FRA. And also, we found that the probability of selecting model subsets with a low coverage was higher for precipitation than for temperature, even if the number of models selected increases.

The number of models used in CORDEX are unequal among the regions, especially only three in Central Asia (Gutowski et al. 2016). When we add three, five, or seven randomly selected models to the three current models in Central Asia, the FRA for ΔP increases from 15% to 30%, 50%, and 65%, respectively, at the median of the FRAs from the random samples (not shown).

4 Summary and conclusions

We explored the ability for the subsets of CMIP5 multimodel ensemble used in ISIMIP2b and CORDEX to reproduce the observed temperature and precipitation, and how the subsets capture the uncertainty in projected change of temperature and precipitation obtained from the full set of the ensemble. In addition, we discussed whether each subset shows a high coverage of the uncertainty in projected climate change compared with the possible subsets generated using 10,000 random samples.

The spreads of the bias and Taylor's skill score from the subsets used in ISIMIP and CORDEX are smaller than those obtained from the full set of CMIP5 ensemble for the annual mean temperature and precipitation. However, despite of the smaller number of models in ISIMIP and CORDEX, the largest bias and the worst skill score distribute beyond the biases and the scores obtained from the half member subsets with less bias or high score of the full set. Therefore, although the ISIMIP and CORDEX approaches were able to select models that acceptably performed to represent the historical state, our results suggest that better subsets can be selected by focusing on smaller biases and/or higher scores for representing the historical climate.

https://doi.org/10.5194/gmd-2019-143
Preprint. Discussion started: 28 June 2019

For the projected change in annual mean temperature, the subsets capture more than 60% of the uncertainty for the full set in the 13 terrestrial regions in ISIMIP and the 10 regions in CORDEX, from the total of 14 regions. The coverage of the uncertainty range by the ISIMIP subset exceeds the coverage by the CORDEX subset in 10 regions by using only four models that are common to all regions. The FRA of the current CORDEX subset tends to be lower than the 50th percentile of the FRAs obtained from the possible 10,000 random samples in the regions where a large number of models are used. ISIMIP selected the subset of models with relatively high coverage of the uncertainty from the full set in all regions, compared with the 50th percentile from the random samples.

On the other hand, for the projected change in annual mean precipitation, the FRA for the CORDEX subset are around the 25th percentile or less of the FRAs from the random samples with the same number of models in half of all regions. However, CORDEX broadly captures the uncertainty range more than ISIMIP, differing from the temperature change. Additionally, a relatively high coverage (>50%) was obtained for the projections of both temperature and precipitation in eight regions when using nine models or more.

Compared with the random samples, the ISIMIP subset shows high coverage for the temperature change in all regions and, by contrast, low coverage for the precipitation change in more than 60% of the regions. Therefore, the global common model set used in ISIMIP could have difficulty in capturing the uncertainty in regional precipitation change projections with capturing most of the uncertainty in the temperature change projections. The region-specific model subset, like CORDEX, yields relatively wide coverage of both uncertainties, but this depends on the number of models used.

Code and data availability

CMIP5 multimodel dataset is publicly available via the website of Earth System Grid Federation (http://pcmdi9.llnl.gov/). Observation products are publicly available online via each website: CRU (https://crudata.uea.ac.uk/cru/data/hrg/cru_ts_4.01/), CPC (https://ftp.cpc.ncep.noaa.gov/precip/CPC_UNI_PRCP/), GPCC (https://www.dwd.de/EN/ourservices/gpcc/gpcc.html), PRECL (http://ftp.cpc.ncep.noaa.gov/precip/50yr/gauge/), CMAP (https://ftp-cpc.ncep.noaa.gov/precip/cmap), GPCP (ftp://meso.gsfc.nasa.gov/pub/gpcp-v2.2/), MSWEP (http://www.gloh2o.org). Code for analysis is available to the editor and reviewers for the purpose of the review. Public access to the code is limited due to the property of TOUGOU program, MEXT, Japan and, however, we can provide the code from the corresponding author upon request under the condition of collaborative research.

