**Hydrology and
Earth System
Sciences**

Climate model bias correction and the role of timescales

J. O. Haerter[1], S. Hagemann[1], C. Moseley[1], and C. Piani[2]

[1]Max Planck Institute for Meteorology, Hamburg, Germany
[2]Abdus Salam International Centre for Theoretical Physics, Trieste, Italy

Abstract. It is well known that output from climate models cannot be used to force hydrological simulations without some form of preprocessing to remove the existing biases. In principle, statistical bias correction methodologies act on model output so the statistical properties of the corrected data match those of the observations. However, the improvements to the statistical properties of the data are limited to the specific timescale of the fluctuations that are considered. For example, a statistical bias correction methodology for mean daily temperature values might be detrimental to monthly statistics. Also, in applying bias corrections derived from present day to scenario simulations, an assumption is made on the stationarity of the bias over the largest timescales. First, we point out several conditions that have to be fulfilled by model data to make the application of a statistical bias correction meaningful. We then examine the effects of mixing fluctuations on different timescales and suggest an alternative statistical methodology, referred to here as a *cascade bias correction method*, that eliminates, or greatly reduces, the negative effects.

1 Introduction

One of the greatest challenges facing modern society in a changing climate is the management of risk associated with hydrological extremes, namely floods and droughts (Vorosmarty et al., 2000; Oki and Kanae, 2006). Risk is a concept expressed in statistical terms, hence, proper management of risk tied to future events must be informed by statistically correct forecasts. Numerical simulations are the principal tool in climate forecasting and hydrological models are used

to obtain simulations of future components of the hydrological cycle. Ordinarily, output fields from climate models, regional or global, are used to force future hydrological simulations. To varying extent, all numerical models suffer from systematic error, i.e. the difference between the simulated value and the observed. Bias is defined as the time independent component of the error. It is well known that some form of pre-processing is necessary to remove biases present in the simulated climate output fields before they can be used for this purpose (Sharma et al., 2007; Hansen et al., 2006; Christensen et al., 2008). However, bias correction cannot correct for incorrect representations of dynamical and/or physical processes and, as will be detailed in this article, model data must provide an adequate representation of the physical system from the outset, to make statistical bias correction applicable.

In the simplest formulations of bias correction only the changes in a specific statistical aspect of the simulated fields is used. The change is applied directly to present day observations to obtain a field which is then used to force the hydrological models. Often the change in mean value or the variance is employed. This is tantamount to correcting the observations with an additive or multiplicative gridded constant. More advanced bias correction methodologies correct for more than one explicitly chosen statistical aspect (Leander and Buishand, 2007 and applied by Hurkmans et al., 2010; Widmann et al., 2003; Schmidli et al., 2006).

Hydrological processes depend on the entire distribution function of precipitation intensity and temperature. For example, extreme hydrological conditions are often caused by unusual precipitation amounts or high temperatures. Persistent heavy precipitation over several days can lead to floods while the absence of precipitation along with high temperatures is often the cause of drought. Hence, improvements on simple bias correction methods can be made when adjusting the entire probability density function (pdf) of the simulated

fields to that of the observations. Consequently, adjusting the likelihood of the occurrence of a given magnitude of daily precipitation or temperature, allows a more adequate representation of the risk of flood and drought by the corrected data (Wood et al., 2002; Hay and Clark, 2003; Dobler and Ahrens, 2008; Piani et al., 2010a,b). These methods are also sometimes referred to as "quantile mapping" (Deque, 2007), "histogram equalization" and/or "rank matching". A recent review is given by Maraun et al. (2010). In Piani et al. (2010a) daily precipitation and temperature fields were corrected by fitting probability density functions to the modeled and observed data. A mapping of the corresponding fit-coefficients was then defined. The method was developed and tested on distinct periods within the control period. In Piani et al. (2010b) the method was refined by first sub-dividing the climatological year into monthly segments and performing bias-corrections within each segment separately. Also, the method was improved by employing *transfer functions* (TFs) to map modeled to observed quantities directly, which reduces the number of required parameters.

While these existing approaches do offer a means of equalizing the statistical properties of modeled and observed climate data, they do not take into account that oscillations on different timescales are caused by disparate physical mechanisms. When a bias correction is performed where all data are grouped into one joint dataset, the fluctuations on different timescales are mixed. This can blur the interpretation for future scenario corrections. Therefore, in the current study we propose a modification of the existing methodology to separate different timescales by performing a cascade of bias corrections.

In Sect. 2 we present the general methodology of the statistical bias correction and outline some possible obstacles. In Sect. 3 actual model and observational data are used to probe to what extent these obstacles are relevant. In Sect. 4 we offer an improvement of the method by producing a cascade of bias corrections. Section 5 features a more general discussion on bias correction methodologies and Sect. 6 concludes.

2 Statistical bias correction

Statistical bias correction (SBC) is a mathematical procedure (a functional) that maps the probability density function (pdf) of model data onto that of the observations:

$$\text{TF}(x_{\text{mod}}) = F_{\text{obs}}^{-1}(F_{\text{mod}}(x_{\text{mod}}))$$

where F_{obs} (F_{mod}) is the cumulative distribution function of the observed (modeled) data x_{obs} (x_{mod}). Hence, in general it is an operation that acts on all moments of the distribution. In this sense, SBC is not a model of the physical world in itself. It completely relies on information being contained in the climate model data albeit with a systematic discrepancy from the observational data. SBC can hence not make up
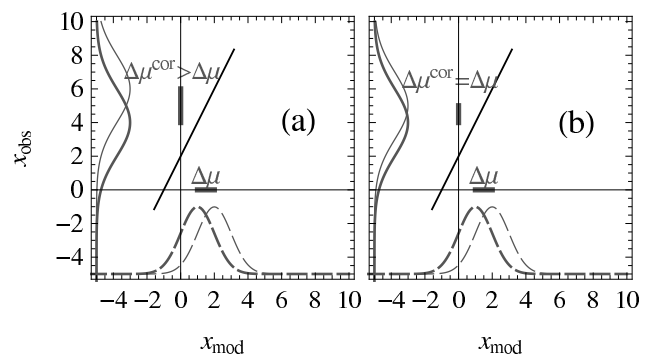
Fig. 1. Two possible options for performing a statistical bias correction on the distribution of the variable x in arbitrary units; Heavy (thin) lines are control (scenario) period data; Solid (dashed) lines are observed (modeled) data; solid straight line represents the linear transfer function. Note that in **(a)** the climate change signal is enhanced by applying the bias correction ($\Delta\mu^{\text{cor}} > \Delta\mu$) while in **(b)** it is not ($\Delta\mu^{\text{cor}} = \Delta\mu$).

for fundamental qualitative flaws of the climate model. To conceptually illustrate the procedure, in the following sections we repeatedly make reference to a bias correction of a given normally distributed climate variable x with mean μ and standard deviation σ. In the case of normal distributions of daily data, a *perfect* bias correction need only adjust the first two moments of the distribution. To construct a mapping between the observed and the modeled data, a *transfer function* is derived (as described in Piani et al., 2010b).

2.1 Construction of transfer functions

In Fig. 1 the TF is shown for the simple Gaussian example. In this case, the control period (heavy lines) means and standard deviations are $\mu_{\text{mod,con}} = 1$, $\mu_{\text{obs}} = 4$, $\mu_{\text{mod,sc}} = 2$, $\sigma_{\text{mod,con}} = \sigma_{\text{mod,sc}} = 1$, $\sigma_{\text{obs}} = 2$ where the subscripts mod (obs) indicate model (observations), and con (sc) refer to the control (scenario) period. Hence, the slope of the TF is the ratio of the standard deviations of observed and modeled data, namely $\sigma_{\text{obs}}/\sigma_{\text{mod,con}}$. This factor stretches the distribution of modeled data to match the width of the observed. In the case of normal distributions, the TF is always linear and the corrected mean $\mu_{\text{mod}}^{\text{cor}}$ and variance $\sigma_{\text{mod}}^{\text{cor}}$ are

$$\mu_{\text{mod,sc}}^{\text{cor}} = \mu_{\text{obs}} + \frac{\sigma_{\text{obs}}}{\sigma_{\text{mod,con}}}\left(\mu_{\text{mod,sc}} - \mu_{\text{mod,con}}\right) \quad (1)$$

$$\sigma_{\text{mod,sc}}^{\text{cor}} = \frac{\sigma_{\text{obs}}}{\sigma_{\text{mod,con}}}\sigma_{\text{mod,sc}}. \quad (2)$$

Such TFs are derived using observational and model data of the same *control* period, hence a period of known boundary conditions, such as carbon dioxide concentrations. Once the TF is computed, it can be used to produce bias corrected data for the future (scenario) model run, where observational data are not available. To apply the bias correction to a modeled

time series, for each individual value of the model data, the transfer function is used to map this value onto a modified (bias corrected) value. A more detailed description of the method is available from Piani et al. (2010b).

