

ISIMIP2b bias-correction fact sheet

Stefan Lange May 24, 2018

1	Introduction							
2	Observational dataset							
3	Bias correction methods							
4	Evaluation							
	4.1	Annual global mean, minimum and maximum values	4					
 4.2 Statistical dependencies between variables, in space and in time								
							Reference period statistics	6
		4.4.1 Multi-year monthly mean values	6					
		4.4.2 Multi-year monthly standard deviations	6					
		4.4.3 Multi-year monthly 5th and 95th percentiles	$\overline{7}$					
		4.4.4 Multi-year monthly mean number of wet days	7					
		4.4.5 Interannual variability of monthly mean values	$\overline{7}$					
	4.5	Preservation of trends	8					
		4.5.1 Multi-year monthly mean values	8					
		4.5.2 Multi-year monthly standard deviations	8					
		4.5.3 Multi-year monthly 5th and 95th percentiles	8					
		4.5.4 Multi-vear monthly mean number of wet days	8					
	4.6	Discussion of identified methodological issues	8					
		4.6.1 General issues	8					
		4.6.2 Variable-specific issues	9					
5	Comparison to ISIMIP fast track 9							

If you use bias-corrected ISIMIP2b climate input data then please cite (where applicable) Frieler et al. (2017) and Lange (2018) for the bias correction methods, Lange (2016), Frieler et al. (2017) and Lange (2018) for the observational dataset EWEMBI and Frieler et al. (2017) for the ISIMIP2b simulation protocol.

1 Introduction

In this document we describe and evaluate the bias correction of climate input data that was carried out in order to facilitate ISIMIP's contribution to the IPCC's special report on the impacts of global warming of 1.5°C above pre-industrial levels that is due in 2018. The document is structured as follows. We first describe the observational dataset used for the bias correction (Sec. 2) and outline the bias correction methods (Sec. 3). We then provide present basic evaluation results and discuss identified methodological issues (Sec. 4). Finally, we summarise the innovations that distinguish the ISIMIP2b bias correction from the one carried out in the ISIMIP fast track (Sec. 5).

2 Observational dataset

While WATCH forcing data (WFD; Weedon et al., 2011) were employed for bias correction in the fast track, ISIMIP2b climate input data were corrected using the newly compiled reference dataset EWEMBI (E2OBS, WFDEI and ERAI data Merged and Bias-corrected for ISIMIP; Lange, 2016), which covers the entire globe at 0.5° horizontal and daily temporal resolution from 1979 to 2013. Data sources of EWEMBI are ERA-Interim reanalysis data (ERAI; Dee et al., 2011), WATCH forcing data methodology applied to ERA-Interim reanalysis data (WFDEI; Weedon et al., 2014), eartH2Observe forcing data (E2OBS; Dutra, 2015) and NASA/GEWEX Surface Radiation Budget data (SRB; Stackhouse Jr. et al., 2011). The SRB data were used to bias-correct E2OBS shortwave and longwave radiation using a new method that has been developed particularly for this purpose (Lange, 2018), in order to reduce known deviations of E2OBS radiation statistics from the respective SRB estimates over tropical land (Dutra, 2015). EWEMBI data sources for the variables that were bias-corrected in ISIMIP2b are given in Tab. 1.

The main differences between EWEMBI and WFD are that (i) EWEMBI covers the entire globe and not only land as WFD, (ii) EWEMBI is based on ERA-Interim while WFD is based its predecessor ERA-40 (Uppala et al., 2005) and (iii) EWEMBI radiation data were bias-corrected using the SRB primary algorithm data products whereas Weedon et al. (2011) considered a correction of WFD radiation with the less reliable (Stackhouse Jr., personal communication) SRB quality control data products unnecessary.

Multi-year monthly mean differences between EWEMBI and WFD hurs, pr, prsn, ps, rlds, rsds, sfcWind, tas, tasmax and tasmin are shown in Figs. 1–10. Substantial differences exist for hurs, rlds, rsds, sfcWind, tas, tasmax and tasmin. In particular, average shortwave radiation over central Africa is more than 50% higher in EWEMBI than in WFD (Fig. 11) and average wind speeds in EWEMBI exceed those in WFD by more than 50% in many regions of the world (Fig. 12).

3 Bias correction methods

All variables that were bias-corrected for ISIMIP2b are listed in Tab. 1. The bias correction was performed on the regular 0.5° EWEMBI grid, to which raw CMIP5 GCM data were interpolated using a first-order conservative remapping scheme (Jones, 1999; we abstained from using any higher than first-order scheme because those occasionally produce values outside the range of the interpolated values, such as negative precipitation values).

Temporally, the method works at daily resolution and uses the proleptic Gregorian calendar. Raw CMIP5 output of GCMs using a 365-day calendar was adjusted by filling leap-day gaps with averages of the values of Feb 28 and Mar 31 of the respective year. Similarly, for GCMs using a 360-day calendar, the additionally missing five days were inserted into the raw data after the first 36, 108, 180, 252 and 324 days of each 360-day year and filled with averages of the values of the respective preceding and following days.

The correction was done based on biases identified by comparing simulated to observed data from a historical reference period, which was 1979–2013 here and used to be 1960–1999 in the fast track. We used simulated data from historical CMIP5 runs for 1979–2005 and from RCP8.5 projections for 2006–2013.

Table 1: Variables bias-corrected for ISIMIP2b and their data sources in the EWEMBI dataset (Lange, 2016). Note that E2OBS data are identical to WFDEI and ERAI data over land and ocean, respectively, except for precipitation over the ocean, which was bias-corrected using GPCPv2.1 monthly precipitation totals (Balsamo et al., 2015; Dutra, 2015). WFDEI-GPCC means WFDEI with GPCCv5/v6 monthly precipitation totals used for bias correction (Weedon et al., 2014; note that the WFDEI precipitation products included in E2OBS were those that were bias-corrected with CRU TS3.101/TS3.21 monthly precipitation totals). E2OBS-SRB means E2OBS with SRB daily mean radiation used for bias correction (Lange, 2018). E2OBS-ERAI means E2OBS everywhere except over Greenland and Iceland (cf. Weedon et al., 2010, p. 9), where monthly mean diurnal temperature ranges were restored to those of ERAI using the Sheffield et al. (2006) method. Note that precipitation here means total precipitation, i.e., rainfall plus snowfall.