Author contribution

All authors conceptualized the study and participated in the discussion. RI analysed the data and prepared the manuscript and all authors revised the manuscript.

Competing interests

The authors declare no conflict of interest.

Acknowledgements

This work was conducted under the TOUGOU Program of the Ministry of Education, Culture, Sports, Science and Technology, Japan and ERTDF 2-1904 of the Environmental Restoration and Conservation Agency, Japan. The authors acknowledge Dr N. N. Ishizaki for useful suggestions. All figures are created by the Generic Mapping Tools (GMT; http://gmt.soest.hawaii.edu) ver. 4.5.12.

Supplement

Supplement 1 is a list of the CMIP5 models used in CORDEX and the number of regions where the model is selected.
Supplement 2 describes the regional classification defined in CORDEX.
Supplement 3 describes the skill score for annual mean model precipitation over land.
Supplement 4 describes the annual mean model temperature bias over land.
Supplement 5 describes the skill score for the annual mean model temperature over land.

References

Beck, H. E., Wood, E. F. Pan, M., Fisher, C. K., Miralles, D. G., van Dijk, A. I., McVicar, T. R. and Adler, R. F.: MSWEP V2 global 3-hourly 0.1° precipitation: methodology and quantitative assessment, Bull. Amer. Meteor. Soc., doi:10.1175/BAMS-D-17-0138.1, 2019.

Cannon, A. J.: Selecting GCM scenarios that span the range of changes in a multimodel ensemble: application to CMIP5 climate extremes indices, J. Clim., 28, 1260–1267, doi:10.1175/JCLI-D-14-00636.1, 2015.

Chen, M., Xie, P., Janowiak, J. E., and Arkin, P. A.: Global land precipitation: a 50-yr monthly analysis based on gauge observations, J. Hydrometeor., 3, 249–266, doi:10.1175/1525-7541(2002)003<0249:GLPAYM>2.0.CO;2, 2002.

Frieler, K., Lange, S., Piontek, F., Reyer, C. P. O., Schewe, J., Warszawski, L., Zhao, F., Chini, L., Denvil, S., Emanuel, K., Geiger, T., Halladay, K., Hurtt, G., Mengel, M., Murakami, D., Ostberg, S., Popp, A., Riva, R., Stevanovic, M., Suzuki, T., Volkholz, J., Burke, E., Ciais, P., Ebi, K., Eddy, T. D., Elliott, J., Galbraith, E., Gosling, S. N., Hattermann, F., Hickler, T., Hinkel, J., Hof, C., Huber, V., Jägermeyr, J., Krysanova, V., Marcé, R., Müller Schmied, H., Mouratiadou, I., Pierson, D., Tittensor, D. P., Vautard, R., van Vliet, M., Biber, M. F., Betts, R. A., Bodirsky, B. L., Deryng, D., Frolking, S., Jones, C. D., Lotze, H. K., Lotze-Campen, H., Sahajpal, R., Thonicke, K., Tian, H., and Yamagata, Y.: Assessing the impacts of 1.5 °C global warming – simulation protocol of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b), Geosci. Model Dev., 10, 4321–4345, doi:10.5194/gmd-10-4321-2017, 2017.

Giorgi, F. and Gutowski, W. J.: Regional Dynamical Downscaling and the CORDEX Initiative, Annu. Rev. Environ. Resour., 40, 467–490, doi:10.1146/annurev-environ-102014-021217, 2015.

Giorgi, F., Jones, C., and Asrar, G. R.: Addressing climate information needs at the regional level: the CORDEX framework, WMO Bull., 58,175–183, 2009.

Geoscientific Model Development Discussions
Open Access
EGU

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

Harris, I., Jones, P. D., Osborn, T. J. and Lister, D. H.: Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset, Int. J. Climatol., 34, 623–642, doi:10.1002/joc.3711, 2014.

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., and Adler, R. F.,: GPCP Version 2.2 Combined Precipitation Data Set. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, Boulder, CO., doi:10.5065/D6R78C9S, 2015. (Accessed 27 Dec 2015)

Knutti, R.: The end of model democracy? Clim. Change, 102, 395–404, doi://10.1007/s10584-010-9800-2, 2010.

