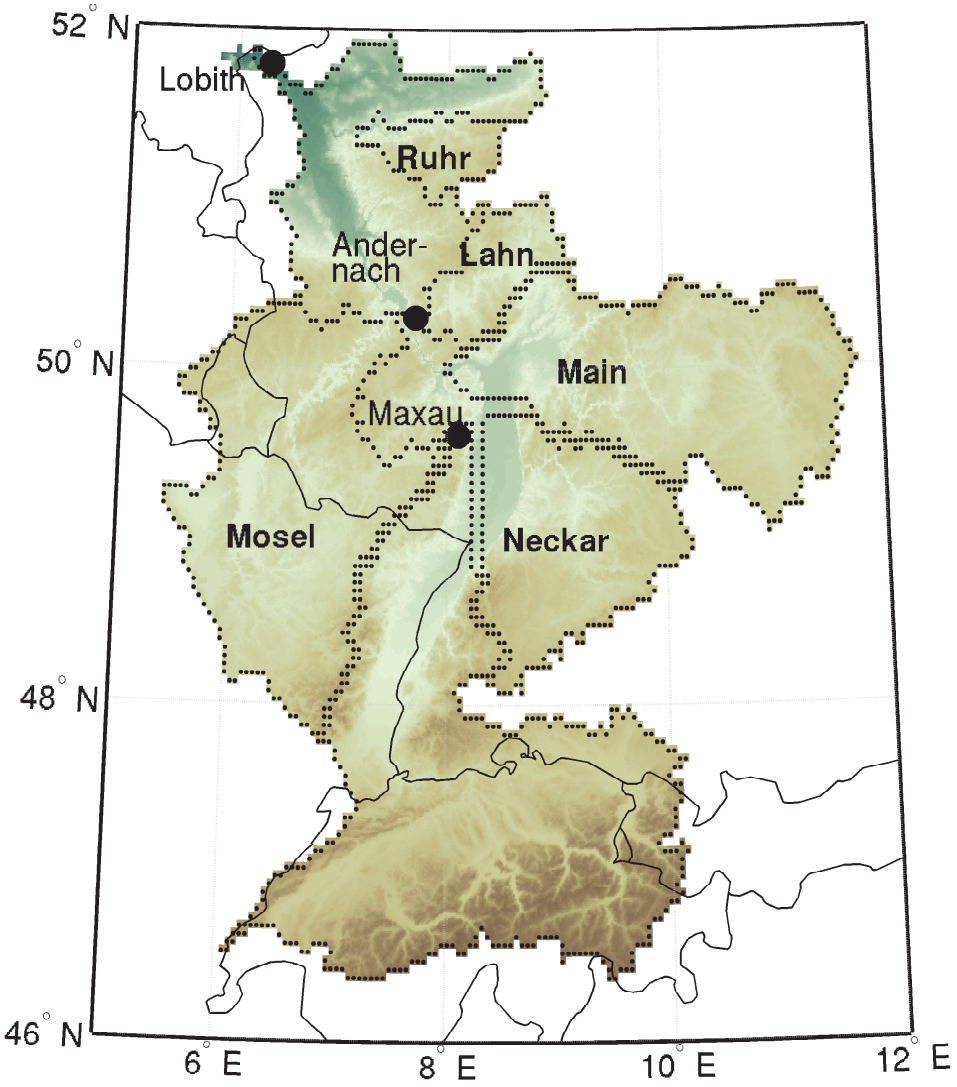
# Bias correction of temperature and precipitation data for regional climate model application to the Rhine basin

W. Terink*, R.T.W.L. Hurkmans*, R. Uijlenhoet*, P.M.M. Warmerdam*, P.J.J.F. Torfs*

October 27, 2008




*Hydrology and Quantitative Water Management Group, Wageningen University

**Abstract**

Nowadays hydrological models have become an important tool in predicting streamflow generation. Most of these models need to be calibrated with the correct meteorological input before they can predict reliable streamfow generation. This report focuses on the hydrological analysis of three different meteorological forcing data sets for the Rhine basin. These are observed data, downscaled ERA15 data and a regional climate model run, known as the reference scenario. The goal of this report is to analyse the difference between the last two and the observed data such that a bias correction can be applied to minimize these differences. The bias correction used here corrects for the mean and the coefficient of variation of the precipitation data. The temperature data is corrected for the mean and the standard deviation.

