Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-107
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 25 June 2018

Geoscientific Model Development Discussions

Bias correction of multi-ensemble simulations from the HAPPI model intercomparison project

Fahad Saeed[1,2], Ingo Bethke[3], Stefan Lange[4], Ludwig Lierhammer[5], Hideo Shiogama[6], Dáithí A. Stone[7], Tim Trautmann[8], Carl-Friedrich Schleussner[1,4,9]

1. Climate Analytics, Berlin, Germany
2. Center of Excellence for Climate Change Research, King Abdulaziz University, Jeddah, Saudi Arabia
3. Uni Research Climate, Bjerknes Centre for Climate Research, Bergen, Norway
4. Potsdam Institute for Climate Impact Research, Potsdam, Germany
5. Deutsches Klimarechenzentrum, Hamburg, Germany
6. Center for Global Environmental Research, National Institute for Environmental Studies, Tsukuba, Japan
7. Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA
8. Institute of Physical Geography, University of Frankfurt, Frankfurt D-60054, Germany
9. IRITHESys, Humboldt University, Berlin

Abstract

Prior to using climate data as input for sectoral impact models, statistical bias correction is commonly applied to correct climate model data for systematic deviations. Different approaches have been adopted for this purpose, however the most common are those based on the transfer functions, generated to map the distribution of the simulated historical data to that of the observations. Here, we present results of a novel bias correction method, developed for Inter-Sectoral Impact Model Intercomparison Project Phase 2b (ISIMIP2b) and applied to outputs of different GCMs generated within the HAPPI (Half A degree Additional warming, Projections, Prognosis and Impacts) project. We have employed various analysis measures including mean seasonal differences, ensemble variability, annual cycles, extreme indices as well as a global hydrological model to assess the performance of ISIMIP2b bias correction technique. The results indicate substantial improvements after the application of

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-107
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 25 June 2018

Geoscientific Model Development
Discussions
Open Access

bias correction when compared against observational data. Moreover, the extreme indices as well as output of global hydrological model also reveal a marked improvement. At the same time, the ensemble spread of the original data is preserved after the application of bias correction. We find that the bias corrected HAPPI data can provide a reliable basis for sectoral climate impact projections.

Introduction

Global climate models (GCMs) are the most commonly used tools to assess changes in future climate. However, due to their coarser resolution (~ 200 km), many regional and local scale climate features go beyond the scope of GCMs and give rise to biases in different climate variables against observation in historical period (Flato *et al* 2013). Besides resolution, imperfect representation of physical processes, incorrect initialization or errors in the parameterization chain etc., can also act as contributing factors to these biases (Ehret *et al* 2012). Global models are benchmarked against global datasets and a set of comprehensive variables, not for specific regions or sector relevant outputs. Despite biases in absolute values, relative changes in GCMs have been shown to resemble observed trends well (Flato *et al* 2013). However, when GCM output is directly used to force impact models (e.g. crop models, hydrological models etc.), which are often certain processes are based on specific thresholds, these absolute biases limit the applicability of the climate data by affecting the calibration and validation process, which is an important aspect of impact modeling (Warszawski et al 2013).

Different approaches have been adopted to overcome the biases generated by GCMs. One of them is the stochastic downscaling in which a functional relationship is established between the most robust and reliable fields provided by GCMs and the observed meteorological variables for a specific region. However, this approach is criticized due to one critical assumption implicit to all statistical downscaling methods which is 'statistical stationarity' (Dixon et al 2016). Moreover, in certain cases (e.g. ISIMIP experiment) in which the global impact models are forced with the GCM data, stochastic downscaling does not remain a very useful option (Warszawski et al 2013). Another physically more consistent approach to overcome these biases is regional climate modelling (Giorgi and Mearns 1999). A regional

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-107
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 25 June 2018

climate model (RCM) can not only bridge the resolution gap for the better representation of regional scale features, but it also gives the flexibility for the representation of characteristic processes of a certain region (Saeed et al 2012). However, despite improvements during recent years, the output of RCMs is still afflicted with biases (systematic errors) to a degree that preclude their direct use in the impact models (Piani et al 2010).

Another well-established approach to overcome the biases of GCMs to make them suitable for their use in impact models is the post-processing of GCM data by correcting with and towards observations (Sippel et al 2016). This approach has become standard in impact studies and is known as bias correction (or bias adjustment). There are different methods which have been published in earlier literature ranging from simple adjustment of the means to flexible, potentially multivariate, quantile mapping approaches (Maraun et al 2017). However, like other methods, there have been many problems associated with the bias correction methods. Few of those already identified include alteration of spatiotemporal field consistency, relations among variable and violation of conservation laws (Ehret et al 2012). Therefore this approach is prone to misuse and hence a careful practice of the bias correction is generally recommended (Maraun et al 2017).

For this study, we have used the applied bias correction technique developed for ISIMIP (Inter Sectoral Impact Modelling Inter-comparison Project) and recently extended for ISIMIP2b (Frieler et al 2016), hereafter referred to as ISIMIP2b-BC approach. ISIMIP bias correction is a trend preserving statistical bias correction approach which adjusts the monthly mean and daily variability of simulated climate data to observations (Hempel et al 2013)(Lange 2017). Here we applied the ISIMIP2b-BC to GCM output from the HAPPI (Half a degree Additional warming, Projection, Prognosis and Impacts) model intercomparison project (Mitchell et al. 2017).

Following the adoption of the Paris Agreement, there has been a growing interest for quantifying impacts at discrete levels of global mean temperature (GMT) increase such as 1.5°C and 2.0°C above pre-industrial levels (Schleussner et al 2016). By now, there has been a dearth of research to address this issue because many available experiments in the CMIP (Couple Model Inter-comparison Project) are not specifically designed for informing this

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-107
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 25 June 2018

report. This has led to the HAPPI experiment, an international effort specifically designed to provide a framework for the generation of climate data describing how the climate, in particular extreme weather, might differ from the present day in worlds those are 1.5° and 2.0°C warmer than pre-industrial conditions (Mitchell et al 2017). The quasi-stationary multi-ensemble design also allows for assessments of climate change signals against a highly variable background.

In this paper, we present the performance of ISIMIP2b-BC in correcting the biases associated with different variables from four different GCMs analyzed in the HAPPI experiment. Besides validation, this paper also focuses on whether the application of ISIMIP2b-BC preserves the ensemble spread of the original simulations. This paper also provides a reference document for the future users of HAPPI data.

Data and Methodology

The HAPPI modelling setup considers three time periods (historical, +1.5°C and +2.0°C) each spanning over 10 years (Mitchell et al 2017). All the runs are executed atmosphere only under prescribed sea-surface temperatures and sea-ice forcing conditions. For each of the three periods considered, multi-ensemble GCM realizations are provided. The 'historical period' taken through 2006-2015 for HAPPI. We employed ISI-MIP2b bias correction methodology to bias correct HAPPI data in order to improve the representation of regional features using (Hempel et al 2013), (Lange 2017). Following the modelling protocol of the Intersectoral Impact Model Intercomparison Project (Frieler et al 2016), the resultant projections are re-gridded to a 0.5°x0.5°C regular grid and then bias corrected using the EWEMBI dataset (Lange 2017). In total 20 ensemble members per GCM have been bias corrected.