McSweeney, C. F. and Jones, R. G.: How representative is the spread of climate projections from the 5 CMIP5 GCMs used in ISI-MIP? Clim. Serv., 1, 24–29, doi:10.1016/J.CLISER.2016.02.001, 2016.

McSweeney, C. F., Jones, R. G., Lee, R. W., and Rowell, D. P.: Selecting CMIP5 GCMs for downscaling over multiple regions, Clim. Dyn., 44, 3237–3260, doi:10.1007/s00382-014-2418-8, 2015.

Mendlik, T. and Gobiet, A.: Selecting climate simulations for impact studies based on multivariate patterns of climate change, Clim. Change, 135, 381–393, doi:10.1007/s10584-015-1582-0, 2016.

Reichler, T. and Kim, J.: How well do coupled models simulate Today's climate? Bull. Amer. Meteor. Soc., 89, 303–312, doi:10.1175/BAMS-89-3-303, 2008.

Schellnhuber, H. J., Frieler, K., and Kabat, P.: The elephant, the blind, and the intersectoral intercomparison of climate impacts, Proc. Natl. Acad. Sci., 111, 3225–3227, doi:10.1073/PNAS.1321791111, 2014.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., and Ziese, M.: GPCC Full Data Reanalysis Version 7.0: Monthly land-surface precipitation from rain gauges built on GTS based and historic data, research data archive at the national center for atmospheric research, computational and information systems laboratory, doi:10.5065/D6000072, 2016. (Accessed 6 JAN 2015)

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., and Hsu, K.-L.: A review of global precipitation data sets: data sources, estimation, and intercomparisons, Rev. Geophys., 56, 79–107, doi:10.1002/2017RG000574, 2018.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res. Atmos., 106, 7183–7192, doi:10.1029/2000JD900719, 2001.

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

The intersectoral impact model intercomparison project: ISIMP project design and simulation protocol, 2018. Retrieved from https://www.isimip.org/protocol/#isimip-fast-track on 23 Jan. 2019.

Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O., and Schewe, J.: The inter-sectoral impact model intercomparison project (ISI–MIP): project framework, Proc. Natl. Acad. Sci., 111, 3228–3232, doi:10.1073/pnas.1312330110, 2014.

https://doi.org/10.5194/gmd-2019-143
Preprint. Discussion started: 28 June 2019

Geoscientific Model Development Discussions
Open Access EGU

Xie, P., Chen, M., and Shi, W.: CPC unified gauge-based analysis of global daily precipitation, Preprints, 24th Conf. on Hydrology, Atlanta, GA, Amer. Meteor. Soc., 2.3A, 2010. Retrieved from https://ams.confex.com/ams/90annual/techprogram/paper_163676.htm on 23 Jan 2019.

Xie, P. and Arkin, P. A.: Global precipitation: a 17-Year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs, Bull. Amer. Meteor. Soc., 78, 2539–2558, doi:10.1175/1520-0477(1997)078<2539:GPAYMA>2.0.CO;2, 1997.

Table 1: Number of CMIP5 models used in the CORDEX regions.

Region		Region	
South America	9	Central Asia	3
Central America	10	Australasia	13
North America	6	Antarctica	9
Africa	15	Arctic	5
Europe	13	Mediterranean	5
South Asia	15	Middle East and North Africa (MENA)	5
East Asia	7	Southeast Asia	12