Variable	Short name	Unit	Source dataset	Source dataset
			over land	over the ocean
Near-Surface Relative Humidity	hurs	%	E2OBS	E2OBS
Near-Surface Specific Humidity	huss	$\rm kg \ kg^{-1}$	E2OBS	E2OBS
Precipitation	pr	$kg m^{-2} s^{-1}$	WFDEI-GPCC	E2OBS
Snowfall Flux	prsn	$kg m^{-1} s^{-1}$	WFDEI-GPCC	E2OBS
Surface Air Pressure	\mathbf{ps}	Pa	E2OBS	E2OBS
Surface Downwelling Longwave	rlds	${ m W}~{ m m}^{-2}$	E2OBS-SRB	E2OBS-SRB
Radiation				
Surface Downwelling Shortwave	rsds	${ m W}~{ m m}^{-2}$	E2OBS-SRB	E2OBS-SRB
Radiation				
Near-Surface Wind Speed	sfcWind	${\rm m~s^{-1}}$	E2OBS	E2OBS
Near-Surface Air Temperature	tas	К	E2OBS	E2OBS
Daily Maximum Near-Surface Air	tasmax	К	E2OBS-ERAI	E2OBS
Temperature				
Daily Minimum Near-Surface Air	tasmin	K	E2OBS-ERAI	E2OBS
Temperature				

The correction was done independently for each variable, grid cell and month to the end of adjusting the respective distributions to those observed in the reference period if applied in the reference period, and to climate change versions of these reference distributions if applied outside the reference period.

The variables listed in Tab. 1 are of various natures and this is reflected in the use of different bias correction methods for different variables. If we abstract away from oversaturation, hurs can only take values in [0%, 100%] while pr, prsn, ps, rlds, rsds and sfcWind are also non-negative but do not possess any well-defined upper bounds. The non-negavity of tas can be safely neglected. Special constraints hold for huss which is a dependent variable once hurs, ps and tas are known, and for tasmin and tasmax, namely tasmin \leq tas \leq tasmax.

The methods used to correct pr, prsn, tas, tasmax and tasmin are identical to those used in the fast track (Hempel et al., 2013), except that we defined dry days using a modified threshold value of 0.1 mm/day here since this value was used to correct WFDEI dry day frequencies (cf. Sec. 2 and Tab. 1; Harris et al., 2013; Weedon et al., 2014). Also, in order to prevent the bias correction from creating unrealistically extreme temperatures, we introduced a maximum value of 3 for the correction factors of tas – tasmin and tasmax – tas [cf. Hempel et al., 2013, Eq. (25)] and limited tas, tasmin and tasmax to the range $[-90^{\circ}C, 60^{\circ}C]$, in line with historical record near-surface temperature observations (WMO; cf. Sec. 4.1).

Known deficiencies of the method used to correct rlds and sfcWind in the fast track were that it produced (i) too low annual global minimum values (cf. Sec. 4.1) and (ii) discontinuous daily climatologies as described by Rust et al. (2015) for the WATCH forcing datasets. Problem (i) was solved by setting $\delta \hat{P}_{\min}^{\text{GCM}} = 0$ in the transfer function used for the correction of daily variability around monthly means (Hempel et al., 2013, Eq. (14)). Problem (ii) was solved by equipping the method with daily (instead of monthly) climatologies using linearly interpolated monthly mean values as in the tas correction method (Hempel et al., 2013, Eqs. (16–20)).

Bias-corrected ps was obtained from CMIP5 output of sea level pressure (psl) in three steps. First, EWEMBI ps was reduced to EWEMBI psl using EWEMBI tas, WFDEI and ERAI surface elevation over land except Antarctica and the rest of the earth's surface, respectively, and

$$psl = ps \, \exp\left(\frac{g \, z}{R \, tas}\right),\tag{1}$$

where z is surface elevation, g is gravity and R is the specific gas constant of dry air. Simulated psl was then corrected using EWEMBI psl and the tas correction method described by Hempel et al. (2013). Using the tas correction method for psl is fine in terms of avoiding negative bias-corrected values – which was the rationale behind utilising the precipitation method for pressure corrections in the fast track – since psl values are generally well above zero, and it is advantageous because it produces continuous daily climatologies (cf. previous paragraph). Finally, the bias-corrected psl was transformed to a bias-corrected ps using (1) with WFDEI and ERAI surface elevation and bias-corrected tas.

Note that in the fast track, ps was corrected directly after it had been retrieved from raw CMIP5 psl and tas using a relationship similar to (1). We decided to change that since the reduction of surface to sea level pressure is done differently in every GCM, so in fact also the retrieval of ps should be done differently for every GCM, but these retrievals are impossible in cases where the relationships used for the reduction from ps to psl are more sophisticated than (1) and require the knowledge of variables that have not been stored in the CMIP5 archive. The new approach circumvents this problem.