As the observed data have a larger standard deviation than the model data (Fig. 1a), the slope of the TF becomes greater than 1. The original model data are hence stretched, both in the control and scenario period. We call the change in the distribution mean $\Delta\mu \equiv \mu_{\mathrm{sc}} - \mu_{\mathrm{con}}$ between the control and scenario period the *mean climate change signal*. In this case, the bias correction leads to an increase in the mean climate change signal from $\Delta\mu = 1$ to $\Delta\mu^{\mathrm{cor}} = 2$ after the bias correction is applied. Hence, the bias correction has caused a modification in the model produced mean climate change signal.

Another option to correct the data would be to apply different types of corrections to the mean and to the standard deviation. In Fig. 1b we show the same case as in Fig. 1a, but the mean is corrected by simply adding the mean uncorrected climate change signal produced by the model to the original observed mean. The standard deviation is adjusted as before. The correction is still a two-parameter correction, with one parameter multiplicatively adjusting the variance and the other additively adjusting the mean.

To define the TF, one could equally well choose other fluctuations instead of daily fluctuations, for example seasonal fluctuations. In the case of temperature, seasonal fluctuation, similar to diurnal fluctuations, are caused – directly or indirectly – by the changes in the solar radiance and could be seen as a response of the system to such changes in the energy budget. However, climate change is usually assumed to be due to changes in boundary conditions, which bring about the greenhouse effect. Greenhouse gases – such as carbon dioxide – capture long-wave radiative energy emitted by Earth.

In the following, a simple example is given to demonstrate the consequences of choosing a certain timescale at which statistics are produced. Let us take model and observational data with matching pdfs of monthly mean data. However, take the day-to-day variability of the model to be larger than that of the observations. We exemplify this situation in Fig. 2 where we show synthetic data sampled randomly from Gaussian distributions. Consequently, the histograms of the daily data (right column of Fig. 2) have significantly different widths. If a TF were constructed from these data, the slope of the line in Fig. 1a would be greater than unity and exaggerate the variance in the monthly means for both the control and scenario period. The choice of day-to-day fluctuations in developing the bias-correction would consequently lead to vastly different results than the alternative choice of monthly mean statistics.

Hence, it is important to note that a bias correction mixing statistics that occur at different timescales may lead to unwanted results. In Sect. 3 we discuss how strong such effects are in actual model and observational data.
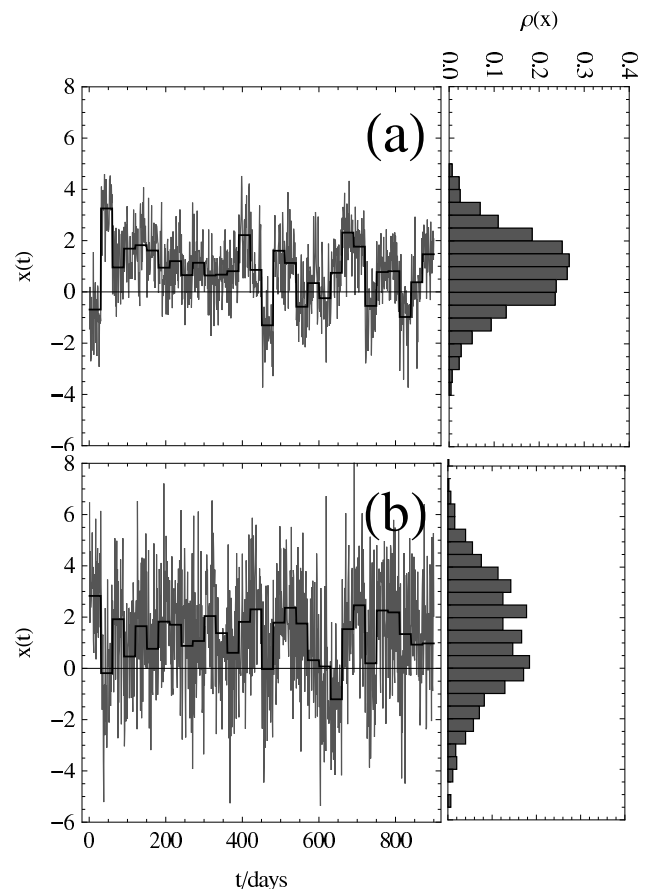
Fig. 2. Example of modeled – **(a)**, gray – and observed – **(b)**, gray – time series of daily data, monthly means indicated by heavy black lines; Panels on the right are normalized histograms of the daily time series.

3 Bias correction with GCM data

An actual bias correction is now performed with daily data from a GCM and global observational data. The GCM is the Max Planck Institute for Meteorology ECHAM5/MPI-OM model (Roeckner et al., 1999; Jungclaus et al., 2006). We use the data generated for the fourth assessment report of the Intergovernmental Panel on Climate Change (IPCC). While this GCM is a state-of-the-art global climate model, not all physical processes present in the real system can be captured. In fact, in some regions of the globe fundamental misrepresentations are likely, especially when performing one-to-one comparisons between model gridboxes and observational data. Such comparisons are especially misleading when processes have small spatial extent – e.g. in the case of storm track patterns. Also, the representation of the timing and frequency associated with the El Nino Southern Oscillation phenomenon may not be reproduced adequetely at a fundamental level. Global observational data are taken from the dataset synthesized within the EU-WATCH (WATer and
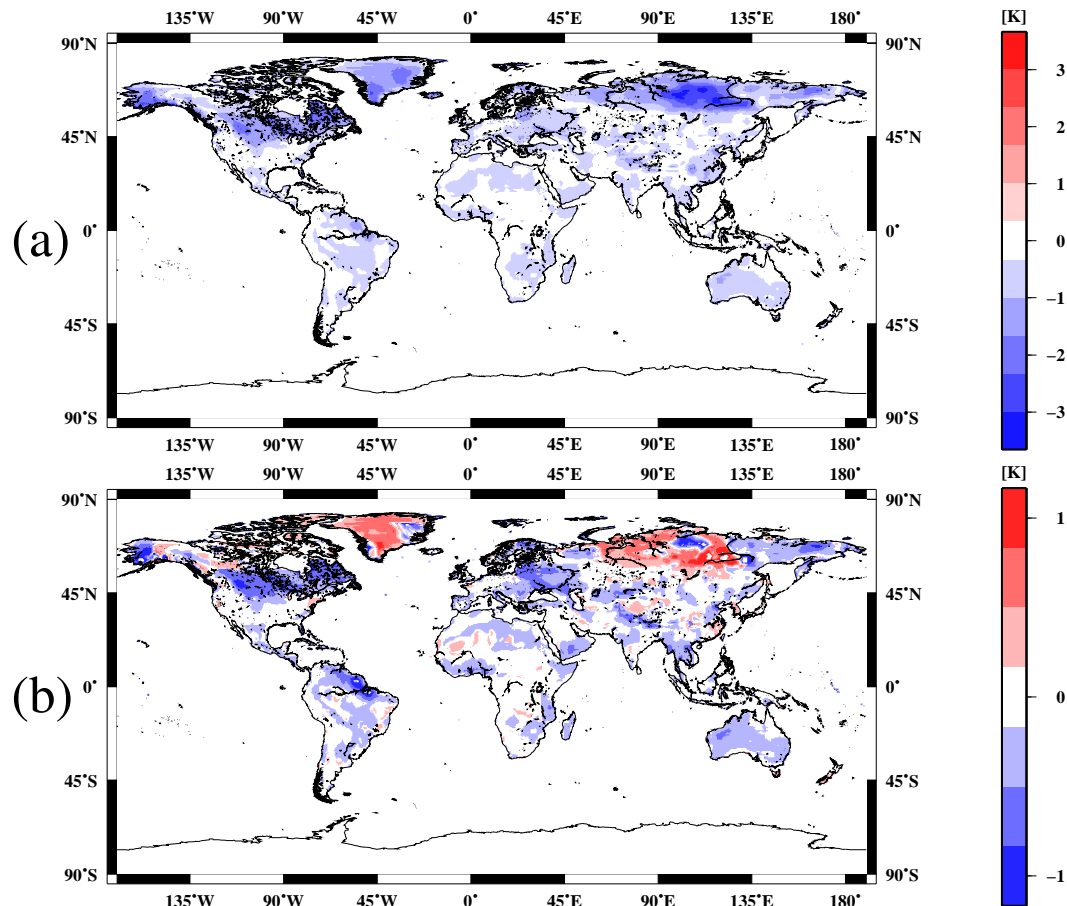
Fig. 3. Temperature bias correction data for ECHAM5 corrected using EU-WATCH forcing data for the years 1960–1999 for December: Improvement in standard deviation between original and bias corrected model (negative values mean improvement); **(a)**, daily values; **(b)**, monthly mean values.

global CHange, http://www.eu-watch.org) project, which is sometimes referred to as *WATCH forcing data* (WFD, Weedon et al., 2010) in the following. This dataset is a combination of monthly observed data and daily reanalysis data and is taken as the *best guess* approximation to actual observations that is available for the globe and for several decades. Both datasets provide overlapping data for the 40-year period from 1960 to 1999.

3.1 Temperature bias correction

For temperature, we have performed a bias correction for this period using the linear transfer function method as described in Sect. 2 and in Fig. 1a. In the following we will refer to this as the *standard bias correction*. Hence, within each gridbox, for every month we group all available data into one joint dataset (say, all Januaries from 1960 to 1999 yielding 1240 daily values) and perform a mapping of the statistical properties of the model data to those of the observational data. The bias correction – by construction – will make the mean and the variance of corrected daily model and

observational data equal. As taking the mean is a linear operation, both the daily and monthly distribution means will be equal when the corrected model and the observations are compared. This does not apply to the variances as we have exemplified in Sect. 2.