https://doi.org/10.5194/gmd-2019-143
Preprint. Discussion started: 28 June 2019

Geoscientific Model Development Discussions
Open Access

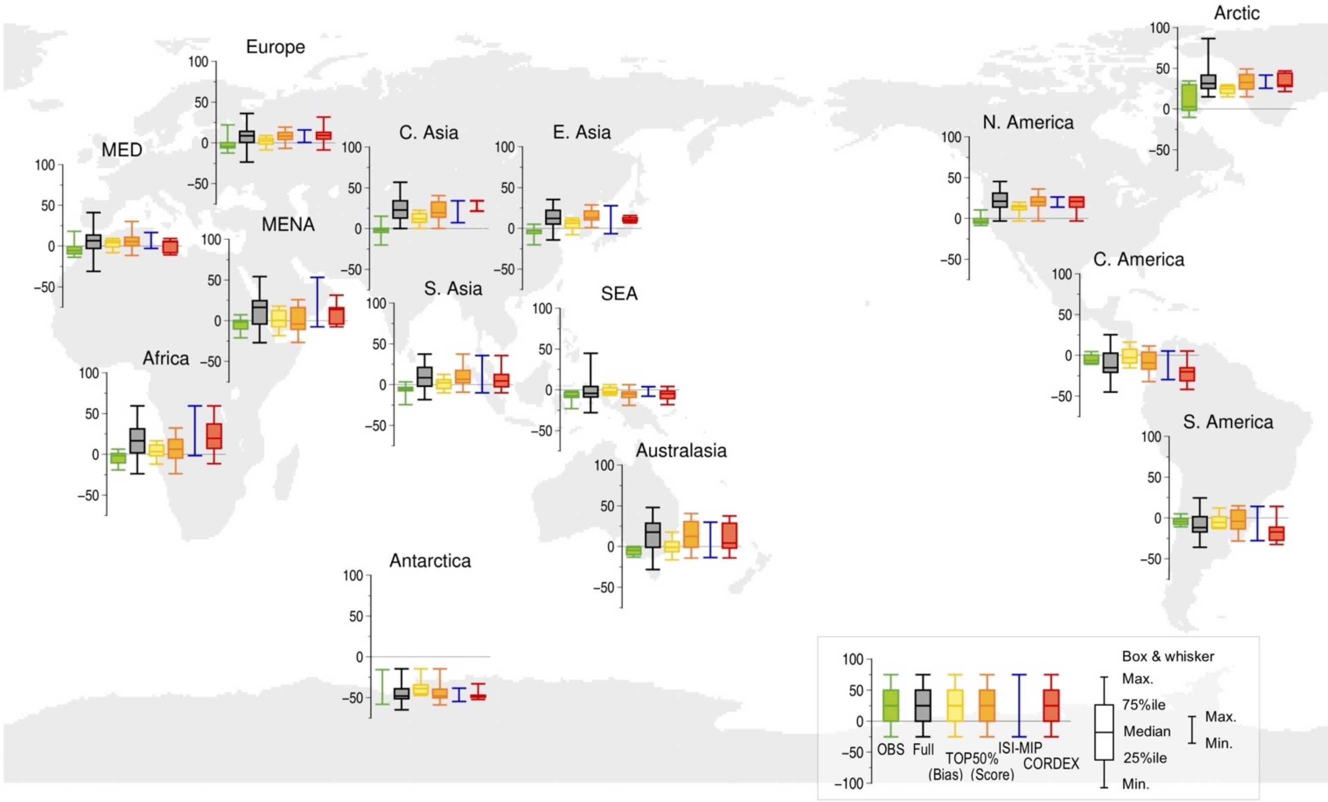

Figure 1: Normalized annual mean model precipitation bias over land from the GPCC reference data (%). The bias was normalized using the areal mean of the reference data. The whiskers of the box plots show the range between the maximum and the minimum biases. The boxes and the lines within the boxes indicate the 25th to 75th percentile range and the median, respectively. Green plots indicate the deviations of six observation data from the reference data. The other plots indicate the model bias in the full set of 49 CMIP5 model set (black), the model sets with a bias with is less than the 50th percentile of biases of the full set (yellow), the model sets with Taylor's skill score with is larger than the 50th percentile of the scores of the full set (orange), and the model sets selected for ISIMIP (blue) and CORDEX (red).