A new and approximately trend-preserving bias correction method was developed for rsds because the old method produced unrealistically high rsds values too frequently. The new method fits beta distributions to the observed and simulated daily rsds data and then transforms the simulated data based on these fitted distributions via quantile mapping as described by Lange (2018). Reflecting the physical limits of rsds, the lower bounds of the beta distributions were set to zero and their upper bounds were estimated by rescaled climatologies of downwelling shortwave radiation at the top of the atmosphere. Details of the distribution fitting are given in Lange (2018, method BCsda1). Approximate trend preservation was achieved as follows. Let $F_{\rm ref}^{\rm to}$, $F_{\rm ref}^{\rm from}$ denote the beta distributions fitted to rsds observed during the reference period, simulated during the reference period and simulated during any other period, respectively. Then the target beta distribution used for quantile mapping of simulated rsds during that other period, $F_{\rm other}^{\rm to}$. Specifically, let x, m and v denote the upper bound, the relative mean value ($m = \mu/x$, where μ is the mean value) and the relative variance ($v = \sigma^2/(\mu(x - \mu))$), where μ and σ are mean value and standard deviation, respectively) of a beta distribution. Then $0 \le m \le 1$ and $0 \le v \le 1$ (Wilks, 1995), and we defined the upper bound of $F_{\rm other}^{\rm to}$ by

$$x_{\text{other}}^{\text{to}} = \begin{cases} 0 & \text{if } x_{\text{ref}}^{\text{from}} = 0, \\ x_{\text{ref}}^{\text{to}} x_{\text{other}}^{\text{from}} / x_{\text{ref}}^{\text{from}} & \text{otherwise,} \end{cases}$$
(2)

its relative mean value by

$$m_{\text{other}}^{\text{to}} = \begin{cases} m_{\text{ref}}^{\text{to}} & \text{if } m_{\text{other}}^{\text{from}} = m_{\text{ref}}^{\text{from}}, \\ m_{\text{ref}}^{\text{to}} m_{\text{other}}^{\text{from}} / m_{\text{ref}}^{\text{from}} & \text{if } m_{\text{other}}^{\text{from}} < m_{\text{ref}}^{\text{from}}, \\ 1 - (1 - m_{\text{ref}}^{\text{to}})(1 - m_{\text{other}}^{\text{from}}) / (1 - m_{\text{ref}}^{\text{from}}) & \text{if } m_{\text{other}}^{\text{from}} > m_{\text{ref}}^{\text{from}}, \end{cases}$$
(3)

and its relative variance, v_{other}^{to} , in the same way as the relative mean value, i.e., using Eq. (3) with m replaced by v.

Using beta distributions with fixed lower and upper bounds of 0% and 100%, respectively, the new rsds bias correction method was also applied to hurs, which had not been corrected in the fast track. A bias-corrected huss consistent with bias-corrected hurs, ps and tas was then calculated using the equations of Buck (1981), as described by Weedon et al. (2010).

The most important feature of the fast-track methods is that they preserve differences between multi-year monthly mean values over any two periods in absolute terms for temperature and in relative terms for pr, psl, rlds, sfcWind, tas – tasmin and tasmax – tas. Since

$$\frac{\mathrm{prsn}_{\mathrm{corrected}}}{\mathrm{pr}_{\mathrm{corrected}}} = \frac{\mathrm{prsn}_{\mathrm{raw}}}{\mathrm{pr}_{\mathrm{raw}}} \tag{4}$$

is used at the daily level to correct prsn, also relative trends in multi-year monthly mean prsn are approximately preserved.

An important consequence of the trend preservation is that corrected time series from different GCMs diverge as they depart from the reference period going both forward and backward in time. For global mean values this is illustrated in Fig. 13.

4 Evaluation

The ISIMIP2b bias correction methods were applied to CMIP5 output of IPSL-CM5A-LR, GFDL-ESM2M, HadGEM2-ES and MIROC5. In the fast track, ISIMIP provided bias-corrected CMIP5 output of IPSL-CM5A-LR, GFDL-ESM2M, HadGEM2-ES, MIROC-ESM-CHEM and NorESM1-M. The GCM selection for ISIMIP2b was mainly motivated by data availability. Nevertheless, compared to the fast track, the new set of GCMs better represents the CMIP5 GCM ensemble in terms of both horizontal model resolution and equilibrium climate sensitivity (Fig. 29). In this section we provide a basic evaluation of the bias-corrected data from each of the four models.

We begin with comparisons of annual global mean, minimum and maximum values before and after the correction (Sec. 4.1). We then briefly discuss how our bias correction methods influence statistical dependencies between variables, in space and in time (Sec. 4.2), and how large correction factors can inflate extreme precipitation events (Sec. 4.3). We then look at reference-period biases before and after the correction for a range of statistics (Sec. 4.4) and check how well trends in these statistics are preserved by the bias correction (Sec. 4.5). We conclude with a discussion of identified methodological issues (Sec. 4.6).

Note that figures from this section that have been produced in the same way for every GCM are distributed to one supplementary document per GCM using identical figure numbers for the analogous figures in the different supplementary documents.

Note also that the NetCDF files containing the bias-corrected data have been checked for proper time, latitude and longitude axes. Moreover, we made sure that they do not contain any infs, nans, missing values, hurs values outside [0%, 100%] or negative huss, pr, prsn, ps, psl, rlds, rsds or sfcWind values.

4.1 Annual global mean, minimum and maximum values

Time series of annual global mean values of all variables before and after the correction are shown in Fig. 13 for all GCMs and CMIP5 experiments. With the exception of prsn, the bias correction makes the time series from different GCMs coincide during the reference period. The preservation of snowfall-to-precipitation ratios at the daily level [Eq. (4)] apparently prevents the correction of long-term mean snowfall. Another observation is that for some variables, most notably for hurs, the EWEMBI data exhibit considerable trends over the reference period which are not present in the uncorrected GCM time series. By design, this discrepancy is preserved by the bias correction.

Similar time series of annual global minimum and maximum values are shown in Figs. 14 and 15. For all GCMs, minima of corrected hurs and sfcWind are persistently too close to zero. For GFDL-ESM2M, minima of corrected hurs and (consequently) huss are always zero due to negative hurs values in the raw GFDL-ESM2M data that were mapped to zero by the bias correction. For a discussion of the particularly low MIROC5 sfcWind minima see Sec. 4.4.1. For IPSL-CM5A-LR and MIROC5, annual global minima of corrected daily mean and minimum temperatures sometimes touch the lower limit of -90°C. Otherwise the minima are well corrected.