We found that the downscaled ERA15 data are too warm and wet for the largest part of the Rhine basin when compared with the observations. This is also the case for the reference scenario. The results of the applied bias correction were quite satisfactory. The spatial differences between the corrected data sets and the observations decreased enormously. Even in the Alps, where these differences were large before applying the bias correction, these differences decreased considerably. When analyzing the monthly average precipitation sums we found that the downscaled ERA15 data are especially too wet in the summer months. The reference scenario is too wet for the entire year. The correction does not have a positive effect in September for the ERA15 data. The reason for this might be that the observations show an increase in precipitation from August to September, while the ERA15 precipitation decreases during this period. The same is true in March for the reference scenario. The comparison between the observed and corrected average monthly temperature for both data sets is quite satisfactory. The corrected average monthly temperature is almost similar to that of the observations.

Besides the fact that the bias correction only corrects for the mean, the standard deviation and the coefficient of variation, it seems to correct other important statistics as well for both data sets. This is especially the case for the fraction of wet days, the lag-1 autocorrelation and exceedance probabilities. The correlation coefficients between the observed and corrected data sets have increased reasonably regarding the fraction of wet days and the lag-1 autocorrelations. The exceedance probabilities of the corrected 10-day winter and summer precipitation sums match those of the observations very well. Even for small exceedance probabilities, the corrected data sets match the observations better than the uncorrected data sets.

# Contents

# 1 Introduction

To investigate the impact of climate change on river streamflow, hydrological models have become an important tool. For a correct understanding of future discharges, it is important to have a hydrological model which is calibrated such that the simulated discharge matches the observed discharge as good as possible. A valuable calibration is only possible when both the observed discharge and climate input (*precipitation, temperature, ...*) contain few errors.

The hydrological model used in this report, known as VIC (*Variable Infiltration Capacity*), is a spatially distributed model. This means that the river catchment is divided into several grid cells, and for each grid cell a calculation of baseflow and runoff is performed. Considering a division into grid cells with a resolution of 0.05 x 0.05 degrees, there are no measurements available at this resolution for the considered period. However, the ECMWF (*European Centre for Medium-Range Weather Forecasts*) has provided a re-analysis data set known as ERA15. The forcing of the re-analyses, in addition to observations, came from the ocean surface temperature (SST) analyses and the ocean ice cover analysis. The ERA15 data set was downscaled at MPI (*Max Planck Institute, Hamburg*) with the RCM (*Regional Climate Model*) REMO. The resulting downscaled ERA15 data set, ERA15d, covers the entire Rhine basin on a resolution of 0.088 x 0.088 degrees (approximately 9 x 9 km). *Hurkmans et al.* (2008) compared the ERA15d data for the period 1993-1995 with observations from the International Commission for the Hydrology of the Rhine basin. This analysis showed that the ERA15d data is a little bit too warm and wet when compared with the (CHR) observations. Using such data for model calibration will automatically result in poor model performance. Therefore this data set needs to be corrected in such a way, that the modeled temperature and precipitation data match the observed temperature and precipitation data regarding their most important statistics.

The goal of this report is to analyse the observed CHR- and ERA15d data set, and apply a correction to both the temperature and precipitation of the latter, such that the most important statistics (mean, std and CV) of the ERA15d data match those of the observations better. With the corrected ERA15d data set, the VIC-model will be re-calibrated.