In previous applications, bias correction was generally applied on a single transient simulation by generating transfer functions for a variable in the base period. These transfer functions were then applied to the projected data of the transient simulation. In HAPPI, the transfer functions have been derived for one long-term member over a climatological time scale of more than 25 years (Table 1). This long-time frame is chosen in order to capture the effects of natural variability. The GCM specific transfer function is then applied to each ensemble member (of 10 years period) to attain the bias corrected data. It is noteworthy that the

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-107
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 25 June 2018

Geoscientific Model Development Discussions
Open Access

individual ensemble members of the different periods are initialised stochastically. This implies that numbering of ensemble members is purely statistical and does not imply any physical relationship in terms of natural variability. Table 2 shows the HAPPI variables which are bias corrected using ISIMIP2b-BC in the present study.

Different analysis approaches, including mean seasonal differences, mean standard deviation across ensemble, annual cycle and extreme indices, are employed to find out the performance of ISIMIP2b-BC. In addition, both the original (non-bias corrected) and the bias corrected data sets have been used to force WaterGAP (Müller Schmied *et al* 2014), a global water use and availability model, which calculates freshwater fluxes and storages for all continents except for Antarctica. A detailed and comprehensive model description of WaterGAP can be found in (Müller Schmied *et al* 2014) and (Müller Schmied 2017). The version used here is WaterGAP2.2c.

Results

Mean Seasonal Differences:

Seasonal plots of ISIMIP2b-BC of 5 variables for different GCMs are shown in Figure 1, which are plotted by taking the mean of 20 ensemble members for each variable and compared against the EWEMBI dataset (2001-2010). Moreover, the complete results for all the variables across all the 4 GCMs are presented in Figure S1-S8. MIROC5 exhibits the strongest bias for temperature over most parts of the globe (Figure 1 and Figure S1). For certain regions, like Russia, the sign of the seasonal bias is swapped between MAM and JJA (Figure 1). Moreover, the direction of the bias is also different among different GCMs over different regions (Figure S1). Application of ISIMIP2b-BC improved the results drastically by bringing the biases down to 0.5°C over most of the globe, irrespective of the direction of the bias. A similar behavior is observed for the results of tasmin and tasmax where large biases are reasonably corrected by ISIMIP2b-BC (Figure S2 and S3). A common feature across all the GCMs is are systematic biases over the high latitude sea ice regions (warm bias over the Arctic, cold bias over Antarctica) for the three temperature variables (i.e. tas, tasmin and tasmax) during boreal autumn and winter, which stays consistent even after the application of ISIMIP2b-BC. This feature might be the result of the substantial alterations in sea ice coverage between the bias correction

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-107
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 25 June 2018

period and the recent past (decrease in the Arctic, increase in the Antarctic (Nghiem et al. (2016)) that will ultimately alter affect the applicability of the transfer function. Despite visible improvements in terms of model bias, we identify the high latitudes as a region where the assumption of a time invariant transfer function is of limited validity.

A similar behavior is observed for precipitation as in Figure 1 for ECHAM6 (and full set of models in Figure S4). Once again, the bias correction yielded substantial improvement of precipitation. A few differences remain over the tropical and subtropical arid regions which is attributed to the lack of precipitation in those regions, and therefore a small difference appears substantial when looking at relative changes for certain seasons. The results for other GCMs also show similar patterns (Figure S4).

CAM4-2degree shows an overestimation of near surface wind speed over most of the land area which is substantially reduced by ISIMIP2b-BC across all the seasons (Figure 1). ECHAM6 and NorESM1 show better results for surface wind speeds which is further improved after the application of ISIMIP2b-BC (Figure S5). A special case is MIROC5 which shows underestimation of near surface wind speeds over tropical forested regions of Africa and South America, which remains the same after the application of ISIMIP2b-BC. It appears that in MIROC5, the wind speeds are calculated inside the rainforest canopy and, therefore, are one to two orders of magnitude too weak in the original model data (Tatsuo Suzuki (Personal Communication)).

For other variables, such as rsds, rlds and hurs, ISIMIP2b-BC has also done a very satisfactory job in reducing the biases for all the seasons across all the four GCMs (Figure 1 and Figure S5-S8).

Seasonal variability across the ensemble:

A key research question for HAPPI simulations are changes in the extremes. Although the ISIMIP2b-BC methodology should leave the stochastic ensemble variability unaffected, we therefore explicitly assessed its effects. The seasonal ensemble standard deviation of daily data for selected variables is presented in Figure 2. The complete results for all the variables across all the 4 GCMs are presented in Figure S9-S16.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-107
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 25 June 2018

All the GCMs have simulated higher temperature variability over higher latitudes as compared to tropics, especially during the corresponding winter season (Figure 2 and Figure S9). ISIMIP2b-BC has ably kept the ensemble variability for all the GCMs. There are few small differences between the original and bias corrected data over some region, however the overall spatial pattern of the variability is reasonably kept after the application of ISIMIP2b-BC. Similar results are obtained for tasmax and tasmin, and are presented in Figure S10 and S11 respectively.

In comparison to temperature variable, precipitation shows different behavior with the monsoon regions showing highest variability in the original simulation with varying magnitude among different GCMs (Figure 2 and Figure S12). The variability is prominent especially during the rainy season. ISIMIP2b-BC has largely kept the variability of the original data with consistent spatial pattern, however there are few differences at regional scale (for example over the South Asian Summer Monsoon region (SASM)). Looking at the summer season plot (JJA) in Figure S12, it can be noticed that all the GCMs have different spatial patterns of variability over land as well as Indian ocean in the original data. High and low variability over India can be seen for MIROC5 and ECHAM6 respectively. Moreover, magnitude and location of high variability spot over the Arabian Sea vary between the GCMs. Application of ISIMIP2b-BC resulted in consistent pattern for all the four GCMs with highest variability occurring over central India for JJA. Therefore, change in variability for precipitation arise as the result of the modification of highly dynamical features of the climate system, such as Monsoon, that are not equally represented in the GCMs. While ISIMIP2b-BC is affecting the ensemble variability in regions where dynamical features dominate, this to some extent is more correcting for misrepresentations of these features in the underlying GCMs without changing their dynamical characteristics. However, it presents a limitation to the applicability of a purely thermodynamic and statistical correction in the presence of strong dynamical features.

For rsds, the highest variability follows the seasonal march of the sun in northern and southern hemisphere in the original data for all the GCMs (Figure 2 and Figure S14). ISIMIP2b-BC has once again done a good job in keeping the seasonal spatial pattern as well as magnitude of the variability. However, ISIMIP2b-BC show a small reduction in the ensemble variability of rlds towards higher latitudes, by keeping the spatial pattern consistent (Figure S15).