Annual global maxima of corrected huss are too high for all GCMs. The accordance of annual global maxima of corrected hurs, ps and tas with corresponding EWEMBI values suggests that this overestimation is due to misrepresented statistical dependencies between these variables. The bias correction caps precipitation at 400 mm/day and temperature at 60°C. Figure 15 shows that these

upper limits are reached rather frequently. Maps of corresponding limit exceedance frequencies during the reference period (Figs. S2 and S3) show that most of the limit precipitation events occur in the intertropical convergence zone and in the Indian summer monsoon region (further occurrences are scattered across tropical and subtropical latitudes), and that the limit temperature events occur over Australia, Morocco, Pakistan and Turkmenistan, depending on the GCM. Otherwise the maxima are well corrected.

4.2 Statistical dependencies between variables, in space and in time

As alluded to above, the ISIMIP2b bias correction methods adjust distributions independently for each variable, grid cell and month. Statistical dependencies between variables, in space and in time are therefore not corrected.

As an example for inter-variable dependencies, Pearson correlations between daily mean temperature tas and diurnal temperature range dtr = tasmax – tasmin from EWEMBI, raw and corrected IPSL-CM5A-LR data are shown in Figs. 16–18. The statistical dependencies between tas and dtr are similarly poorly represented before and after bias correction.

As an example for spatial dependencies, the second empirical orthogonal functions of sea level pressure from EWEMBI, raw and corrected IPSL-CM5A-LR data are shown in Figs. 19–21. Again, the bias correction hardly alters the misrepresented spatial covariances.

As an example for temporal dependencies, autocorrelations of binarised precipitation time series (pr \mapsto 1 if pr \geq 0.1 mm/day else 0) from EWEMBI, raw and corrected IPSL-CM5A-LR data are shown in Figs. 19–21. In this case, the bias correction yields minor improvements since it corrects wet day frequencies (cf. Sec. 4.4.4) but major deviations in the temporal structure of precipitation time series persist.

4.3 Large correction factors and extreme precipitation events

Hagen Koch from the regional water sector brought to our attention the following issue that he came across when he prepared hydrological impact assessments for the São Francisco river basin in the semiarid Northeast of Brazil. Fed with bias-corrected fast-track climate input data based on NorESM1-M future climate projections under RCP8.5 his hydrological model SWIM simulated discharge with much too high interannual variability and produced one particularly extreme flood event in March 2031 that we want to focus on here.

EWEMBI-to-NorESM1-M ratios of 1979–2013 March mean precipitation over Northeast Brazil are shown in Fig. 25. Values greater than 4 occur over large parts of the region. In order to reduce this underestimation of monthly mean rainfall the bias correction method developed by Hempel et al. (2013) rescales simulated monthly precipitation values with these ratios as long as the ratios do not exceed a cap value (or fall below the cap value's inverse in the case of overestimation), which was set to 10 in the fast track, and as long as the resulting daily precipitation values do not exceed a maximum value of 400 mm/day.

It now happened that under RCP8.5 NorESM1-M simulated some days with rather extreme precipitation in March 2031. Time series of raw and bias-corrected regional mean and maximum precipitation falling in this month are shown in Fig. 26. Rainfall amounts on 7 March 2031 correspond to a twosigma event relative to the historical EWEMBI distribution and to a four- to five-sigma event relative to the historical NorESM1-M distribution (Fig. 27).

The Hempel et al. (2013) bias correction method tries to adjust these sigma levels at the grid scale, similar to how a simple quantile mapping would do, i.e., high-sigma events relative to the historical NorESM1-M distribution are mapped to high-sigma events relative to the historical EWEMBI distribution. The present example shows that this can turn moderately extreme into very extreme precipitation events if correction factors are large.

We investigated if a reduction of the cap value for the correction factors can alleviate the problem. Results for cap values ranging from 2 to 6 are shown in Figs. 26 and 27. Apparently, a considerable improvement requires a reduction of the cap value to 4 or lower. But this comes at the cost of increased biases in precipitation climatologies as shown for IPSL-CM5A-LR in Fig. 28. It was therefore decided not to change the cap value for ISIMIP2b. Future method development shall yield a more fundamental solution of the problem.

In the meantime, beware of extreme precipitation events in regions and months where and when the bias correction method worked with large correction factors. Monthly maps of precipitation correction factors for the different GCMs are shown in Fig. S1.

4.4 Reference period statistics

4.4.1 Multi-year monthly mean values

Maps of biases of multi-year monthly mean values before and after bias correction are shown in Figs. S4–S25. The bias correction does a good job at adjusting this most basic statistic with the following exceptions. For IPSL-CM5A-LR, some overestimations of relative humidity are introduced over the tropical oceans. Some underestimations of precipitation remain over tropical South America for IPSL-CM5A-LR and GFDL-ESM2M and over India for IPSL-CM5A-LR and HadGEM2-ES. These underestimations coincide with capped correction factors (cf. Fig. S1). Snowfall flux biases are well reduced over land, but over the mid-latitude oceans and over the Southern Ocean the bias correction is not able to reduce biases for any of the four GCMs.

A special case are MIROC5 near-surface wind speeds over tropical rainforests, which are underestimated before and after the correction, most visibly over South America. The reason is that sfcWind was calculated inside the rainforest canopy in MIROC5 (Tatsuo Suzuki, personal communication) and was therefore one to two orders of magnitude too weak in the raw data. Since we capped sfcWind correction factors at 10 (just like pr correction factors) these strong underestimations were not fully corrected. We abstained from increasing the cap value because that could have resulted in unrealistically strong winds outside the reference period (cf. Sec. 4.5.1) as the spatial distribution of tropical rainforests in MIROC5 changes during the RCP runs, in line with projected land-use changes (Tatsuo Suzuki, personal communication).