A second bias correction will be appplied to a data set which is available on the MPI website. This is the "Climate of the $20^{th}$ century" run, which concerns the period 1951-2000. This data set will be used in a later phase to compare the future climate scenarios with. Hereafter this data set will be referred to as the reference scenario. The bias correction, retrieved from the reference scenario and the observations, then will finally be applied to the three IPCC (*Intergovernmental Panel on Climate Change*) future climate scenarios B1, A1B and A2, to be able to investigate whether or not climate change will affect the River Rhine streamflow behaviour. These three scenarios are generated with the same RCM (REMO) as the reference scenario. It is assumed that the bias between the REMO model output and the observations remains invariant when climate changes.

# 2 Observations

## 2.1 Introduction

The observed data set contains daily values of precipitation and temperature data for 134 subbasins and is made available by the Bundesanstalt für Gewässerkunde (BfG) (*Beersma*, 2002). This data set will be referred to as the CHR data set hereafter. These observations span the period 1961-1995. The bias correction needs to be applied to 2 data sets; the ERA15d data set and the reference scenario. The ERA15d data set and the CHR data set overlap during the period 1979-1995. The reference scenario covers the entire CHR observed period 1961-1995. For this reason, the CHR data set is analyzed both for the periods 1979-1995 and 1961-1995.

## 2.2 Precipitation

For the analysis of the observed precipitation, the Rhine catchment has been divided into 134 subbasins. For each of these subbasins, an average daily precipitation amount has been calculated over the periods 1961-1995 and 1979-1995. Both temperature and precipitation exhibit seasonality in most European countries. Therefore the analysis is performed for the summer months (AMJJAS) and winter months (ONDJFM) separately.

The average winter precipitation sum for each subbasin is shown in Figure 1 for the period 1961-1995. From this figure it can be seen that the Alps and the Black Forest in Germany receive the largest precipitation amounts during winter. These areas catch precipitation amounts up to 900 mm. However, the largest part of the Rhine basin receives an average precipitation amount between 300 and 600 mm. The area-weighted average precipitation sum is 467 mm for the entire Rhine basin, with an area-weighted standard deviation of 129 mm.

A similar analysis is done for the summer period. This result is shown in Figure 1 as well. Again a similar pattern of high precipitation amounts can be seen in the Alps and the southern part of Germany. Some areas in the Alps receive precipitation amounts up to 1200 mm during summer. When comparing the average precipitation sum in winter with that during summer, it can be concluded that summer rainfall is larger than that in winter. The difference between summer and winter precipitation is small in the largest part of Germany, when compared to some areas in the Alps where these differences can reach values of 400 mm. During summer, the area-weighted average precipitation sum is 524 mm. The corresponding standard deviation for summer is 141 mm, which means that the spread in summer is larger than in winter.

The same precipitation analysis is performed for the period 1979-1995. The average winter precipitation sum for each subbasin is shown in Figure 2. Again the Alps and the Black Forest in Germany are the areas with the highest average precipitation amounts. The area-weighted average precipitation amount for this period seems to be higher than for the period 1961-1995. This leads to the conclusion that at least one, or even more than one winter period, is drier than average during the period 1961-1978. The average summer precipitation sum for each subbasin is shown in Figure 2. When compared to the average winter precipitation sums, summer is receiving more precipitation. The Alps receive the highest amounts of precipitation. A maximum average precipitation sum of 1150 mm is reached in this area, compared to an amount of 350-600 mm in the largest part of Germany. It turns out that the period 1961-1978 is characterized by one or more drier winters. This is also true for the summer season. However, these differences are smaller than for the winter season. The difference in the area-weighted average precipitation sum for the winter season between the periods 1979-1995 and 1961-1995 is 32 mm, instead of the 4 mm difference for the summer season.