To examine this aspect, we investigate whether the bias correction has improved on the discrepancy between the modeled and the observed standard deviation. In Fig. 3a we first present the change in discrepancy of standard deviation of the daily values caused by the bias correction:

$$\Delta \, \mathrm{SD}(T) = \left| \mathrm{SD} \left(T_{\mathrm{mod,cor}} \right) - \mathrm{SD} \left(T_{\mathrm{obs}} \right) \right| \qquad (3)$$
$$- \left| \mathrm{SD} \left(T_{\mathrm{mod,org}} \right) - \mathrm{SD} \left(T_{\mathrm{obs}} \right) \right|.$$

Here, SD denotes the standard deviation of the distribution, $T_{\mathrm{mod,cor}}$ ($T_{\mathrm{mod,org}}$) represent the corrected (original) model temperature data and T_{obs} are the observed temperature data. As Fig. 3a shows, in all regions of the globe the value is not positive, meaning the standard deviation in the corrected case is closer to that of the observations. This is not surprising as this feature is built into the bias correction.

The computation in Eq. (3) is now repeated for the monthly mean values of temperature, obtained before and after applying the bias correction based on daily values. The result is shown in Fig. 3b. While large regions of the globe still show a substantial improvement – meaning negative values – there are also some areas, such as Greenland or Siberia, with a substantial increase of the deviation from observations in the monthly mean standard deviation.

Hence, while the bias correction has led to an improvement of the day-to-day variance, the variance of the monthly means has in fact become less realistic after performing the bias correction.

In Fig. 4a we show the daily and monthly mean standard deviations for the WFD. The first striking feature of this plot is that fluctuations at the daily and monthly scales are not independent (similar patterns in the two panels of Fig. 4a). Regions where there are large day-to-day fluctuations generally also have larger fluctuations of the monthly means from one year to the next. The high latitudes show larger fluctuations than the tropics. In the low and mid latitudes, the difference between the modeled temperature fluctuations and the observed (Fig. 4b) shows no systematic pattern, neither for the daily nor the monthly mean standard deviation and there is no clear dominance of a positive or negative signal. However, the ECHAM5 model appears to generally underestimate day-to-day variability in the high latitudes while the bias is more mixed in the case of interannual fluctuations. Comparing the panels in Fig. 4b, we find there are some regions – such as South America, parts of Africa and Australia – where the bias in the daily and interannual fluctuations is rather similar. We now turn to the comparison of the bias corrected data (Fig. 4c) with the observations. Clearly, the bias in the day-to-day fluctuations has been all but removed. However, the improvement of the interannual fluctuations is only obvious in regions of overlapping day-to-day and interannual biases. Hence, only when short-term and long-term fluctuations are aligned, the bias correction will lead to improvements on both timescales. In some regions, most notably Greenland and parts of Siberia, the bias corrected signal has led to a worsening of the interannual fluctuations as the day-to-day and interannual bias have opposite signs in those regions (Fig. 4b, blue and red colors). We return to the remaining panels of this figure in Sect. 4.2.

To better understand the effect of mixing timescales, we now choose a single gridbox in Siberia (61.25° N, 112.25° E), where it is particularly pronounced. In Fig. 5 we present the daily (Fig. 5a) and monthly mean (Fig. 5b) time series for the observations as well as the original and corrected ECHAM5 simulation data. The daily time series shows that the observations produce rather strong oscillations as compared to the original model data. Therefore, the corrected model data become somewhat *stretched* in the vertical direction to equalize the variances. In the case of the monthly mean time series the oscillations of the observations are not very strong and perhaps even smaller than those produced by

the model. However, due to the adjustment of the day-to-day variance, the corrected monthly time series acquires an even larger amplitude of oscillation than before. Hence, the worsening of the monthly mean statistics in this region is caused by the underestimation of the day-to-day variability by the model compared to the observations.

In Fig. 5a we also present the histograms of the original and corrected model data in comparison with the observations. They show that there is a clear equalization of the mean and variance. Conversely, the standard deviation of the observed monthly means is 3.1 K, and for the original (corrected) model, it is 3.5 K (4.7 K). Hence, while the original model variance was rather close to that of observations, the corrected value is nearly 50% too large.

3.2 Precipitation bias correction

In the case of daily precipitation, the discussion of Sect. 3.1 is less relevant. On the one hand, the distribution of precipitation intensities is never of a symmetric or even Gaussian shape as in the case of temperature and the precipitation distribution is bounded from below by zero. To approximate the distribution function of daily precipitation intensity, Gamma distributions or other rapidly decaying functions of intensity have been used in the past (Piani et al., 2010a; Gutowski et al., 2007; Wilson and Toumi, 2005; Haerter et al., 2010). The common feature of such functions is that they have well-defined means (as opposed to power-law distribution functions on short timescales as reported in Haerter et al., 2010) and the mean and variance are coupled.[1]

On the other hand, the monthly precipitation mean is not the average of 30 values of the random variable as is the case for temperature. For precipitation, non-zero measurements are recorded only on a few days of the month. Hence, the monthly mean value is often dominated by only a small number of daily precipitation records and hence is often rather well approximated by one or two large events. Furthermore, precipitation processes on the daily and monthly timescales are often closely related, e.g. no rain events (short range timescale) over northern Europe during strong Euro-Atlantic blocking regimes (medium range timescale). Hence, computing the variance of monthly means and that of daily data consequently often leads to rather similar results. To check this, we have computed the variance of monthly means in the WFD, the original and the corrected model data (Fig. 6). The figure confirms that the corrected model data more closely agree with the WFD than the original data, unlike the temperature case shown in Fig. 3.

---

[1]In the case of the Gamma distribution of a random variable x: $\rho(x) \equiv x^{k-1} \frac{\exp(-x/\theta)}{\Gamma(k)\theta^k}$ with the shape parameter k and scale parameter θ the mean is $k\theta$ and the variance $k\theta^2$. Hence, the mean and variance are not independent as in the case of a Gaussian distribution.
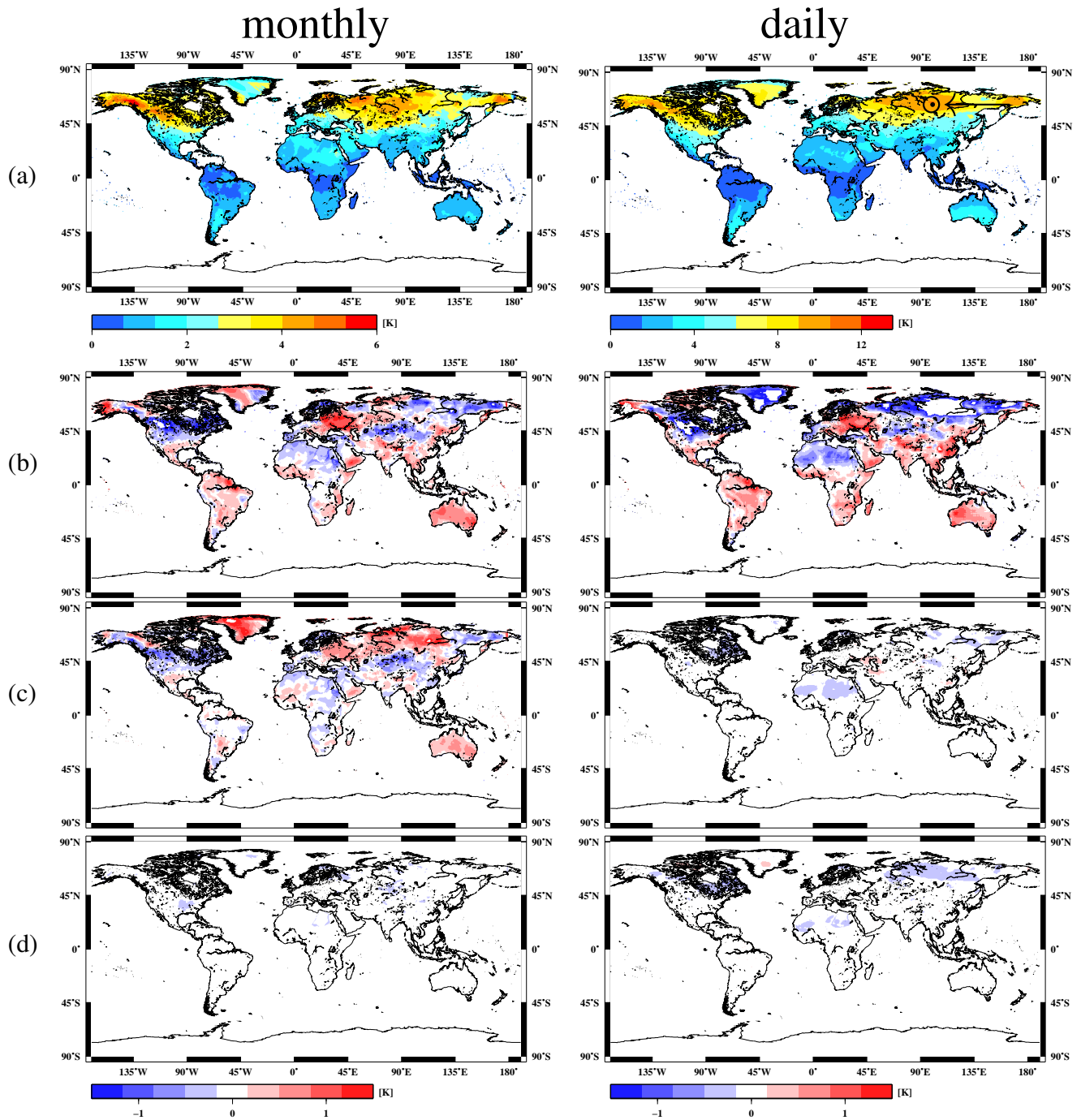
Fig. 4. Comparison of December standard deviation of monthly mean temperature (left) and standard deviation of daily temperature (right) for observations **(a)** and the ECHAM5/MPI-OM model: **(b)**, difference plot of standard deviation of the original ECHAM5 minus standard deviation of observations; **(c)**, same as **(b)** but using bias corrected data; **(d)**, same as **(b)** but using cascade bias correction. Note the change in scale between monthly and daily data in **(a)**. Circle and arrow in **(a)**, right panel, mark a point used for time series in Fig. 5.