4.4.2 Multi-year monthly standard deviations

Maps of biases of multi-year monthly standard deviations of daily data before and after bias correction are shown in Figs. S26–S47. For hurs, underestimations of standard deviations are well corrected while some overestimations persist or are created in particular over the Arctic Ocean. Generally, biases of huss standard deviations are well reduced at high latitudes whereas at tropical latitudes and in particular over tropical oceans, huss standard deviations are mostly overestimated after the correction.

For pr, standard deviation biases are generally well reduced by the bias correction. Strong biases remain over the subtropical dry regions. For prsn, standard deviation biases are mostly well reduced over land. The correction works less well over mid-latitude oceans and over the Southern Ocean for GFDL-ESM2M and HadGEM2-ES as well as over Europe and over the eastern US for IPSL-CM5A-LR, GFDL-ESM2M and MIROC5.

For psl, standard deviation biases are well corrected outside the tropics except for IPSL-CM5A-LR over Antarctica. Considerable biases remain over tropical regions where the interannual variability of monthly means is large relative to the intramonthly variability of daily means (cf. Secs. 4.4.5 and 4.6). This problem is most severe for GFDL-ESM2M.

For rlds and rsds, standard deviation biases are generally well corrected. Remaining biases in rlds standard deviations are found over high latitudes in winter and in the intertropical convergence zone. For rsds, some slight underestimations remain or are created over tropical and subtropical land, in particular for GFDL-ESM2M.

For sfcWind, standard deviation biases are well corrected everywhere for HadGEM2-ES while some overestimations persist over tropical oceans for IPSL-CM5A-LR, GFDL-ESM2M and MIROC5 and underestimations remain over tropical rainforests for MIROC5 (cf. discussion in Sec. 4.4.1).

For tas, most underestimations of standard deviations are well corrected. Substantial overestimations persist over tropical oceans for IPSL-CM5A-LR, GFDL-ESM2M and MIROC5, over tropical South America for GFDL-ESM2M and MIROC5, over subtropical oceans for IPSL-CM5A-LR and HadGEM2-ES, over the Arctic Ocean for IPSL-CM5A-LR and over the Southern Ocean for HadGEM2-ES. These biases persist because – as for psl – the current tas correction method only adjusts the intramonthly variability of daily means and leaves the interannual variabilities of monthly means unchanged (cf. Secs. 4.4.5 and 4.6).

For tasmax, standard deviation biases are generally reduced. Remaining biases include overestimations over tropical oceans for IPSL-CM5A-LR, GFDL-ESM2M and MIROC5, over subtropical oceans for IPSL-CM5A-LR and HadGEM2-ES, over the Arctic Ocean for IPSL-CM5A-LR and over the Southern Ocean for HadGEM2-ES as well es underestimations over tropical land for IPSL-CM5A-LR, GFDL-ESM2M and HadGEM2-ES. For tasmin, the bias correction changes standard deviation biases in a similar manner as for tasmax, but with smaller remaining biases. In most cases, the bulk of the remaining variability biases of tasmax and tasmin are inherited from tas.

4.4.3 Multi-year monthly 5th and 95th percentiles

Maps of biases of multi-year monthly 5th percentiles of daily data before and after bias correction are shown in Figs. S48–S69. The 5th percentiles are generally well corrected with the following exceptions. For GFDL-ESM2M and IPSL-CM5A-LR, the bias correction introduced some underestimations of 5th percentiles of huss over tropical oceans, and for IPSL-CM5A-LR, the bias correction was not able to fully correct underestimations of 5th percentiles of rlds over the Arctic Ocean.

Maps of biases of multi-year monthly 95th percentiles of daily data before and after bias correction are shown in Figs. S70–S91. The 95th percentiles are generally well corrected with the following exceptions. For hurs, some minor biases remain at high latitudes for IPSL-CM5A-LR. For all GCMs, some overestimations of 95th percentiles of huss are introduced over the tropics and subtropics. Over the tropics, the 95th percentiles of pr were not fully corrected or changed sign, and for IPSL-CM5A-LR and GFDL-ESM2M, considerable underestimations remain over tropical South America in some months. Generally, biases in 95th percentiles of prsn only slightly improve; particularly strong biases persist over the oceans. For psl, biases in 95th percentiles were not fully or overcorrected at high latitudes. For rlds, some biases in 95th percentiles remain over high latitudes and some overestimations were introduced over the Arctic Ocean. For IPSL-CM5A-LR and MIROC5 this also applies to tas and tasmax. For GFDL-ESM2M, the bias correction introduced some underestimations of 95th percentiles of rsds over land. For MIROC5, underestimations of 95th percentiles of sfcWind remain over tropical rainforests.

4.4.4 Multi-year monthly mean number of wet days

The precipitation bias correction method (Hempel et al., 2013) adjusts the distribution of precipitation amounts on wet days as well as the multi-year monthly mean number of wet days. Maps of biases in these wet day frequencies before and after bias correction are shown in Figs. S92 and S93, respectively. The widespread overestimations of wet day frequencies were generally well corrected. For IPSL-CM5A-LR and GFDL-ESM2M, the more challenging correction of too low wet day frequencies was not successful over tropical South America in its dry season.

4.4.5 Interannual variability of monthly mean values

The interannual variability of monthly mean values is not explicitly corrected by any of the ISIMIP2b bias correction methods. Nevertheless, the bias correction alters these variabilities and in this section we want to check if it does so for the better or worse. Maps of biases of standard deviations of monthly mean values before and after bias correction are shown in Figs. S94–S115. The bias correction reduces the overall magnitude of the biases for hurs, rsds, pr and prsn while no such improvements are found for huss, psl, rlds, tas, tasmax and tasmin. For sfcWind, the bias correction increases the overall magnitude of the biases for IPSL-CM5A-LR, decreases it for MIROC5 and does not changed it much for GFDL-ESM2M and HadGEM2-ES. In some cases, the bias correction changes the sign of the variability bias.