https://doi.org/10.5194/gmd-2019-143
Preprint. Discussion started: 28 June 2019

Geoscientific Model Development
Discussions
Open Access

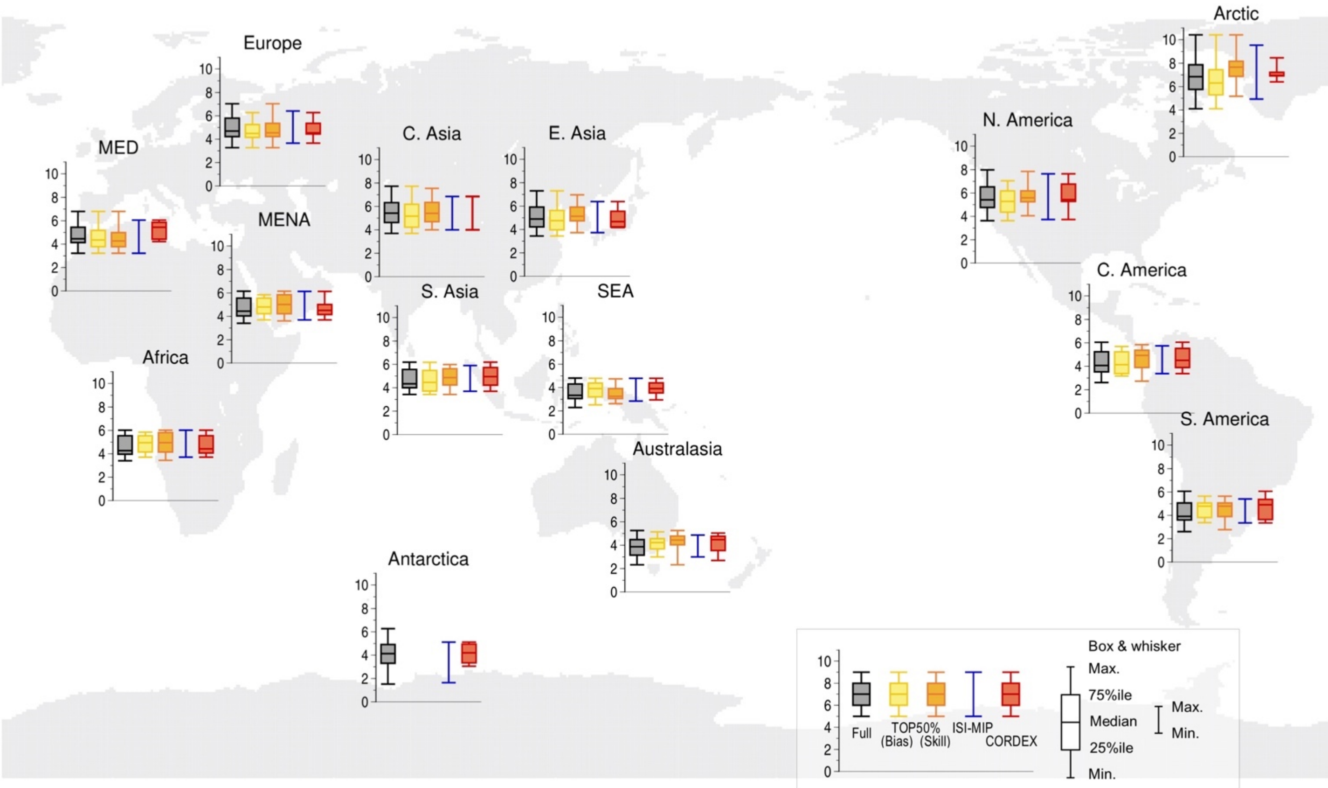

Figure 2: Annual mean temperature increments in the future climate projection (K). The whiskers of the box plots show the range between the maximum and the minimum biases. The boxes and the lines within the boxes show the 25th to 75th percentile range and the median, respectively. Box plots indicate the model bias in the full set of 42 CMIP5 models (black), the model sets with the top 50% of the CMIP5 models for the bias (yellow) or Taylor's skill score (orange), and the model sets selected for ISIMIP (blue) and CORDEX (red).