## 2.3 Temperature

A map of the average winter temperature for each subbasin is shown in Figure 3 for the period 1961-1995. The variation in average temperature over the entire basin is large. The lowest temperature is measured in the Alps, due to the high altitudes in this mountainous area. The lowest average temperature in this

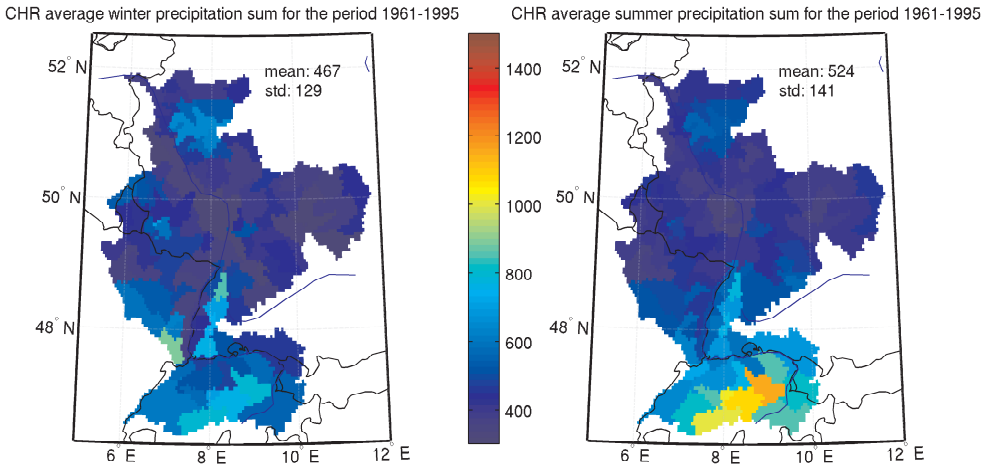



Figure 1: Average winter and summer precipitation sum [mm] of the CHR observed data for the period 1961-1995. The area-weighted mean and standard deviation are shown as well.

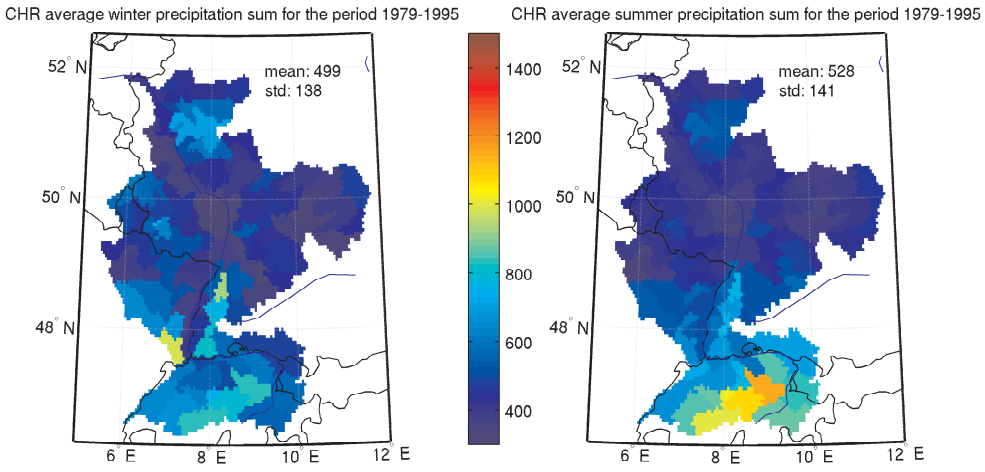



Figure 2: Average winter and summer precipitation sum [mm] of the CHR observed data for the period 1979-1995. The area-weighted mean and standard deviation are shown as well.