4.5 Preservation of trends

4.5.1 Multi-year monthly mean values

All ISIMIP2b bias correction methods were designed to preserve trends in multi-year monthly mean values – absolute trends for psl and tas, relative trends for pr, psl, rlds, sfcWind and the diurnal temperature range, and trends within boundaries for hurs and rsds. In this section we want to check if the methods are successful at trend preservation (Figs. S116–S137).

The method used to correct hurs and rsds preserves trend directions and attempts to also preserve trend magnitudes as far as that is possible given the upper and lower bounds of these variables. We find high spatial correlations between trends before and after bias correction even though changes of trend magnitude can be substantial for both variables.

For some GCMs, trend changes are also considerable for huss and prsn. Otherwise, magnitudes and spacial patterns of trends in multi-year monthly mean values are well preserved.

4.5.2 Multi-year monthly standard deviations

Maps of relative trends in multi-year monthly standard deviations before and after bias correction are shown in Figs. S138–S159. The ISIMIP2b bias correction methods were designed to also preserve these trends in some way or another. The results show that the direction of standard deviation trends was generally well preserved. Larger changes in spatial trend patterns are found for hurs and huss (in particular for IPSL-CM5A-LR in boreal summer), prsn and rsds.

4.5.3 Multi-year monthly 5th and 95th percentiles

Since the ISIMIP2b bias correction methods were designed to preserve trends in multi-year monthly mean values and standard deviations, they can also be expected to preserve trends in multi-year monthly 5th percentiles to some extent. Maps of such trends before and after bias correction are shown in Figs. S160–S181. Generally, spatial trend patterns are well preserved. Major changes are only found for hurs, huss, rsds and particularly pr, whose spatial trend patterns before and after bias correction are not at all correlated for IPSL-CM5A-LR.

Bias correction-induced changes of trends in multi-year monthly 95th percentiles (Figs. S182–S203) are similar to the 5th percentile ones.

4.5.4 Multi-year monthly mean number of wet days

Spatial correlations between trends in multi-year monthly mean numbers of wet days before and after bias correction (Figs. S204 and S205) are typically greater than 0.7 but can go down to 0.64 in individual months. The larger changes in trend magnitudes associated with these low pattern correlations predominently occur over the ocean.

4.6 Discussion of identified methodological issues

4.6.1 General issues

The methods used here to correct pr, psl, rlds, sfcWind and tas adjust the variability of daily values about the respective monthly mean but leave the interannual variability of the monthly means unadjusted. Consequently, the variability of daily values about multi-year monthly means is not well corrected where the interannual variability of monthly means is large relative to the intramonthly variability of daily means. This is particularly the case at tropical latitudes. We need correction methods that also adjust the interannual variability of monthly means to overcome this problem.

We found too low minimum and/or too high maximum values in bias-corrected hurs and (consequently) huss, rlds, sfcWind, tas, tasmax and tasmin values. The use of non-parametric instead of parametric bias correction methods (such as empirical quantile mapping) might help to mitigate this issues. The ISIMIP2b bias correction methods do not adjust statistical dependencies between variables. We think that a considerable fraction of the remaining biases in huss maximum values and standard deviations are due to hurs, ps and tas being corrected univariately. Future ISIMIP phases should apply multivariate bias correction methods.

4.6.2 Variable-specific issues

In some cases, the hurs correction method considerably changed trends in multi-year monthly standard deviations. We found that these changes are reduced if empirical instead of parametric cumulative distribution functions (CDFs) are employed for the description of simulated hurs distributions. We nevertheless used the parametric CDFs to reduce the risk of overfitting.

We found that the method used here to correct pr (i) is not always successful at enhancing underestimated wet day frequencies, (ii) sometimes completely fails to preserve trends in extreme percentiles and (iii) potentially amplifies out-of-sample extreme events to unrealistic levels where large correction factors are needed to adjust precipitation climatologies. These problems need to be tackled by future correction method development.

The method used here to bias-correct prsn was shown to perform poorly where prsn/pr ratios are neither close to 0 nor close to 1. In the future, the distribution of daily prsn/pr ratios should be corrected instead of retained.

Finally, we found considerable remaining variability biases and too extreme global annual maximum and minimum values in bias-corrected tasmax and tasmin data, respectively. In the future, the simple rescaling method used here to correct climatological mean values of tas – tasmin and tasmax – tas should be replaced by a method that corrects the distributions of these quantities.

5 Comparison to ISIMIP fast track

While in the fast track, WATCH forcing data (WFD) were employed for bias correction, the ISIMIP2b climate input data were corrected using the newly compiled observational dataset EWEMBI, which is based on ERA-Interim (WFD is based on ERA-40), covers land and ocean (WFD covers land only) and features bias-corrected radiation data based on SRB observations (WFD does not). Multi-year monthly mean differences between WFD and EWEMBI are substantial for humidity, radiation, wind speed and temperature. Very similar differences will be found between bias-corrected climate input data from the fast track and ISIMIP2b.

The reference period used to identify the biases to be corrected was 1979–2013 here and used to be 1960–1999 in the fast track. Therefore, bias-corrected climate time series from different GCMs diverge from a focal point in time that is about 15 years later in ISIMIP2b than in the fast track. Please note that this leads to a possibly ostensible reduction of uncertainties in future climate projections (cf. Hawkins and Sutton, 2016, Fig. SB1).

We here used a first-order conservative remapping scheme to interpolate raw CMIP5 output from the native grid of each GCM to the 0.5° EWEMBI grid while this interpolation was done bilinearly in the fast track. The new approach is advantageous as it conserves spatial averages.