Note also that the representation of precipitation extremes and local precipitation patterns by GCMs is often not realistic. Qualitative misrepresentation of precipitation features can be caused by the lack of spatial resolution and

the inability to adequately model subgrid-scale processes involved in precipitation formation. Strong topographic gradients cannot be considered appropriately by most GCMs and the proper parameterization of convective precipitation is an
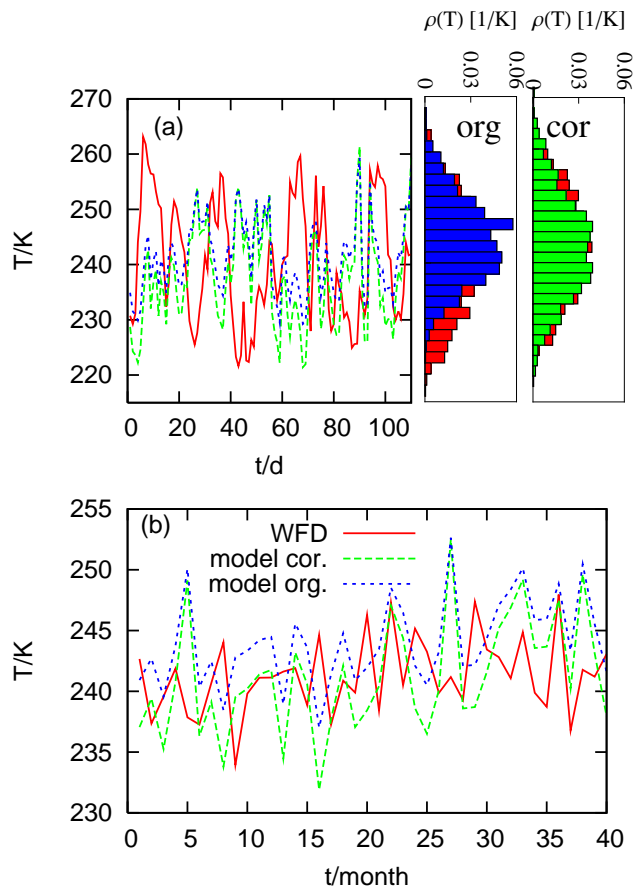
Fig. 5. Time series of observational, original and corrected model data for daily **(a)** and monthly **(b)** values at location (61.25° N, 112.25° E) marked in Fig. 4a; in **(a)** we also show the histograms of the observational (red), the original (blue) and corrected (green) model data.

active field of research. These shortcomings of current-day GCM model data make it difficult to apply grid-point by grid-point statistical bias correction in all regions of the globe.

In summary, we have shown that the statistical bias correction as used in Piani et al. (2010b) does equalize the statistics of the daily observations. However, this can lead to unwanted results at longer timescales. In the following Sect. 4 we propose an extension of the method.

4 Improved statistical bias correction

Before discussing a possible improvement of the statistical bias correction we caution that application of the procedure requires several conditions to be fulfilled: the model data at hand must constitute a realistic simulation of the actual physical system. The model bias is a systematic quantitative time-independent transformation of the probability distribution function of modeled data. The bias correction can neither improve on the representation of fundamentally misrepresented

physical processes nor can it account for misrepresentation of the transient response to green house gas emissions. To give an example, it is not to be expected that biases caused by a misrepresentation of phenomena such as the *El Nino Southern Oscillation* can be corrected by bias correction. If the extent of the warm tongue or the timing of El Nino are misrepresented, it is not possible to improve such behavior in any way by employing bias correction. It is further not proven that a biased response to the solar irradiation cycle should be scale dependent. Investigating the cause of model bias is a central question in the climate modeling community and goes beyond the scope of this article. To make our argument, we need to assume that the model data are representing the actual physical processes of the climate system – albeit with a quantitative departure from the actual observational data. This is a strong assumption that may not apply for all model data and should be kept in mind when performing similar corrections on other data.

We now address cases where these assumptions are fulfilled, i.e. the model data do constitute a realistic description of the actual physical phenomena, while the model probability density function may differ from that of the observations. To improve the statistical bias correction and remedy the shortcomings mentioned in Sects. 2 and 3 we propose a *cascade bias correction* methodology. The problems discussed there may be caused by treating data originating from mechanisms that act on different timescales on equal footing: day-to-day variability is caused by fluctuations of local weather systems, the magnitude of the diurnal cycle and evaporative processes.

4.1 Description of cascade bias correction method

The cascade bias correction breaks down the original process into its different timescales thereby avoiding their mixing. As an example, take a long time series (several decades) of daily temperature data. We then break down the time series into segments of months and combine all months of January into a new time series. The motivation for doing so is that we assume the temperature data in January to fluctuate little as a result of the systematic seasonal dependence of statistical expectation value within the month but rather due to the natural fluctuations from one day to the next. The choice of one month is of course somewhat arbitrary, two months or two weeks could be equally acceptable choices, depending perhaps on the geographic region. A trade-off has to be made between the statistical permutability of the values within the segment and the required size of the sample of fluctuations required to produce probability density functions of the natural fluctuations.

In the following we assume the segment to be one month and the discussion focuses on several years of data for this particular month (say January). After such a segmentation has been completed, each daily value can be expressed relative to the mean within its segment, hence a daily temperature
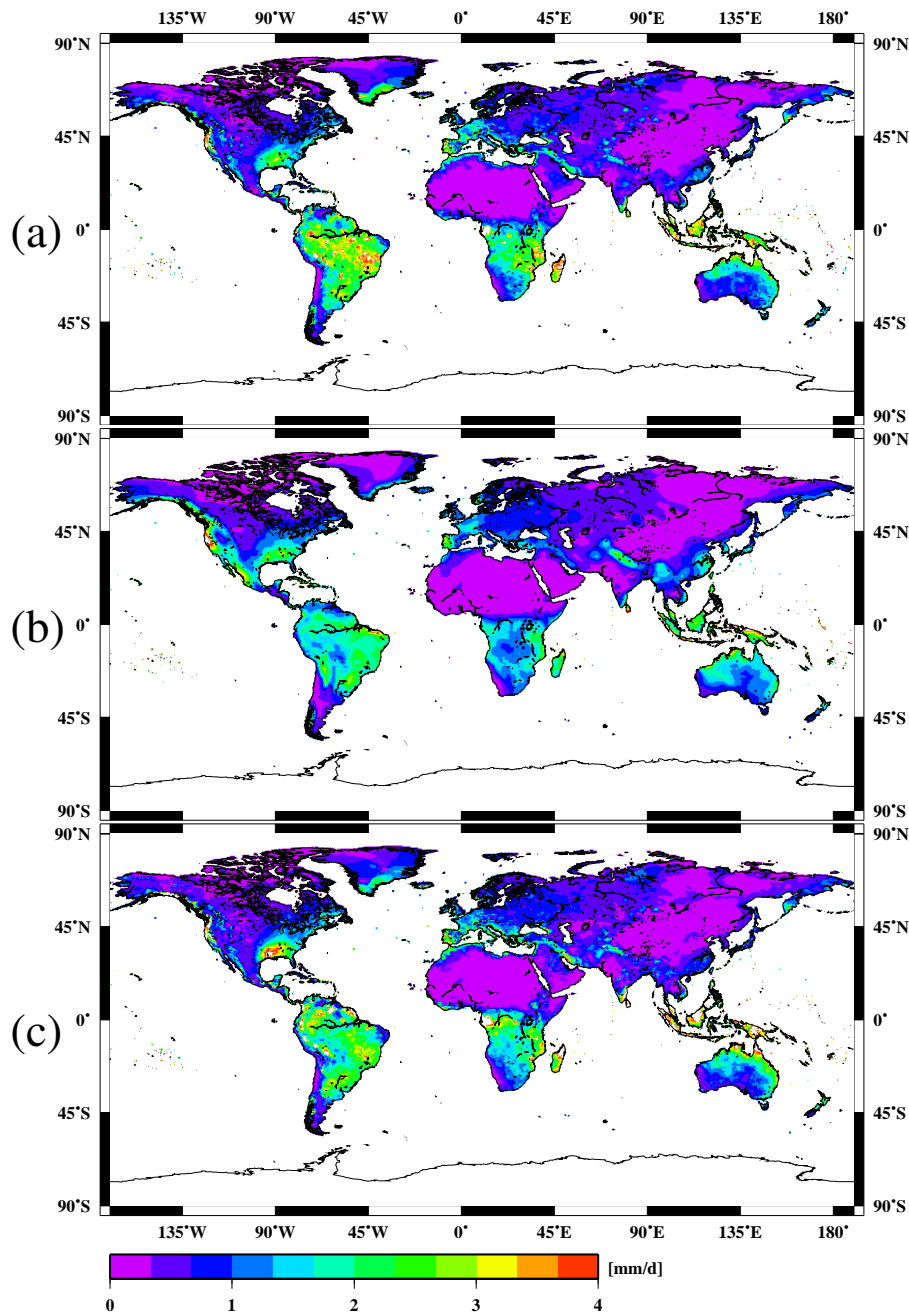
Fig. 6. Standard deviation of monthly mean precipitation from 1960–1999 for December for WFD **(a)**, original model data **(b)** and bias corrected model data **(c)**.