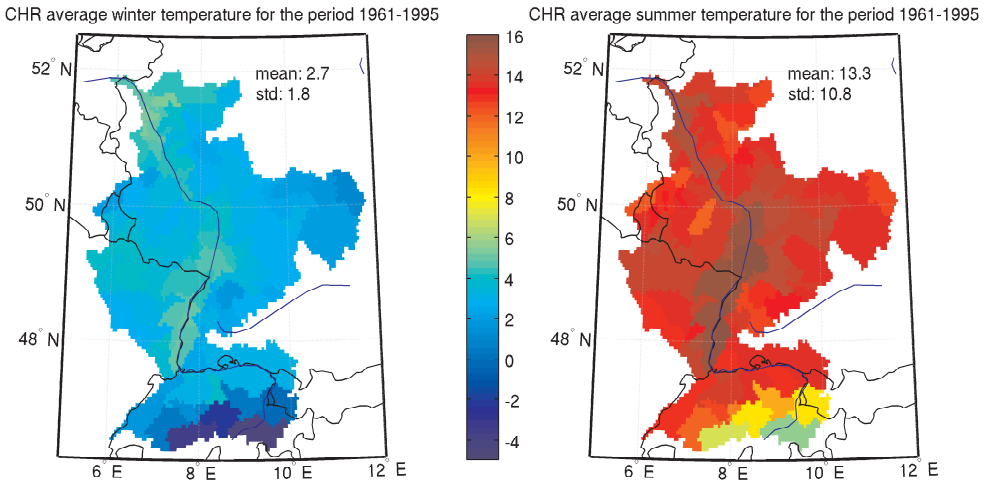



Figure 3: Average winter and summer temperature [°C] of the CHR observed data for the period 1961-1995. The area-weighted mean and standard deviation are shown as well.

area is -4 to -5 °C. Warmer temperatures are measured throughout the largest part of Germany, especially in the area close to the river Rhine. In these areas, the average winter temperature can reach values up to 5-6 °C. An area-weighted average winter temperature of 2.7 °C is calculated, with an area-weighted standard deviation of 1.80 °C. A similar map for the summer season is shown in Figure 3. Again the lowest temperature is measured in the Alps, and the highest is measured near the river Rhine. The maximum average temperature is about 16 °C and is measured in central Germany. The lowest average temperature (6 °C) is measured in the Alps near the Italian border. The area-weighted average temperature for the summer season is 13.3 °C, with a standard deviation of 10.8 °C.

The average winter temperature for each subbasin over the period 1979-1995 is shown in Figure 4. The area-weighted average temperature for this period is 2.7 °C, with an area-weighted standard deviation of 1.8 °C. This means that this period is slightly warmer than the period 1961-1995. The variation in temperature over the Rhine basin for 1979-1995 is comparable to the latter period. Again the lowest temperatures are found in the Alpine areas, whereas the higher temperatures can be found in the area near the Rhine. The map with the average summer temperature for the period 1979-1995 is shown in Figure 4. When comparing the summers of the two selected periods, there are hardly any differences. The area-weighted average summer temperature for the period 1979-1995 is slightly higher than for the period 1961-1995, which suggests that the period before 1979 has at least one, or even more colder years in it. The maximum average temperature is about 16 °C and is measured in central Germany. The lowest average temperature is measured in the Alps near the Italian border, and has a value of 6 °C.

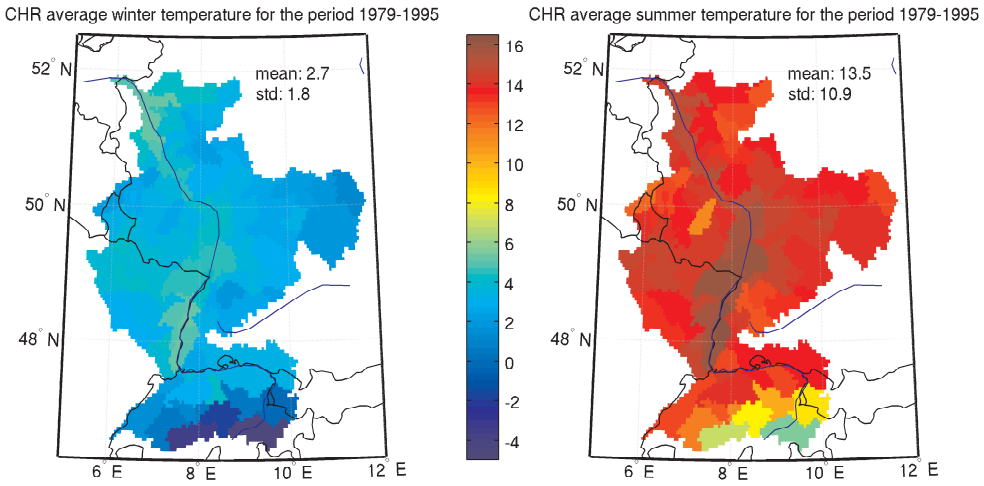



Figure 4: Average winter and summer temperature [°C] of the CHR observed data for the period 1979-1995. The area-weighted mean and standard deviation are shown as well.