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-107
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 25 June 2018

Geoscientific Model Development Discussions
Open Access
EGU

Furthermore, a satisfactory performance of ISIMIP2b-BC in correcting the hurs and sfcWind fields have been observed which can be seen from the Figure S13 and S16 respectively.

Annual Cycle:

In order to further ascertain the ability of ISIMIP2b-BC in keeping the spread of the original simulation, annual cycle of all the 8 variables have been plotted in Figure 3 and Figure S17 to S20 averaged only over the land areas. The temperature (tas, tasmax, tasmin) and radiation (rsds, rlds) variables show small ensemble spread for both original and bias corrected data. ISIMIP2b-BC has done a satisfactory job in bringing these variables closer to the EWEMBI annual curve while keeping the spread reasonably intact. For other variables ISIMIP2b-BC not only corrects the bias, but also the shape of the mean curve (e.g. hurs for MIROC5 and pr for CAM4-2degree in Figure 3). ECHAM6 has reasonably captured the annual cycle of precipitation over the land areas, however ISIMIP2b-BC has shown a marked improvement for the rest of the three GCMs. GCMs have also shown varying performances in capturing the annual cycle of sfcWind with CAM4-2degree and NorESM1 a systematic overestimation which is also corrected by ISIMIP2b-BC quite reasonably (Figure 3 and Figure S17 to S20).

Extreme Indices:

Climate extremes are one of the parameters which are likely to get affected by bias corrections especially if the resolution simulation and the observations are considerably different from each other (Maraun et al 2017). Moreover, the climate extremes are also reported to perform not as good as long-term means and trends after the application of ISIMIP bias correction (Hempel et al 2013). Therefore, it becomes imperative to analyze how does ISIMIP2b-BC perform with HAPPI data in representing climate extremes. For this purpose, we consider different extreme indices based on the recommendation of the Expert Team on Climate Change Detection and Indices (Zhang et al 2011). For each extreme, all the indices are calculated for each individual ensemble member for both bias corrected and original data and the ensemble mean is taken afterwards across all the four GCMs. The presented results in Figure 4 to 7 are the differences from EWEMBI data.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-107
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 25 June 2018

Geoscientific Model Development
Discussions

Consecutive Summer Days (CSD):

CSD is defined as the number of instances for which the maximum temperature remains greater than certain threshold for consecutive period of 5 days or more. Results for each model are presented in Figure 4 for the thresholds 30°C, 35°C and 40°C. Figure 4 shows general improvement in the simulation of CSD after the application of ISIMIP2b-BC, bringing the results closer to the EWEMBI data. Understandably, most of the differences occur in the lower latitudes where these high temperature thresholds are attained more often. ECHAM6 and MIROC5 simulate higher number of CSD instances than EWEMBI which are reasonably corrected by ISIMIP2b-BC.

Consecutive Frost Days (CFD):

CFD is defined as the number of instances for which the minimum temperature remains less than 0°C for consecutive period of 5 days or more. Results for each model for CFD are presented in Figure 5. Contrary to CSD, most of the differences occur in the higher latitudes for CFD. Once again, ECHAM6 has simulated much higher number of CFD than EWEMBI data which is corrected to a large extent after the application of ISIMIP2b-BC. Altogether, ISIMIP2b-BC has brought the CFD results closer to EWEMBI.

Consecutive Dry Days (CDD):

CDD is defined as the number of instances for which the daily precipitation remains less than certain threshold for consecutive period of 5 days or more. Figure 6 shows the CDD results for thresholds of 1mm/day, 1.5 mm/day and 2.0 mm/day. Once again, improvements in CDD after the application of ISIMIP2b-BC can be seen. Although irrelevant in terms of impact model application, however most obvious improvement is in the correction of overestimation and underestimation for MIROC5, over most of the ocean region. Improvements can also be seen over land regions for all the models.

Consecutive Wet Days (CWD):

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-107
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 25 June 2018

CWD is defined as the number of instances for which the daily precipitation remains higher than certain thresholds for consecutive period of 5 days or more. Results for CWD days for thresholds of 1 mm/day, 2 mm/day and 3 mm/day are shown in Figure 7. Like other indices discussed earlier, a general improvement in the results after the application of ISIMIP2b-BC can be witnessed. Similar to CDD, application of ISIMIP2b-BC has shown a marked improvement in the MIROC5 simulation over the ocean areas. Moreover, improvement can also be noticed in the results of CWD for all the models over land areas.

Simulated Streamflow:

The streamflow is calculated by WaterGAP for the bias corrected and non-bias corrected ensemble members and compared to a reference run forced with EWEMBI data. In figure 8 we present the simulation results at outlets of three major river systems (Rhine, Mississippi and Amazon). Due to the bias correction, the annual discharge and the seasonal flows fit better to the reference simulations. Next to the general better fit of discharge the range of the ensemble members is predominantly maintained in the three basins. Some further streamflow simulation results can be found in figure S21, which also support the aforementioned findings.

Summary and Conclusions:

Bias correction of climate model output has remained a controversial issue among the scientists. A vast variety of bias correction methods are already in use in the field of climate and impact modeling, however many problems with the bias correction methods have also been identified. For example, bias correction not only alters the underlying physical characteristics among variables, but also e.g. for extreme tails for the same variable (Sippel et al., 2016). On the other hand, climate scientists are confronted with growing pressure to translate their modeling information into informed adaptation decision by using the impact models. Impact models, in which certain processes are based on threshold, are sensitive to large systematic biases in the climate data making bias correction an essential step. We have assessed the performance of ISIMIP2b-BC in correcting 8 variables simulated under HAPPI project across 4 different GCMs.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-107
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 25 June 2018

Geoscientific Model Development Discussions
Open Access

An overall satisfactory performance of ISIMIP2b-BC is witnessed in correcting the seasonal means as well ensemble spread across the GCMs. A marked improvement in the annual cycle of the variable is also achieved. The results for extreme indicators also come closer to the observations although they have not been explicitly bias corrected. All these improvements are translated into the results of the impact model (WaterGAP), which show significant improvements after the application of ISIMIP2b-BC. Few limitation have also been identified, for example ISIMIP2b-BC failed to correct the deviations in strong dynamical features of the monsoon regions.

The results of this study indicate that the application of bias correction technique is mandatory when forcing the impact models with the data of climate models. Besides an orthodox validation, this study will also serve as a reference document to analyze the performance of bias corrected against the original (non-bias corrected) data for the future HAPPI data users.

Data availability:

The data used in this paper is freely available from the link http://portal.nersc.gov/c20c/data/ClimateAnalytics/

Acknowledgements:

The authors would like to thank the HAPPI initiative and all participating modelling groups that have provided data. This research used science gateway resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. CFS and FS acknowledge support by the German Federal Ministry of Education and Research (01LS1613A) and by the German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (16_II_148_Global_A_IMPACT). NorESM contributions have received support from the Research Council of Norway (261821) and UNINETT Sigma2 (ns9082k). DS was supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, under contract number DE-AC02-05CH11231. The MIROC5 simulations were performed using the Earth Simulator at JAMSTEC and the NEC SX at NIES.