value $T_{i,j}$ corresponding to month i and day j will be mapped onto the *anomaly*

$$T'_{i,j} \equiv T_{i,j} - \overline{T_i} \tag{4}$$

where $\overline{T_i}$ is the monthly temperature mean within the given month i and primed variables are anomalies. More precisely, at any given day j of month i the monthly mean should be defined as the running monthly mean value involving the previous and subsequent 15 days to avoid jumps at the interfaces between months. We skip this point to simplify the discussion but we suggest that it could be included in the algorithm to avoid jumps at the interfaces between calendar months. Note that using running mean values can modify the monthly means of calendar months which are conventionally defined as temporal entities in the community. Hence, the benefits of working with running means may be outweighed by the difficulty of the interpretation of the resulting corrected time

series. The set of anomalies $T'_{i,j}$ of all months i can then be used to compute the distribution of daily anomalies. To perform a bias correction, the same operation will be performed on model and observational data and a TF (which we call f_{daily}) will be derived for the daily anomalies, similar to the method described in Sect. 2. A bias corrected daily value is then obtained as $T'^{\text{cor}}_{i,j} = f_{\text{daily}}(T'_{i,j})$.

In the following step the monthly means $\overline{T}_i$ are considered.[2] Statistics of all available monthly means $\overline{T}_i$ will be constructed for both model and observations. Transfer functions can then be derived in the spirit of Sect. 2 but for the monthly means (which we call f_{monthly}). A bias corrected monthly mean of month i is then obtained as $\overline{T}^{\text{cor}}_i = f_{\text{monthly}}(\overline{T}_i)$. To obtain the bias-corrected daily model data, one would simply combine the corrections on the two levels of the cascade to yield

$$T^{\text{cor}}_{i,j} = \overline{T}^{\text{cor}}_i + T'^{\text{cor}}_{i,j} \tag{5}$$

for the bias-corrected value of $T_{i,j}$. This procedure would apply similarly for the other months of the year.

Furthermore, if a sub-daily – say hourly – bias correction is intended (as in Sect. 4.2.2), the cascade should be continued towards smaller time-intervals. The day would be divided up into hourly segments and within each segment the temperature would be re-defined relative to the daily mean value. The procedure would then begin at the hourly level but continue as stated above. The final bias corrected hourly values would then become

$$T^{\text{cor}}_{i,j,k} = \overline{T}^{\text{cor}}_i + T'^{\text{cor}}_{i,j} + T''^{\text{cor}}_{i,j,k} \tag{6}$$

where $T_{i,j,k}$ is the temperature value at month i, day j and hour k. Note that all TFs except that for the longest time-interval for the various cascade steps will have only a slope parameter as the means are zero. Hence, the number of parameters required for a cascade of n steps will be $n+1$.

4.2 Application of cascade method

The method introduced in Sect. 4.1 is now applied to actual model and observational data. In the first case we intend to improve on the bias correction of Sect. 3.1. In the second, we apply a three-tier cascade to hourly data.

4.2.1 GCM cascade bias correction

We return to the global data discussed in Sect. 3.1. We employ a two-tier correction as described in Sect. 4 and Eq. (5),

---

[2]If sufficiently long time series exist – such as in millenium simulations in some climate models – multi-year fluctuations may be identified and it could be adequate to define multi-year oscillatory intervals and continue in the same fashion as before, hence defining the means over these multi-year intervals and computing the anomalies of the $\overline{T}_i$ w. r. t. these means. However, for most practical purposes, the cascade will end at this level.

hence we first produce monthly means $\overline{T}_i$ of the observed and simulated data for the 40-year period between 1960 and 1999. We then extract the day-to-day anomalies $T'_{i,j}$ relative to these monthly means and produce a two-parameter correction (slope and offset) for the $\overline{T}_i$ and a one-parameter correction (slope) for the anomalies. The results are presented in Fig. 4d. Clearly, both the monthly mean and the day-to-day statistics now agree very well with those of the observed data. A change of slope of the TF implies a change in the climate change signal (Sect. 2). Therefore, we want to assess the changes in the climate change signal brought about by the different correction methods. Using the IPCC B1 emission scenario (Solomon et al., 2007) and the projected ECHAM5 data for the 30-year period 2070–2099 we compare the original model data, the standard bias-correction model data, and the cascade correction model data (Fig. 7). In all three cases, the general warming of the original model data is preserved, with a positive gradient of warming with increasing latitude. However, we note that the extreme northern latitudes generally acquire a stronger warming signal in the standard correction relative to the uncorrected model (Fig. 8a) while the cascade corrected climate signal is not as strongly enhanced (Fig. 8b) in these regions. In some parts of these regions the modification of the climate change signal is actually reversed in the cascade method. In South and Central America there is larger agreement between the two correction procedures since interannual and interday fluctuations are aligned there (compare Fig. 4b).

4.2.2 Three tier cascade

We now intend to produce a three-tier cascade correction. Therefore we employ data at an hourly resolution to allow statistics for the hourly, daily and monthly periods. The observational temperature data are provided by the German Weather Service for a station in Aachen, Germany. The model data are from the Max Planck Institute for Meteorology regional climate model (REMO), which is available at a 10 km horizontal resolution (Jacob et al., 2008). We choose the gridbox closest to the station to achieve an optimal comparison. The overlapping time-period for the two datasets are the 20 years from 1960 to 1979. While the method described in Sect. 4 could also be applied to precipitation statistics, we are here comparing station data to those of model gridboxes. For temperature, this is appropriate as spatial fluctuations are small within a 10 km distance and the region studied has small topographic gradients. For precipitation, the gridbox value has to be interpreted as a spatial mean over its area and the comparison with station rain gauge measurements is cumbersome.