For all variables except snowfall, the ISIMIP2b bias correction methods differ from those used in the fast track. Some threshold and cap values were modified and introduced in the correction methods for precipitation and the three temperature variables, respectively, to the end of keeping the bias-corrected values within realistic limits. In order to circumvent the problem of retrieving the original CMIP5 surface pressure data, pressure was corrected at sea level here instead of at the surface as in the fast track. We replaced monthly with daily climatologies in the correction methods for pressure, longwave radiation and wind speed in order to avoid discontinuities in bias-corrected daily climatologies of these variables. Wholly new methods were developed and applied for the correction of relative humidity and shortwave radiation – the only two variables that have both lower and upper bounds. In the fast track, relative humidity was not corrected at all and shortwave radiation was corrected with a method that too frequently produced unrealistically high values.

The ISIMIP2b bias correction methods were applied to CMIP5 output of IPSL-CM5A-LR, GFDL-ESM2M, MIROC5 and HadGEM2-ES. On average, the new GCMs feature a lower equilibrium climate

sensitivity and an atmospheric model with higher horizontal resolution than those employed in the fast track.

The bias-corrected ISIMIP2b climate input data cover pre-industrial, historical and future climate conditions until 2300 over land and ocean. The new data thus provide greater temporal and spatial coverage than the corresponding fast track data.

References

- Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Brun, E., Cloke, H., Dee, D., Dutra, E., Muñoz Sabater, J., Pappenberger, F., de Rosnay, P., Stockdale, T., and Vitart, F.: ERA-Interim/Land: a global land surface reanalysis data set, Hydrology and Earth System Sciences, 19, 389–407, doi: 10.5194/hess-19-389-2015, 2015.
- Buck, A. L.: New Equations for Computing Vapor Pressure and Enhancement Factor, Journal of Applied Meteorology, 20, 1527–1532, doi:10.1175/1520-0450(1981)020(1527:NEFCVP)2.0.CO;2, 1981.
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, doi:10.1002/qj.828, 2011.
- Dutra, E.: Report on the current state-of-the-art Water Resources Reanalysis, Earth2observe deliverable no. d.5.1, URL http://earth2observe.eu/files/Public%20Deliverables, 2015.
- 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), Geoscientific Model Development, 10, 4321–4345, doi:10.5194/gmd-10-4321-2017, 2017.
- 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, International Journal of Climatology, doi:10.1002/joc.3711, 2013.
- Hawkins, E. and Sutton, R.: Connecting Climate Model Projections of Global Temperature Change with the Real World, Bulletin of the American Meteorological Society, 97, 963–980, doi:10.1175/ BAMS-D-14-00154.1, 2016.
- Hempel, S., Frieler, K., Warszawski, L., Schewe, J., and Piontek, F.: A trend-preserving bias correction – the ISI-MIP approach, Earth System Dynamics, 4, 219–236, doi:10.5194/esd-4-219-2013, 2013.
- Jones, P. W.: First- and Second-Order Conservative Remapping Schemes for Grids in Spherical Coordinates, Monthly Weather Review, 127, 2204–2210, doi:10.1175/1520-0493(1999)127(2204: FASOCR)2.0.CO;2, 1999.
- Lange, S.: EartH2Observe, WFDEI and ERA-Interim data Merged and Bias-corrected for ISIMIP (EWEMBI), doi:10.5880/pik.2016.004, 2016.
- Lange, S.: Bias correction of surface downwelling longwave and shortwave radiation for the EWEMBI dataset, Earth System Dynamics, 9, 627–645, doi:10.5194/esd-9-627-2018, 2018.
- Rust, H. W., Kruschke, T., Dobler, A., Fischer, M., and Ulbrich, U.: Discontinuous Daily Temperatures in the WATCH Forcing Datasets, Journal of Hydrometeorology, 16, 465–472, doi: 10.1175/JHM-D-14-0123.1, 2015.

- Sheffield, J., Goteti, G., and Wood, E. F.: Development of a 50-Year High-Resolution Global Dataset of Meteorological Forcings for Land Surface Modeling, Journal of Climate, 19, 3088–3111, doi: 10.1175/JCLI3790.1, 2006.
- Sherwood, S. C., Bony, S., and Dufresne, J.-L.: Spread in model climate sensitivity traced to atmospheric convective mixing, Nature, 505, 37–42, doi:10.1038/nature12829, 2014.
- Stackhouse Jr., P. W., Gupta, S. K., Cox, S. J., Mikovitz, C., Zhang, T., and Hinkelman, L. M.: The NASA/GEWEX surface radiation budget release 3.0: 24.5-year dataset, Gewex news, 21(1):10-12, URL http://www.gewex.org/resources/gewex-news/, 2011.
- Uppala, S. M., Kållberg, P. W., Simmons, A. J., Andrae, U., Bechtold, V. D. C., Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C. M., Berg, L. V. D., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B. J., Isaksen, L., Janssen, P. A. E. M., Jenne, R., Mcnally, A. P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N. A., Saunders, R. W., Simon, P., Sterl, A., Trenberth, K. E., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.: The ERA-40 re-analysis, Quarterly Journal of the Royal Meteorological Society, 131, 2961–3012, doi:10.1256/qj.04.176, 2005.
- Weedon, G. P., Gomes, S., Viterbo, P., Österle, H., Adam, J. C., Bellouin, N., Boucher, O., and Best, M.: The WATCH forcing data 1958–2001: A meteorological forcing dataset for land surface and hydrological models, Technical report no. 22, URL http://www.eu-watch.org/publications/ technical-reports, 2010.
- Weedon, G. P., Gomes, S., Viterbo, P., Shuttleworth, W. J., Blyth, E., Österle, H., Adam, J. C., Bellouin, N., Boucher, O., and Best, M.: Creation of the WATCH Forcing Data and Its Use to Assess Global and Regional Reference Crop Evaporation over Land during the Twentieth Century, Journal of Hydrometeorology, 12, 823–848, doi:10.1175/2011JHM1369.1, 2011.
- Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., and Viterbo, P.: The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data, Water Resources Research, 50, 7505–7514, doi:10.1002/2014WR015638, 2014.
- Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, Academic Press, San Diego, CA, 1995.