Moreover, MIROC5 contributions have been supported by the TOUGOU project of MEXT, Japan and by the Environment Research and Technology Development Fund (2-1702) of the Environmental Restoration and Conservation Agency, Japan.

References:

Bentsen M, Bethke I, Debernard J B, Iversen T, Kirkevåg A, Seland Ø, Drange H, Roelandt C, Seierstad I A, Hoose C, and Kristjánsson J E 2013 The Norwegian Earth Sys- tem Model, NorESM1-M – Part 1: Description and basic evalu- ation of the physical climate, Geosci. Model Dev., 6, 687–720, doi:10.5194/gmd-6-687-2013.

Dixon K W, Lanzante J R, Nath M J, Hayhoe K, Stoner A, Radhakrishnan A, Balaji V and Gaitán C F 2016 Evaluating the stationarity assumption in statistically downscaled climate projections: is past performance an indicator of future results? *Clim. Change* **135** 395–408

Ehret U, Zehe E, Wulfmeyer V, Warrach-Sagi K and Liebert J 2012 HESS Opinions "should we apply bias correction to global and regional climate model data?" *Hydrol. Earth Syst. Sci.* **16** 3391–404

Flato G, Marotzke J, Abiodun B, Braconnot P, Chou S C, Collins W, Cox P, Driouech F, Emori S, Eyring V, Forest C, Gleckler P, Guilyardi E, Jakob C, Kattsov V, Reason C and Rummukainen M 2013 Evaluation of Climate Models *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* ed T F Stocker, D Qin, G-K Plattner, M Tignor, S K Allen, J Boschung, A Nauels, Y Xia, V Bex and P M Midgley (Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press) pp 741–866

Frieler K, Betts R, Burke E, Ciais P, Denvil S, Deryng D, Ebi K, Eddy T, Emanuel K, Elliot J, Galbraith E, Gosling S, Halladay K, Hatermann F, Hickler T, Hinkel J, Huber V, Jones C, Krysanova V, Lange S, Lotze H, Lotze-Campen H, Mengel M, Mouratiadou I, Müller Schmied H, Ostberg S, Piontek F, Popp A, Reyer C, Schewe J, Stevanovic M, Suzuki T, Thonicke K, Tian H, Tittensor D T, Vautard R, van Vliet M, Warszawski L and Zhao F 2016 Assessing the impacts of 1.5°C global warming - simulation protocol of the Inter-Sectoral Model Intercomparison Project (ISIMIP2b) *Geosci. Model Dev. Discuss.*

Giorgi F and Mearns L O 1999 Introduction to special section: Regional climate modeling revisited *J. Geophys. Res. Atmos.* **104** 6335–52

Hempel S, Frieler K, Warszawski L, Schewe J and Piontek F 2013 A trend-preserving bias correction – the ISI-MIP approach *Earth Syst. Dyn.* **4** 219–36

James R, Washington R, Schleussner C-F, Rogelj J and Conway D 2017 Characterizing half-a-degree difference: a review of methods for identifying regional climate responses to global warming targets *Wiley Interdiscip. Rev. Clim. Chang.* e457

Lange S 2017 Bias correction of surface downwelling longwave and shortwave radiation for the EWEMBI dataset *Earth Syst. Dyn. Discuss.* **2017** 1–30

Maraun D, Shepherd T G, Widmann M, Zappa G, Walton D, Gutiérrez J M, Hagemann S, Richter I, Soares P M M, Hall A and Mearns L O 2017 Towards process-informed bias correction of climate change simulations *Nat. Clim. Chang.* **7** 664–773

Mitchell D, AchutaRao K, Allen M, Bethke I, Beyerle U, Ciavarella A, Forster P M, Fuglestvedt J, Gillett N, Haustein K, Ingram W, Iversen T, Kharin V, Klingaman N, Massey N, Fischer E, Schleussner C-F, Scinocca J, Seland Ø, Shiogama H, Shuckburgh E, Sparrow S, Stone D, Uhe P, Wallom D, Wehner M and Zaaboul R 2017 Half a degree additional warming, prognosis and projected impacts (HAPPI): background and experimental design *Geosci. Model Dev.* **10** 571–83

Müller Schmied H 2017 Evaluation, modification and application of a global hydrological model *Frankfurt Hydrol. Pap.* **16**

Müller Schmied H, Eisner S, Franz D, Wattenbach M, Portmann F T, Flörke M and Döll P 2014 Sensitivity of simulated global-scale freshwater fluxes and storages to input data, hydrological model structure, human water use and calibration *Hydrol. Earth Syst. Sci.* **18** 3511–38

Neale R B, Richter J, Park S, Lauritzen P H., Vavrus S J, Rasch P J, and Zhang M 2013 The mean climate of the Community Atmosphere Model (CAM4) in forced SST and fully coupled experiments, J. Climate, 26, 5150–5168.

Nghiem, S., I. Rigor, P. Clemente-Colon, G. Neumann, and P. Li (2016), Geophysical constraints on the Antarctic sea ice cover, Remote Sens. Environ., 181, 281–292.

Piani C, Weedon G P, Best M, Gomes S M, Viterbo P, Hagemann S and Haerter J O 2010 Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models *J. Hydrol.* **395** 199–215

Saeed F, Hagemann S and Jacob D 2012 A framework for the evaluation of the South Asian summer monsoon in a regional climate model applied to REMO *Int. J. Climatol.* **32** 430–40

Schleussner C-F, Rogelj J, Schaeffer M, Lissner T, Licker R, Fischer E M, Knutti R, Levermann A, Frieler K and Hare W 2016 Science and policy characteristics of the Paris Agreement temperature goal *Nat. Clim. Chang.* **6** 827–835

Shiogama H, Watanabe M, Imada Y, Mori M, Kamae Y, Ishii M and Kimoto M 2014 Attribution of the June-July 2013 heat wave in the southwestern United States, SOLA, 10, 122–126.

Sippel S, Otto F E L, Forkel M, Allen M R, Guillod B P, Heimann M, Reichstein M, Seneviratne S I, Thonicke K and Mahecha M D 2016 A novel bias correction methodology for climate impact simulations *Earth Syst. Dyn.* **7** 71–88

Warszawski L, Frieler K, Huber V, Piontek F, Serdeczny O and Schewe J 2013 The Inter-Sectoral Impact Model Intercomparison Project (ISI-MIP): Project framework. *Proc. Natl. Acad. Sci. U. S. A.* 1–5

Zhang X, Alexander L, Hegerl G C, Jones P, Tank A K, Peterson T C, Trewin B and Zwiers F W 2011 Indices for monitoring changes in extremes based on daily temperature and precipitation data *Wiley Interdiscip. Rev. Clim. Chang.* **2** 851–70

Table 1: Table of GCMs which are bias corrected in HAPPI, with their specifications.