In the following we generally choose the month of January as an example. In Fig. 9 we present the hourly time series of the observational and model data. Note the different characteristics of the time series that immediately strike the eye: The overall spread of the model data is smaller than that of
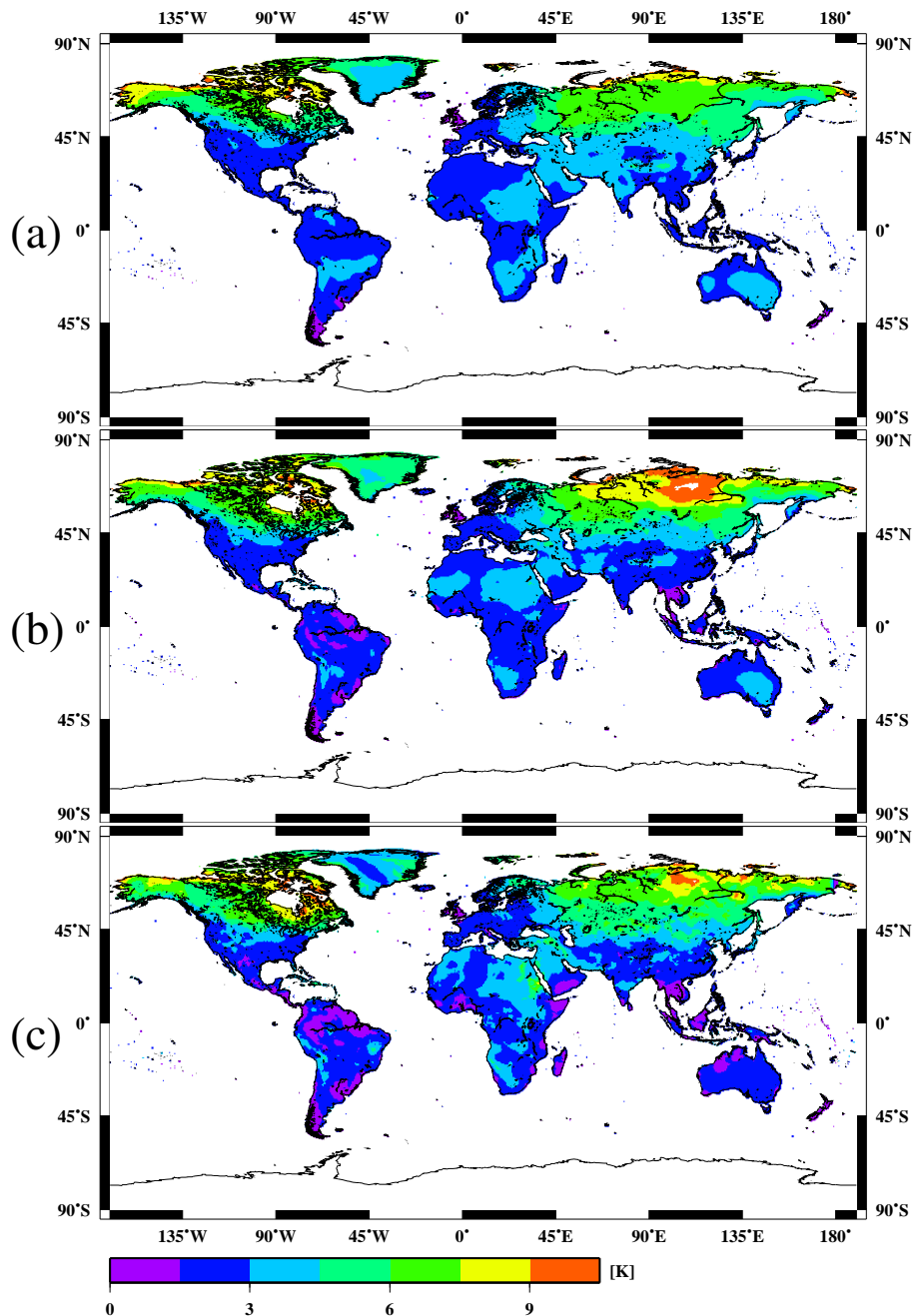
Fig. 7. Temperature climate change signal for the B1 emission scenario projected by ECHAM5 for the period 2070–2099 relative to 1960–1999 for December; **(a)**, model without bias correction; **(b)**, standard bias correction; **(c)**, cascade bias correction.

the observations and the interannual fluctuations (from one January to the next) are more pronounced in the observations.

We proceed by computing the daily anomalies relative to the monthly mean values (histograms in Fig. 9) as described by Eq. (4). We then iterate the same procedure for the sub-daily scale by defining intra-day anomalies relative to the daily mean. To illustrate the result, for the hourly anomalies we display those between 16 and 17 h of each day, hence 20×31 values. The probability density function

of the daily mean model fluctuations is statistically significantly narrower than that of the observations with standard deviations of 2.6 K (3.9 K) respectively. The hourly anomalies (relative to the interpolated daily mean) are small compared to the daily fluctuations and there is no statistically significant difference between the model and observational pdf. Hence, there are different deviations between the model and observed data for the two timescales. Consequently, an algorithm utilizing a blend of fluctuations on all scales to
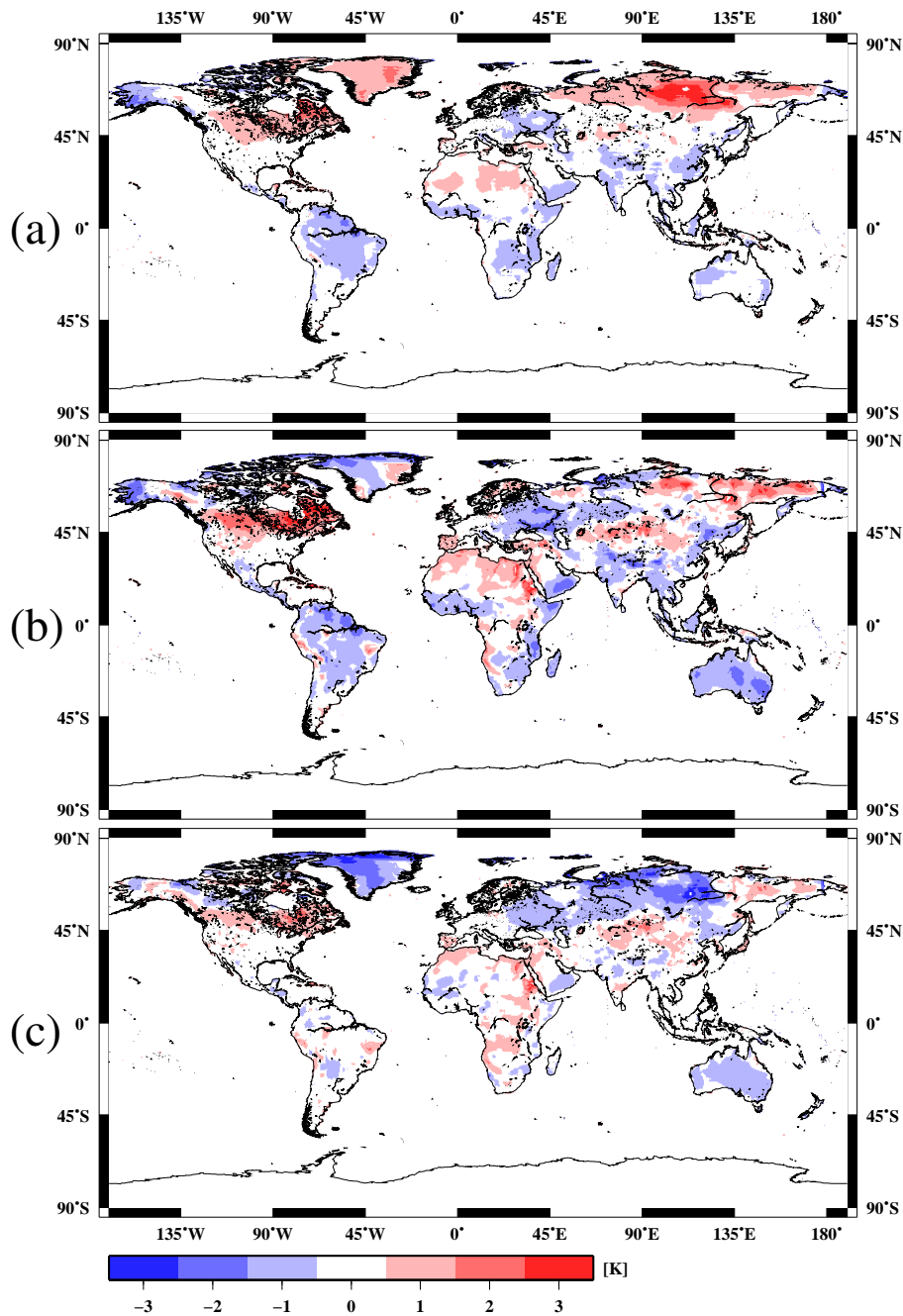
Fig. 8. Modification of climate change signal: **(a)**, by using the standard correction method – panel (b) minus panel (a) of Fig. 7; **(b)**, by using the cascade method – panel (c) minus panel (a) of Fig. 7; and **(c)**, the difference between the correction procedures – panel (c) minus panel (b) of Fig. 7.

derive the correction factors is bound to mix these two statistics and would alter distributions at scales where it should not. This finding underscores the advantage of distinguishing timescales when performing the bias correction.

Now we produce linear TFs for the data at all different timescales, namely hourly fluctuations relative to the daily mean, daily fluctuations relative to the monthly mean and monthly fluctuations. Note that the monthly fluctuations constitute the only case where the TF acquires an offset in addition to the slope. In all other cases, the TF corresponds to a multiplicative correction factor as the means of fluctuations are zero by definition. Hence, in our three-level example, in total four correction parameters are required for each hour of the day and month. After the correction factors
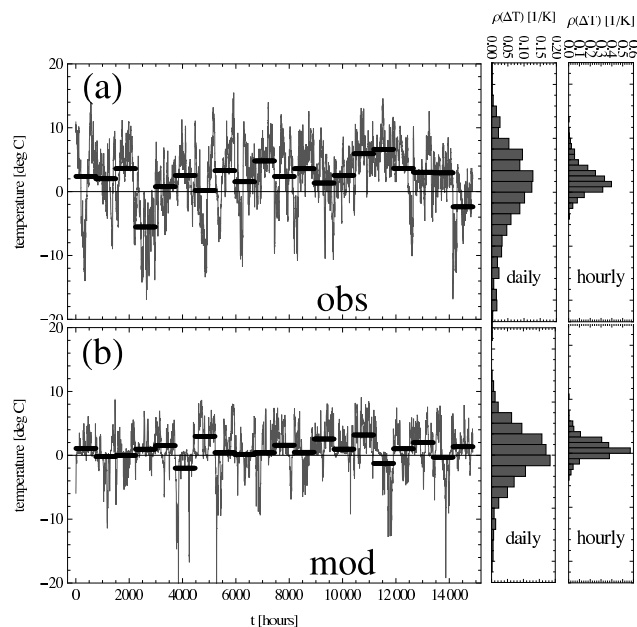
Fig. 9. Hourly time series of 20 Januaries for observations – **(a)**, gray – and model – **(b)**, gray. Black horizontal lines indicate monthly averages. Histograms of daily anomalies (relative to the monthly mean, central column) and hourly anomalies (relative to the daily mean, right column).

are computed, the bias-corrected time series is re-composed by merging the individual components (Fig. 10) as given by Eq. (6). In this figure we show data for both January and July. We also show the comparison with the standard bias correction. In the standard correction, fluctuations are not corrected separately as in the cascade correction, however, a correction is performed for each hour of the day, this means all first hours of all January days (31 days × 20 years) would yield one correction. Note that in January the diurnal fluctuations are enhanced in both corrected series in comparison with the original model data, while the day-to-day fluctuations remain rather unchanged. Furthermore, an overall shift is applied to the data as the model appears to have a general cold bias during these months. When comparing the standard and cascade corrections, we find that on days with extreme fluctuations – such as at the beginning of the time series in panel a – the standard corrected (dashed gray) time-series is 2 K lower than the cascade corrected series. The reason is that the diurnal cycle is enhanced in the corrected version, but in the standard correction the diurnal range is measured in absolute terms rather than relative to the diurnal mean as in the cascade correction. This leads to an exaggeration of the correction in the standard version.