# 3 Bias correction

## 3.1 Objective

As mentioned before in Chapter 1, the goal of this report is to analyse the CHR-, ERA15d- and reference scenario data set, and apply a bias correction to the last two, such that the most important statistics (mean, std and CV) of the bias-corrected data sets match the most important statistics of the observed CHR data sets well. The bias correction applied in this report, is based on that employed by *Leander and Buishand* (2007) for the Meuse basin. They found that a relatively simple nonlinear correction, adjusting both the biases in the mean and variability leads to a better reproduction of observed extreme daily and multi-day precipitation amounts than the commonly used linear scaling correction. This method of bias correction does not correct for the fraction of wet and dry days. The bias correction of temperature was found to be more straigthforward than that of precipitation, involving shifting and scaling to adjust the mean and variance, respectively. In this chapter, the method used to calculate the bias correction for precipitation and temperature will be described in detail.

The bias correction needs to be applied to the ERA15d data set, which will be used for calibration of the VIC model, and to the reference scenario. As mentioned before, the ERA15d data set has the overlapping period 1979-1995 with the CHR observed data set. The overlapping period for the reference scenario is 1961-1995.

## 3.2 Method

### 3.2.1 Precipitation

Because the bias in precipitation and temperature was found to vary spatially, bias corrections were carried out for each of the 134 subbasins individually. *Leander and Buishand* (2007) used a power transformation, which corrects the CV (Coefficient of Variation) as well as the mean. In this nonlinear correction each daily precipitation amount $P$ is transformed to a corrected $P^*$ using:

$$P^* = aP^b \tag{1}$$

The effect of sampling variability is reduced by determing the parameters $a$ and $b$ for every five-day period of the year, including data from all years available, in a window including 30 days before and after the considered five-day period (*Leander and Buishand*, 2007). The determination of the $b$ parameter is done iteratively. It was determined such that the CV of the corrected daily precipitation matches the CV of