Model	Horizontal Resolution	Time-Period for construction of Transfer Functions	References
CAM4	2x2º	1979-2005	Neale et al. (2013)
MIROC5	150x150 km	1979-2010	Shiogama et al. (2014)
MPI-ECHAM6.3	T63	1979-2010	--
NorESM1-Happi	1.25x0.94º	1986-2010	Bentsen et al. (2013)

Table 2: List of HAPPI variables which are bias corrected using ISIMIP2b-BC.

Variable	Short name	Unit
Precipitation	pr	Kg m^2 s^{-1}
Near surface air temperature	tas	K
Near surface maximum air temperature	tasmax	K
Near surface minimum air	tasmin	K

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-107
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 25 June 2018

Geoscientific Model Development Discussions Open Access EGU

temperature		
Surface downwelling longwave radiation	rlds	W m^{-2}
Surface downwelling shortwave radiation	rsds	W m^{-2}
Near surface wind speed	sfcWind	m s^{-1}
Near surface relative humidity	hurs	%

1

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-107
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 25 June 2018

Geoscientific Model Development Discussions
Open Access

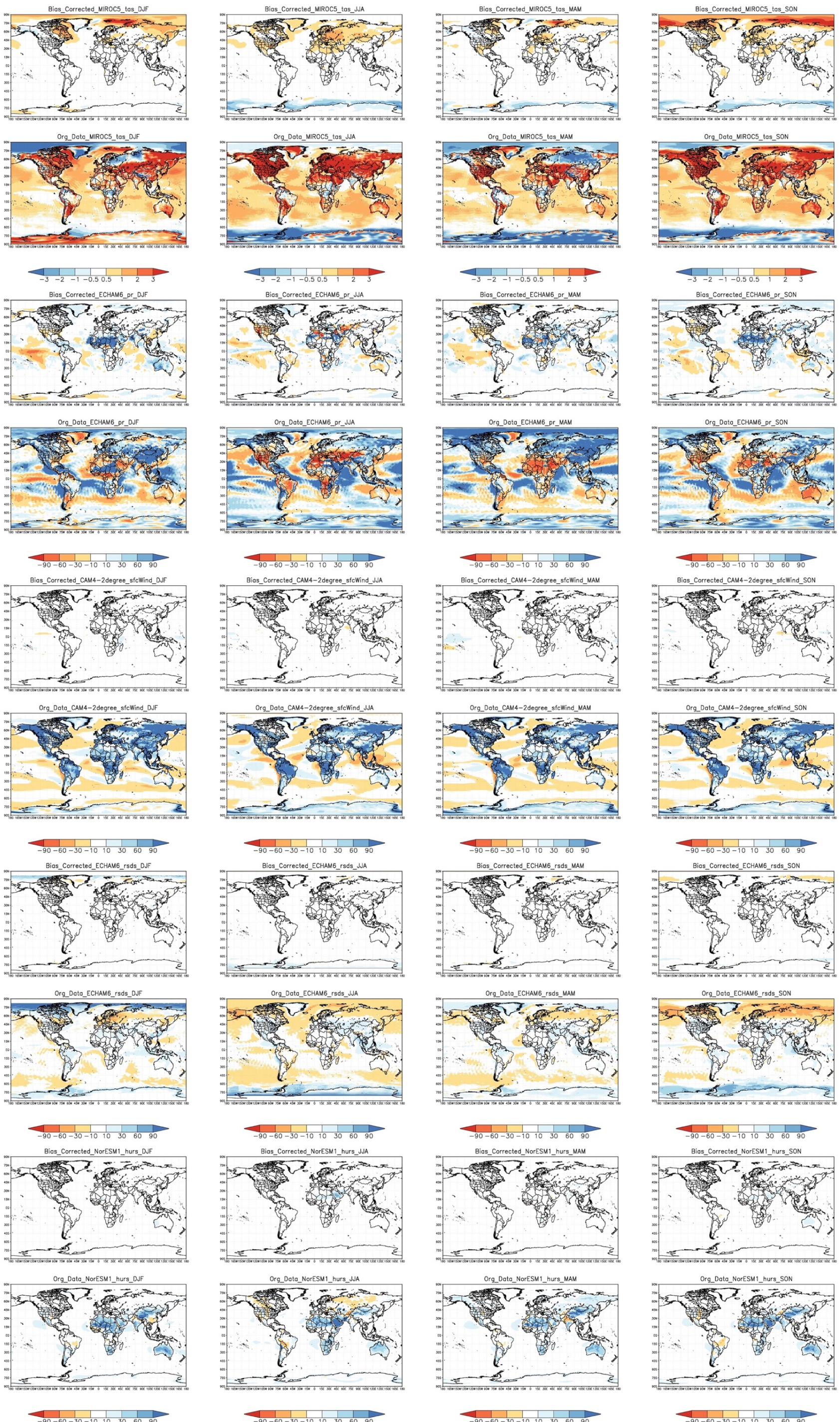

Figure 1: Seasonal absolute (tas) and relative (pr,rsds,hurs,sfcWind) differences against EWEMBI data for different GCMs. For each variable, upper and lower panels show the results of bias corrected and original (non-bias corrected) data respectively.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-107
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 25 June 2018