WMO: Global Weather & Climate Extremes Archive, URL https://wmo.asu.edu/#global.



Figure 1: 1979–1999 monthly mean differences between EWEMBI and WFD hurs (%).



Figure 2: 1979–1999 monthly mean differences between EWEMBI and WFD pr (mm/day).



Figure 3: 1979–1999 monthly mean differences between EWEMBI and WFD prsn (mm/day).



Figure 4: 1979–1999 monthly mean differences between EWEMBI and WFD ps (hPa).



Figure 5: 1979–1999 monthly mean differences between EWEMBI and WFD rlds (W/m^2) .



Figure 6: 1979–1999 monthly mean differences between EWEMBI and WFD rsds (W/m^2) .



Figure 7: 1979–1999 monthly mean differences between EWEMBI and WFD sfcWind (m/s).



Figure 8: 1979–1999 monthly mean differences between EWEMBI and WFD tas (°C).



Figure 9: 1979–1999 monthly mean differences between EWEMBI and WFD tasmax (°C).



Figure 10: 1979–1999 monthly mean differences between EWEMBI and WFD tasmin (°C).



Figure 11: 1979–1999 monthly mean relative differences between EWEMBI and WFD rsds (%).



Figure 12: 1979–1999 monthly mean relative differences between EWEMBI and WFD sfcWind (%).



Figure 13: Uncorrected (left) and corrected (right) time series of annual global mean hurs, huss, pr, prsn, ps, psl, rlds, rsds, sfcWind, tas, tasmax and tasmin (from top to bottom) for IPSL-CM5A-LR (red), GFDL-ESM2M (dark blue), MIROC5 (light blue), HadGEM2-ES (orange) and EWEMBI (black). The grey vertical lines indicate the transitions from pre-industrial to historical to RCP to extended RCP periods at the end of 1860, 2005 and 2099, respectively. Pre-industrial, historical and RCP2.6 time series are shown with 100% opacity, RCP6.0 with 40% opacity and the post-1860 continuation of pre-industrial control with 16% opacity.



Figure 14: Same as Fig. 13 but for annual global minimum values.



Figure 15: Same as Fig. 13 but for annual global maximum values.



Figure 16: Multi-year (1979–2013) monthly Pearson correlations between daily tas and tasmax – tasmin from EWEMBI data.



Figure 17: Multi-year (1979–2013) monthly Pearson correlations between daily tas and tasmax – tasmin from raw IPSL-CM5A-LR data.



Figure 18: Multi-year (1979–2013) monthly Pearson correlations between daily tas and tasmax – tasmin from corrected IPSL-CM5A-LR data.



Figure 19: Multi-year (1979–2013) monthly 2nd empirical orthogonal functions of daily psl from EWEMBI data.



Figure 20: Multi-year (1979–2013) monthly 2nd empirical orthogonal functions of daily psl from raw IPSL-CM5A-LR data.



Figure 21: Multi-year (1979–2013) monthly 2nd empirical orthogonal functions of daily psl from corrected IPSL-CM5A-LR data.



Figure 22: Multi-year (1979–2013) monthly autocorrelations of binary wet-dry sequences (pr \mapsto 1 if pr $\geq 0.1 \text{ mm/day}$ else 0) from EWEMBI data.



Figure 23: Multi-year (1979–2013) monthly autocorrelations of binary wet-dry sequences (pr \mapsto 1 if pr $\geq 0.1 \text{ mm/day}$ else 0) from raw IPSL-CM5A-LR data.



Figure 24: Multi-year (1979–2013) monthly autocorrelations of binary wet-dry sequences (pr \mapsto 1 if pr $\geq 0.1 \text{ mm/day}$ else 0) from corrected IPSL-CM5A-LR data.



Figure 25: Study region (latitude-longitude box) over Northeast Brazil (coastline in black). Colors show correction factors for NorESM1-M March mean precipitation. Red contours indicate EWEMBI 1979–2013 March mean precipitation with dotted/solid/dashed lines at 2/4/6 mm/day, respectively.



Figure 26: Time series of regional mean (left) and maximum (right) precipitation as simulated by NorESM1-M for March 2031 under RCP8.5 (grey) and as bias-corrected with different correction factor cap values (colors, see legend). Spatial mean and maximum values were calculated over the region that is outlined in Fig. 25.



Figure 27: Empirical cumulative distribution functions (eCDFs) of 1979–2013 regional mean (left) and maximum (right) daily precipitation in March for EWEMBI (solid orange) and NorESM1-M (solid black). Dashed lines show fits of these eCDFs with log-normal CDFs. Uncorrected (grey) and bias-corrected (color coding as in Fig. 26) NorESM1-M regional mean and maximum precipitation on 7 March 2031 under RCP8.5 are indicated with vertical solid lines. The numbers next to these lines in the legend are the sigma levels of the corresponding events relative to the fitted 1979–2013 EWEMBI (first number) and NorESM1-M (second number) distributions. Spatial mean and maximum values were calculated over the region that is outlined in Fig. 25.



Figure 28: Changes in 1979–2013 monthly mean precipitation biases after a reduction of the correction factor cap value from 10 to 4 for IPSL-CM5A-LR. Colors show increases of biases in mm/day. Shading indicates areas where biases more than double.



Figure 29: Horizontal resolution of atmospheric model versus equilibrium climate sensitivity (from Sherwood et al. 2014) of various CMIP5 GCMs with ISIMIP2b GCMs in orange and ISIMIP fast track GCMs in red (GCMs included in both selections are orange and red). Note that MIROC-ESM(-CHEM) means that the corresponding point in the plot shows the horizontal resolution of the atmospheric model of MIROC-ESM-CHEM and the equilibrium climate sensitivity of MIROC-ESM which is very similar to that of MIROC-ESM-CHEM (Shingo Watanabe, personal communication).