https://doi.org/10.5194/gmd-2019-143
Preprint. Discussion started: 28 June 2019

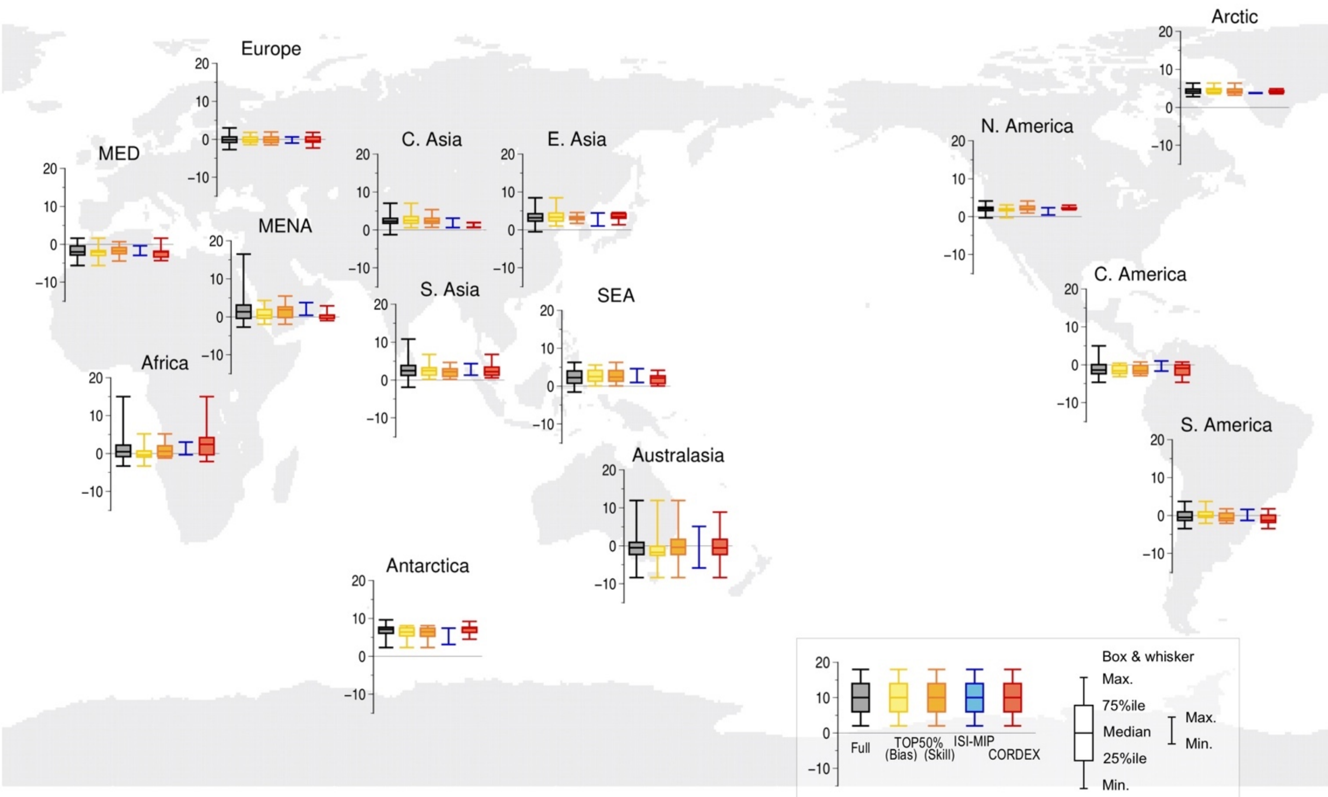

Figure 3: As for Figure 2, but for the projected change in annual mean precipitation scaled to the regional mean temperature increment over the land (% K^{-1}). The change in global precipitation was scaled to the global temperature increment.