The situation is somewhat different in July where the model produces more realistic diurnal fluctuations by itself (more similar to those in the observations), hence the corrected data mirror the original data rather closely and only an overall shift of roughly 2 K towards higher temperatures is applied to correct the cold bias present in the model data.
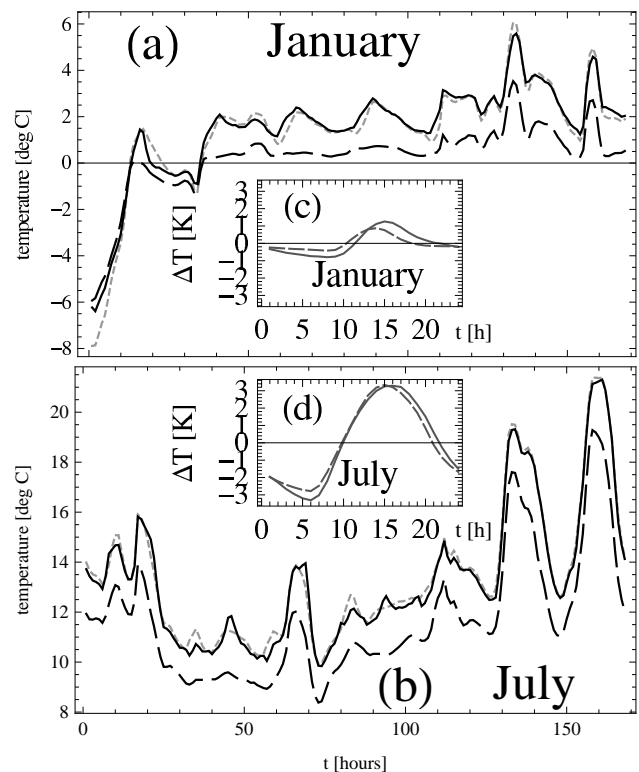
Fig. 10. Original and corrected hourly time series for one week in January **(a)** and July **(b)**, long-dashed line is original model data, solid line is cascade-corrected model data, short-dashed line is data corrected using standard method; Inset: Diurnal cycle for January **(c)** and July **(d)** for model (dashed) and observations (solid), respectively.

5 Discussion

In this paper, as in past articles on this topic, bias is intended as the time independent component of the error. The error is the difference between the simulated value and the observed. Bias correction is done as part of the post processing of simulated data. Hence, it cannot add information or skill to the simulation and, furthermore, it cannot eliminate the error. The sole purpose of bias correction is to eliminate the time independent component of the error if it exists. Crucially, if there is no bias, that is if there is no constant portion to the error, the bias correction methodology leaves the simulation unaltered.

A grid-based bias correction can only be expected to yield meaningful results if there is sufficient correspondence between the modeled and observed behavior within the same gridbox. By *correspondence* we mean a qualitatively adequate representation of the physical phenomenon at hand. For example, if in the model northern and southern latitudes were swapped, such correspondence would be lost and the bias correction would become obsolete. Similarly, if North Atlantic storm tracks were systematically shifted in the

model relative to the observations, the grid-based bias correction would fail. Along with spatial offsets, a grid-based bias correction cannot compensate for temporal offset. If the Indian monsoon were appearing with one or two weeks delay in the model compared to observations, this could not be corrected by this method for similar reasons. Similarly, if the timing or the frequency of the El Nino Southern Oscillation phenomenon were misrepresented by the model, this behavior could not be improved upon by SBC.

Apart from these limitations, one of the main obstacles when applying a statistical bias correction to climate model data is that fluctuations on different scales can mix and lead to unexpected and unwanted behavior in the corrected time series. To tackle this problem, we here propose to eliminate this effect by performing a cascade of corrections that operate on the different timescales present in the system. We have applied the method to both global and station data and shown that it is capable of equalizing the statistics on timescales ranging from hours to years. However, the consequence of such a methodology is that statistical properties obtained from the *control period* time series can only be taken as properties relative to the mean value of that time-interval. For example, the soil-moisture atmosphere interaction (Seneviratne and Stoeckli, 2007; Seneviratne et al., 2006; Fischer et al., 2007) acts on timescales from days to decades and can profoundly impact both on surface temperature and precipitation. As a result, too strong model day-to-day or even inter-annual fluctuations could be reduced by the bias-correction. However, mechanisms functioning on timescales longer than the control period cannot be corrected. One such mechanism that is of crucial importance in most climate change experiments is global warming due to greenhouse gas emissions. It is an open question in climate modeling whether the bias present in current day simulations allows conclusions on the bias in the simulation of future climate change. Such a relationship likely depends on the details of the physical causes behind the bias and whether or not these are relevant to the climate change signal.

To exemplify to what extent a current day temperature bias may affect the climate change signal we resort to the simplest textbook one-dimensional energy balance model incorporating carbon dioxide feedback (Fig. 11). We assume the model misrepresents the planetary albedo α which leads to a bias, both during the current day and in the future. However, the model is taken to represent the mechanism of greenhouse gas induced warming adequately. The energy balance model of the earth system is reduced to the earth surface and a single atmospheric layer that is partially absorbent in the long-wave spectrum – a fraction ϵ is absorbed by the atmosphere – but transparent in the short-wave. Introducing more greenhouse gases will lead to an increased value of absorptivity $\epsilon_{sc} > \epsilon_{con}$, where ϵ_{sc} (ϵ_{con}) refer to the scenario and control period value of absorptivity.

The reader can easily verify that under conditions where ϵ can be changed arbitrarily, a bias correction with an additive
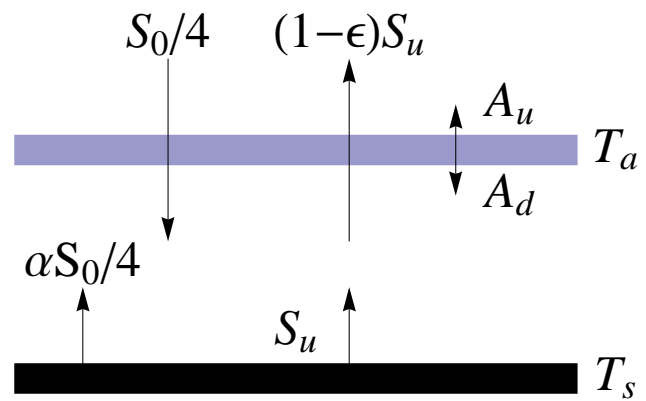
Fig. 11. Simple one-dimensional energy balance model of the earth system with solar constant S_0 and globally integrated incident solar radiation $S_0/4$; planetary albedo α; absorptivity of the atmosphere ϵ. S_u is long-wave radiation emitted by the earth's surface, T_a (T_s) are emission temperatures of the atmosphere (surface); A_u (A_d) are upward (downward) long-wave radiation emitted by the atmosphere.

correction $T_s^{\mathrm{mod,cor}} = T_s^{\mathrm{mod}} + \Delta T$ is not compatible with this demand while a multiplicative correction $T_s^{\mathrm{mod,cor}} = \delta T_s^{\mathrm{mod}}$ is, where δ is a multiplicative factor.

Hence, in the case of the simple energy balance model, applying a bias correction would clearly be beneficial in producing a more realistic climate change signal. For other, more complex models such as a GCM, the conclusion may be less obvious.

The consequences of choosing a certain bias-correction method are much more dramatic in the case of precipitation than in the case of temperature. The model mean often deviates by a factor of two from that of the observations. Furthermore, a simple energy balance model cannot give an indication of which type of correction should be used here (such as multiplicative or additive). As mentioned in the introduction, the reaction of the hydrological cycle to greenhouse gas induced warming – both in observations and in models – would have to be known better, in order to give a definite answer to the question of the adequate correction procedure (Allen and Ingram, 2002; Held and Soden, 2006; Emori and Brown, 2005). However, the use of a multiplicative correction is often triggered by practical considerations, namely the potential effect of an additive correction to produce (unphysical) negative precipitation values which then requires the artificial introduction of dry periods, hence a change in the temporal structure of the time series.