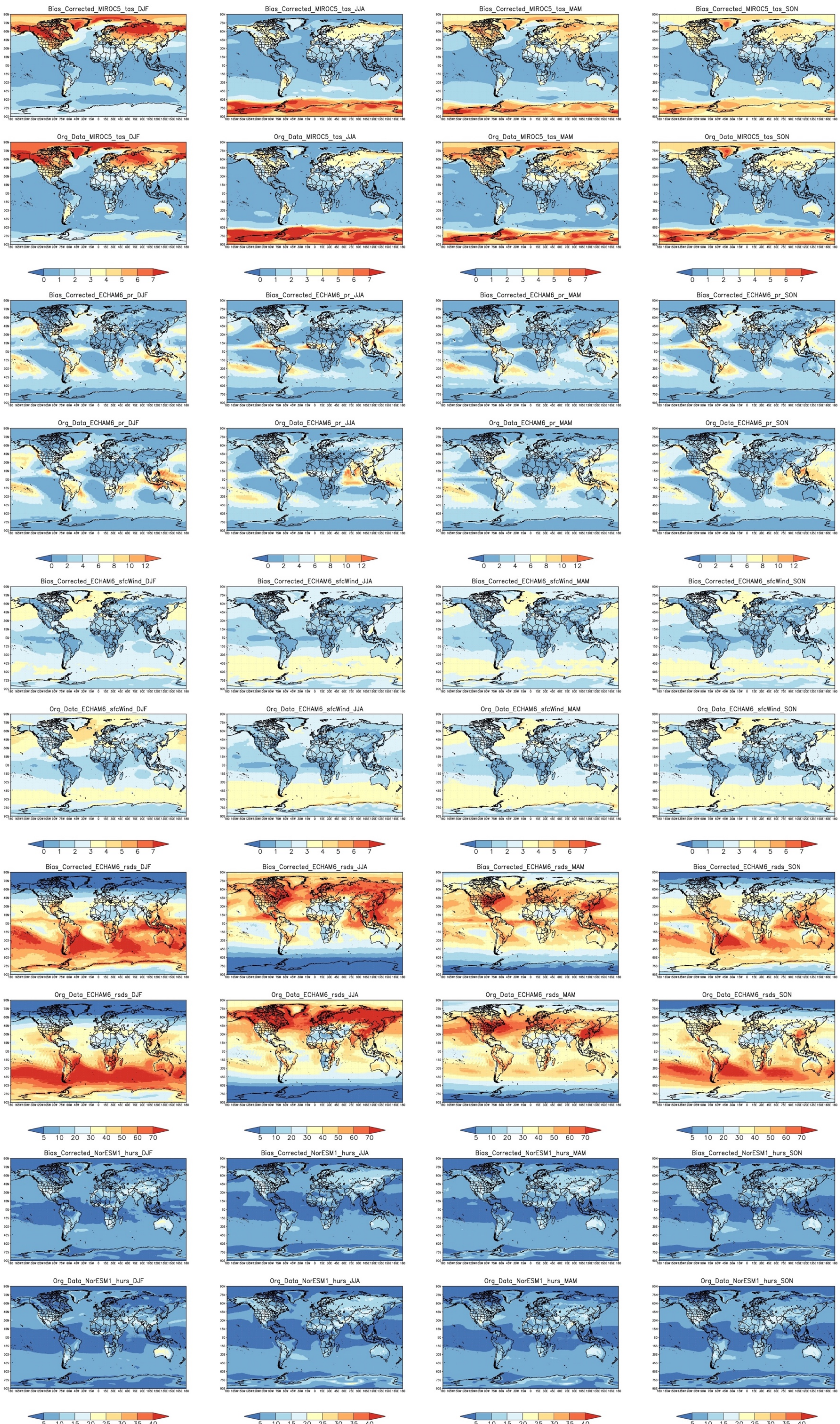

Figure 2: Seasonal ensemble standard deviation across different GCMs for different variables. For each variable, upper and lower panels show the results of bias corrected and original (non-bias corrected) data respectively.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-107
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 25 June 2018

Geoscientific Model Development
Discussions
Open Access

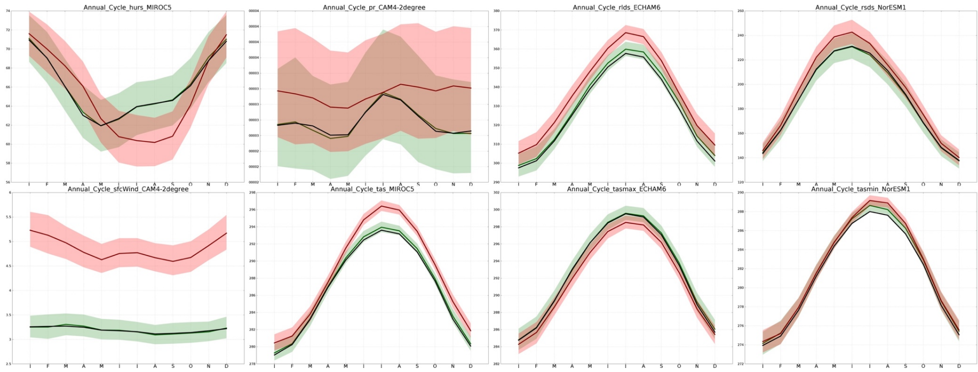

Figure 3: : Annual Cycle for different variables averaged over the land areas across different GCMs. Red, green and black curves represent ensemble median for original (non-bias corrected), bias-corrected and EWEMBI data. The respective coloured band around original and bias corrected curves represent the spread of the 20 Ensemble member for each variable.

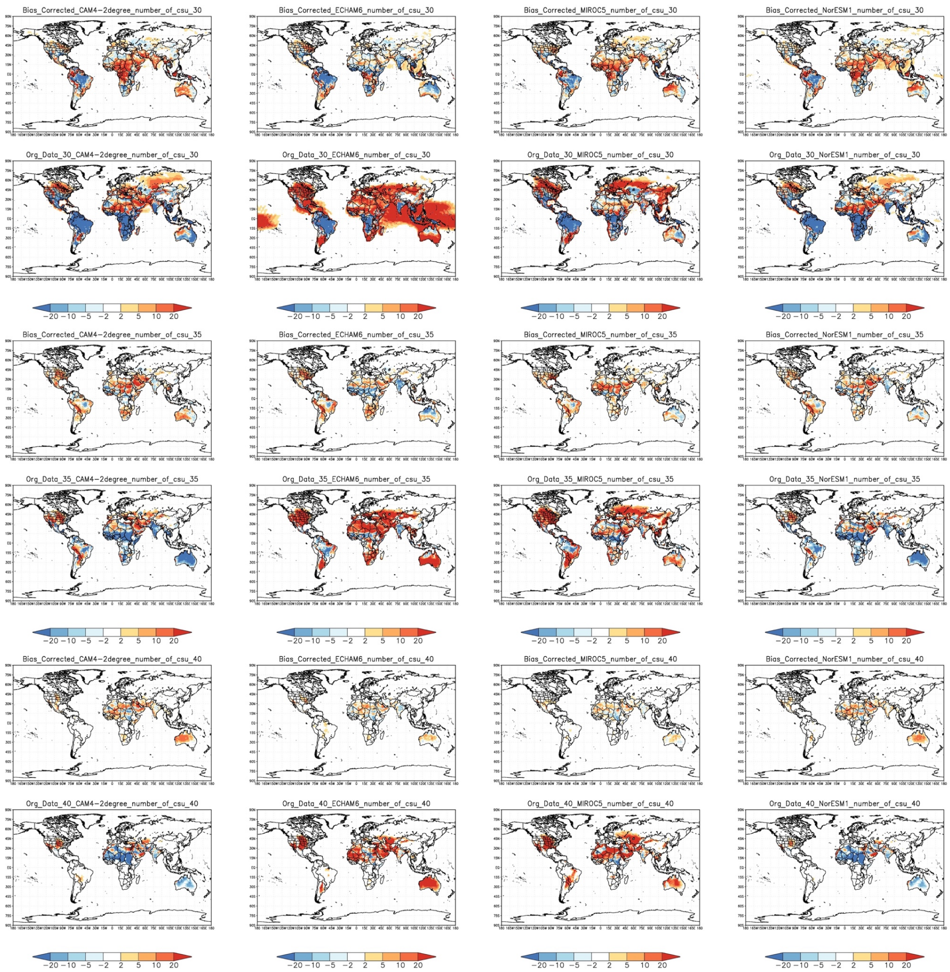

Figure 4: Consecutive Summer Days (CSD) index at the thresholds of 30°C, 35°C and 40°C presented as a difference from EWEMBI data. Presented are the number of instances during which the maximum temperature remains equal or more than the respective thresholds consecutively for 5 days or more. Upper and Lower panels show the results for bias corrected and original (non-bias corrected) data across all the four GCMs.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-107
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 25 June 2018