https://doi.org/10.5194/gmd-2019-143
Preprint. Discussion started: 28 June 2019

Geoscientific Model Development
Discussions

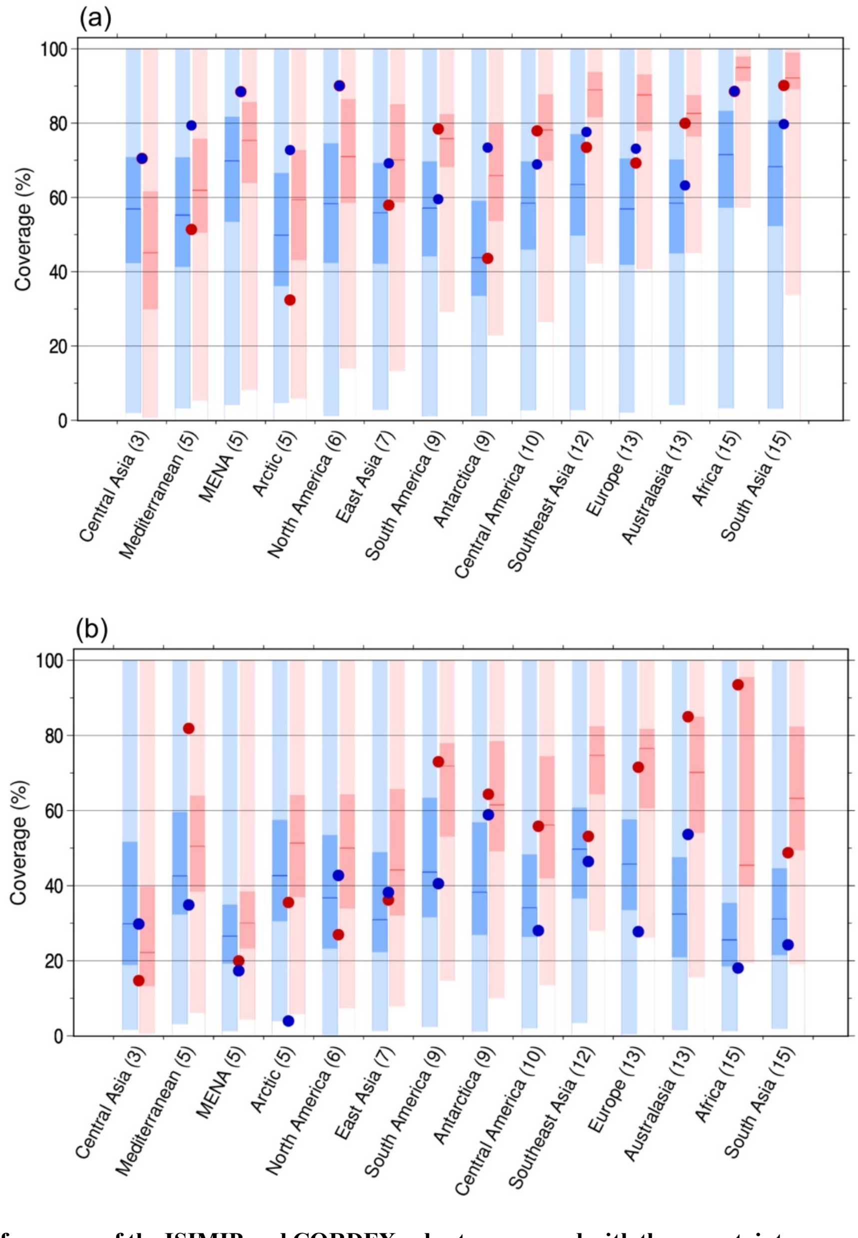

Figure 4: Coverage performance of the ISIMIP and CORDEX subsets compared with the uncertainty range of the full set of CMIP5 models for (a) annual mean temperature increment and (b) precipitation change scaled to the regional mean temperature increment. Blue and red dots indicate the coverage in ISIMIP and CORDEX, respectively for each region. Blue bars indicate the spread of coverage (FRA) when four models, as in ISIMIP, are selected randomly in 10,000 times. Red bars indicate the spread when randomly selecting the same number of models as in CORDEX; e.g., 10 models in Central America. The full range of the coloured bars indicates the minimum to maximum coverage. Dark blue and red bars indicate the 25th to 75th percentile range of the FRA spread. Horizontal lines in the dark blue and red regions indicate the median. Numbers in parentheses are the number of models used in CORDEX.