6   Conclusions

Bias correction procedures are emerging as indispensable tools to render output from climate models useful as input to hydrological and impact assessment models. Statistical bias correction schemes are able to transform the entire probability density function of a given modeled climate variable to match that of the observations. Hence, once a choice of timescale is made, for example daily values, the statistics are equalized. The main point of this paper is that bias corrections could potentially benefit from correcting data from different timescales separately, especially when disparate mechanisms act on the different timescales. For instance daily data and monthly mean data can exhibit completely different statistical behavior. To motivate this statement, we have presented data from a bias correction of daily GCM data corrected with observational data. In some regions, the magnitude of interannual fluctuations of monthly means show the opposite sign of discrepancy between model and observations than the day-to-day fluctuations. Therefore, we have proposed an improved method, which we call *cascade bias correction*, which generates a cascade of bias correction functions, each operating on a different timescale. We have used hourly observational and model data to perform a three-tier cascade bias correction for a single gridbox and a two-tier cascade for GCM data for daily and interannual corrections. Our results show that considering timescales separately substantially improves the bias correction, as the actual statistical behavior of the observed data is reproduced at various timescales.

Statistical bias correction does not replace adequate model representation of physical processes. We therefore reiterate the conditions on climate model data to make the application of statistical bias correction schemes reasonable: At every gridbox where SBC is to be applied, it must be ensured that the model provides a realistic representation of the physical processes involved. Quantitative discrepancies between the modeled and observed probability density function of the quantity at hand must be constant in time.

Furthermore, this study emphasizes that every statistical bias correction makes assumptions on the applicability of statistical TFs from today's climate to the future climate. We caution that it is an open question whether a bias correction should impact on the climate change signal produced by a climate model. The climate change mechanism operates on very long timescales and short-term fluctuations should not be used as a means of assessing shortcomings in the model's climate change projection. A bias correction is inherently assuming specific model characteristics that cause the discrepancy between simulated and observational data. A reliable bias correction should hence adequately involve the consequences of greenhouse gas concentration changes and how these impact on the system. We have sketched such an analysis in the case of temperature for a very simple energy balance model. In the case of a full-scale GCM an analogous operation would involve detailed queries into the climate model's representation of the climate change mechanism.

Edited by: F. Pappenberger

References

Allen, M. R. and Ingram, W. J.: Constraints on future changes in climate and the hydrologic cycle, Nature, 419, 224–232, 2002.

Christensen, J. H., Boberg, F., Christensen, O. B., and Lucas-Picher, P.: On the need for bias correction of regional climate change projections of temperature and precipitation, Geophys. Res. Lett., 35, L20709, doi:10.1029/2008GL035694, 2008.

Deque, M.: Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values, Global Planet. Change, 57, 16–26, doi:10.1016/j.gloplacha.2006.11.030, 2007.

Dobler, A. and Ahrens, B.: Precipitation by a regional climate model and bias correction in Europe and South Asia, Meteorol. Z., 17, 499–509, doi:10.1127/0941-2948/2008/0306, 2008.

Emori, S. and Brown, S. J.: Dynamic and thermodynamic changes in mean and extreme precipitation under changed climate, Geophys. Res. Lett., 32, L17706, doi:10.1029/2005GL023272, 2005.

Fischer, E. M., Seneviratne, S. I., Vidale, P. L., Luethi, D., and Schaer, C.: Soil moisture – atmosphere interactions during the 2003 European summer heat wave, J. Climate, 20, 5081–5099, 2007.

Gutowski, W. J., Kozak, K. A., Patton, J. C., Takleet, E. S., Arritt, R. W., and Christensen, J. H.: A Possible Constraint on Regional Precipitation Intensity Changes under Global Warming, J. Hydrometeorol., 8, 1382, 2007.

Haerter, J. O., Berg, P., and Hagemann, S.: Heavy rain intensity distributions on varying timescales and at different temperatures, J. Geophys. Res., 115, D17102, doi:10.1029/2009JD013384, 2010.

Hansen, J., Challinor, A., Ines, A., Wheeler, T., and Moronet, V.: Translating forecasts into agricultural terms: advances and challenges, Clim. Res., 33, 27–41, 2006.

Hay, L. E. and Clark, M. P.: Use of statistically and dynamically downscaled atmospheric model output for hydrologic simulations in three mountainous basins in the western United States, J. Hydrol., 282, 56–75, doi:10.1016/S0022-1694(03)00252-X, 2003.

Held, I. and Soden, B. J.: Robust Responses of the Hydrological Cycle to Global Warming, J. Climate, 19, 5686–5699, 2006.

Hurkmans, R., Terink, W., Uijlenhoet, R., Torfs, P., Jacob, D., and Troch, P. A.: Changes in Streamflow Dynamics in the Rhine Basin under Three High-Resolution Regional Climate Scenarios, J. Climate, 23, 679–699, 2010.

Jacob, D., Goettel, H., Kotlarski, S., Lorenz, P., and Sieck, K.: Klimaauswirkungen und Anpassung in Deutschland, Bericht des Umweltbundesamt (UBA), 2008.

Jungclaus, J. H., Botzet, M., Haak, H., Keenlyside, N., Luo, J.-J., Latif, M., Marotzke, J., Mikolajewicz, U., and Roeckner, E.: Ocean circulation and tropical variability in the coupled model, J. Climate, 19, 3952–3972, 2006.

Leander, R. and Buishand, T. A.: Resampling of regional climate model output for the simulation of extreme river flows, J. Hydrol., 332, 487–496, doi:10.1016/j.jhydrol.2006.08.006, 2007.

Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Themessl, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M., and Thiele-Eich, I.: Precipitation downscaling under climate change: Recent developments to bridge the gab between dynamical models and the end user, Rev. Geophys., 48, RG3003, doi:10.1029/2009RG000314, 2010.

Oki, T. and Kanae, S.: Global hydrological cycles and world water resources, Science, 313, 1068–1072, 2006.

Piani, C., Haerter, J. O., and Coppola, E.: Statistical bias correction for daily precipitation in regional climate models over Europe, Theor. Appl. Climatol., 99, 187, doi:10.1007/S00704-009-0134-9, 2010a.

Piani, C., Weedon, G. P., Best, M., Gomes, S., Viterbo, P., Hagemann, S., and Haerter, J. O.: Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models., J. Hydrol., 395, 199–215, doi:10.1016/j.jhydrol.2010.10.024, 2010b.

Roeckner, E., Bengtsson, L., Feichter, J., Lelieveld, J., and Rodhe, H.: Transient Climate Change Simulations with a Coupled Atmosphere Ocean GCM Including the Tropospheric Sulfur Cycle, J. Climate, 12, 3004–3032, 1999.

Schmidli, J., Frei, C., and Vidale, P.: Downscaling from GCM precipitation: a benchmark for dynamical and statistical downscaling methods., Int. J. Climatol., 26, 679–689, 2006.

Seneviratne, S. I. and Stoeckli, R.: The role of land-atmosphere interactions for climate variability in Europe, Adv. Global Change Res., 33, 179–193, 2007.

Seneviratne, S. I., Luethi, D., Litschi, M., and Schaer, C.: Land-atmosphere coupling and climate change in Europe, Nature, 443, 205–209, 2006.

Sharma, D., Das Gupta, A., and Babel, M. S.: Spatial disaggregation of bias-corrected GCM precipitation for improved hydrologic simulation: Ping River Basin, Thailand, Hydrol. Earth Syst. Sci., 11, 1373–1390, doi:10.5194/hess-11-1373-2007, 2007.

Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., Tignor, M., and Miller, H. L.: Climate Change 2007: The Physical Science Basis. Contr. of WG I to the 4th AR of the IPCC, Cambridge University Press, Cambridge, UK and New York, NY, USA, 996 pp., 2007.

Vorosmarty, C. J., Green, P., Salisbury, J., and Lammers, R. B.: Global water resources: vulnerability from climate change and population growth, Science, 289, 547, 2000.

Weedon, G. P., Gomes, S., Viterbo, P., Shuttleworthe, W. J., Blyth, E., Österle, H., Adam, J. C., Bellouin, N., Boucher, O., and Best, M.: The WATCH forcing data 1958–2001: a meteorological forcing dataset land-surface and hydrological-models, early online release, http://journals.ametsoc.org/toc/hydr/0/0, last access: March 2011, J. Hydrometeorol., in press, 2011.

Widmann, M., Bretherton, C. S., and Salathé Jr., E. P.: Statistical precipitation downscaling over the Northwestern United States using numerically simulated precipitation as predictor, J. Climate, 16, 799–816, 2003.

Wilson, P. S. and Toumi, R.: A fundamental probability distribution for heavy rainfall, Geophys. Res. Lett., 32, L14812, doi:10.1029/2005GL022465, 2005.

Wood, A. W., Maurer, E. P., Kumar, A., and Lettenmaier, D. P.: Long-range experimental hydrologic forecasting for the eastern United States, J. Geophys. Res., 107, 4429, doi:10.1029/2001JD000659, 2002.