Geoscientific Model Development
Discussions

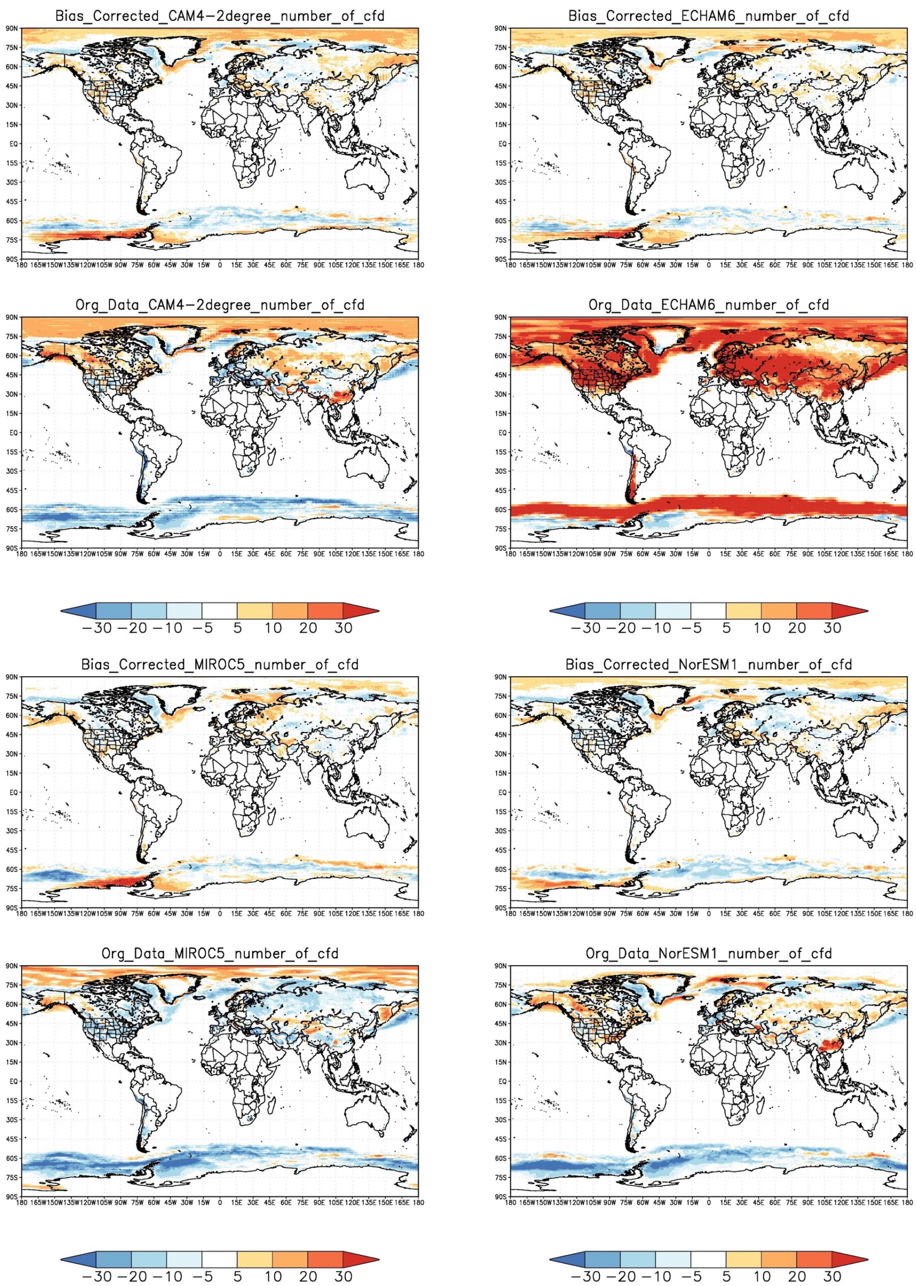

Figure 5: Consecutive Frost Days (CFD) index presented as a difference from EWEMBI data. Presented are the number of instances during which the minimum temperature remains equal or less than 0°C consecutively for 5 days or more. Upper and Lower panels show the results for bias corrected and original (non-bias corrected) data across all the four GCMs.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-107
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 25 June 2018

Geoscientific Model Development
Open Access
Discussions

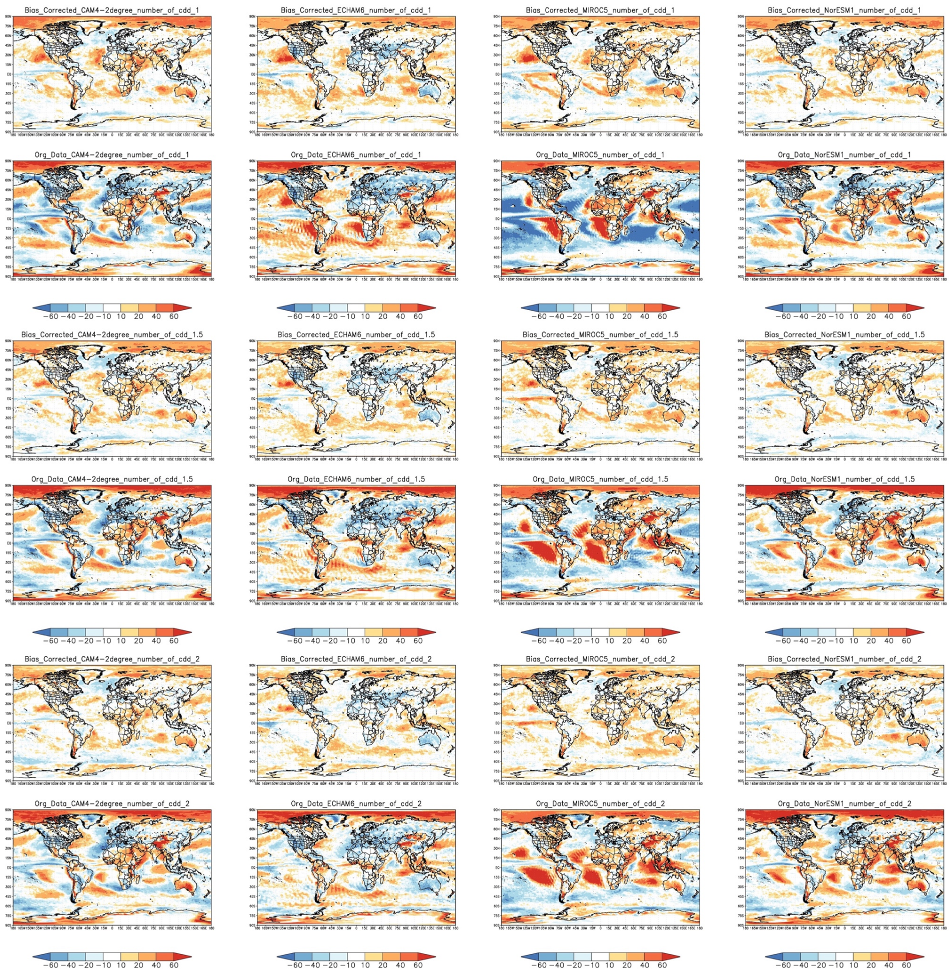

Figure 6: Consecutive Dry Days (CDD) index at the thresholds of 1mm/day, 1.5 mm/day and 2.0 mm/day presented as a difference from EWEMBI data. Presented are the number of instances during which the daily precipitation remains more than the respective thresholds consecutively for 5 days or more. Upper and Lower panels show the results for bias corrected and original (non-bias corrected) data across all the four GCMs.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-107
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 25 June 2018

Geoscientific Model Development
Open Access
Discussions

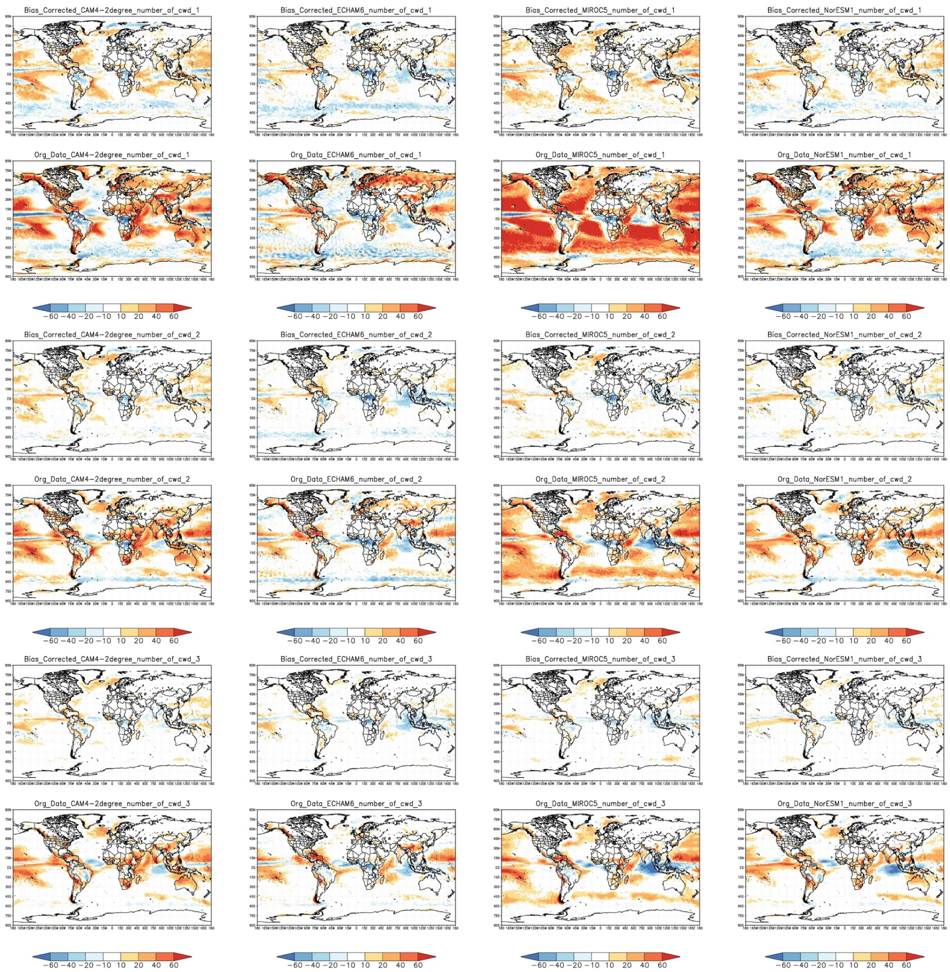

Figure 7: Consecutive Wet Days (CWD) index at the thresholds of 1mm/day, 2 mm/day and 3 mm/day presented as a difference from EWEMBI data. Presented are the number of instances during which the daily precipitation remains more than the respective thresholds consecutively for 5 days or more. Upper and Lower panels show the results for bias corrected and original (non-bias corrected) data across all the four GCMs.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-107
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 25 June 2018

Geoscientific Model Development
Open Access
Discussions

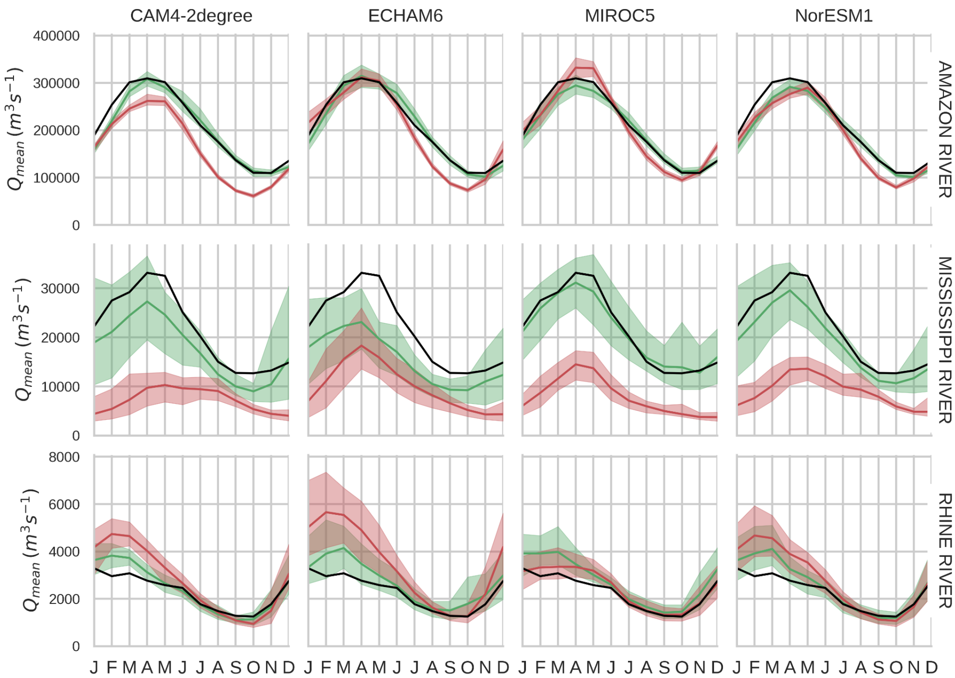

Figure 8: Figure 8: Seasonal discharge at the outlet of Amazon river basin (first row), Mississippi river basin (second row) and Rhine river basin (third row) for 2006 – 2013 for the analysed GCMs. The black line represents the reference simulation forced with EWEMBI, whereas the green and red coloured lines represent the bias corrected and the original (non-bias corrected) ensemble mean. The coloured bands around original and bias corrected curves show the spread of the 20 ensemble members.