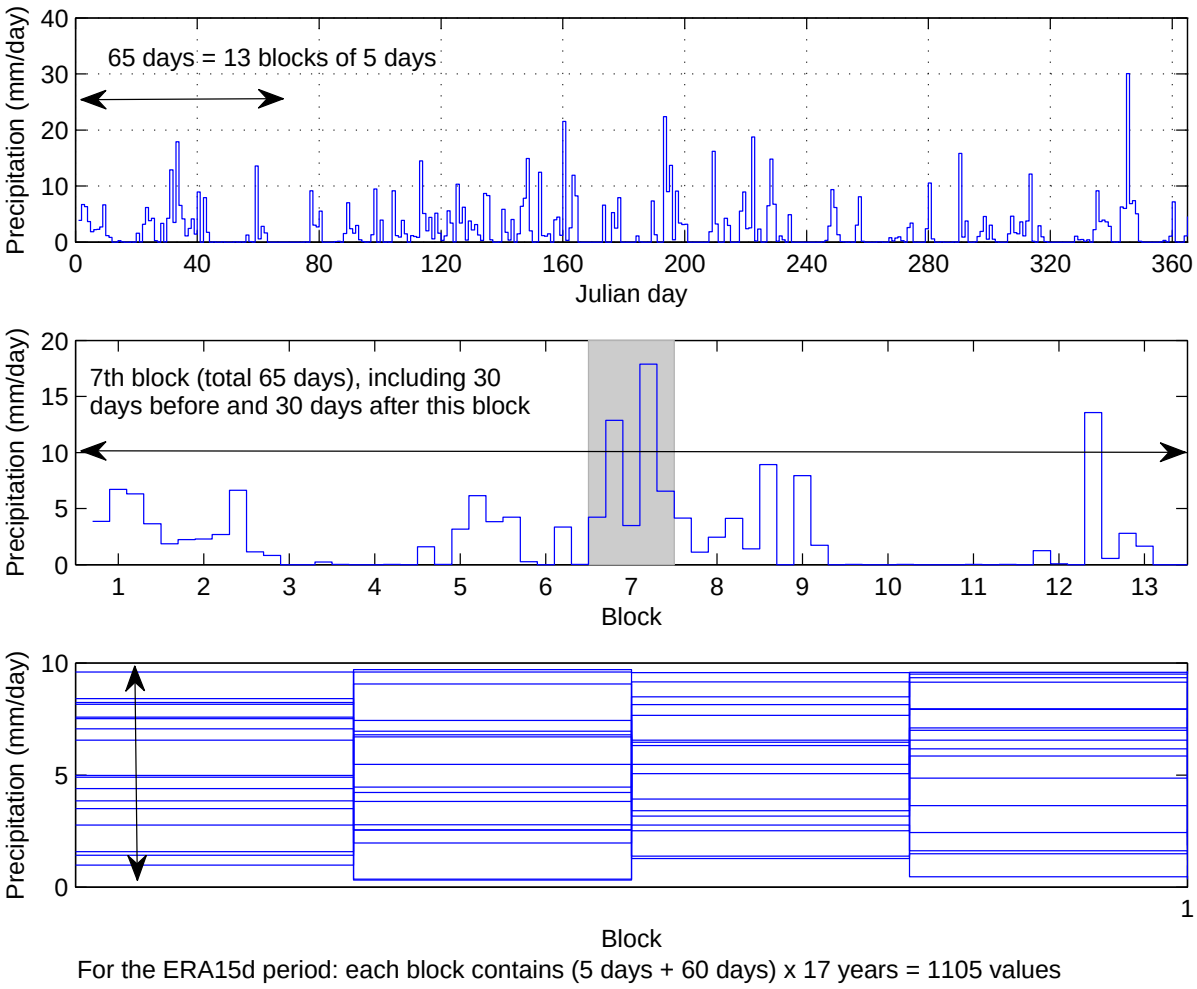



Figure 5: Dividing a year into 73 blocks of 5 days for the calculation of the parameters $a$ and $b$.

the observed daily precipitation. With the determined parameter $b$, the transformed daily precipitation values are calculated using:

$$P^* = P^b \tag{2}$$

Then the parameter $a$ is determined such that the mean of the transformed daily values (*Equation 2*) correspond with the observed mean. The resulting parameter $a$ depends on $b$. The parameter $b$ depends only on the CV and is independent of the value of parameter $a$. At the end each block of 5 days has its own $a$ and $b$ parameter, which are the same for each year.

The bias correction for the ERA15d data set needs to be calculated for the period 1979-1995, which has a total length of 17 years. Figure 5 illustrates the division of 365 days into 73 blocks of 5 days. For every 5-day block, a different set of parameters $a$ and $b$ is determined using the method described above. The figure at the top of 5 represents the daily precipitation throughout the year. The figure in the middle of 5 zooms in to the first 65 days of the year. This results in 13 blocks of 5 days each. The bottom line figure in 5 zooms in into the first block of the year. Referring to the ERA15d period, this period contains 17 years of data. Taken into account the 30 days before and after the considered 5-day period, each block contains 65 x 17 = 1105 values for the calculation of the CV and the mean. For the reference scenario (1961-1995), each block contains 65 x 35 = 2275 values.

### 3.2.2 Temperature

*Leander and Buishand* (2007) applied a different technique for correcting the daily temperature. The correction of temperature only involves shifting and scaling to adjust the mean and variance. For each subbasin, the corrected daily temperature $T^*$ was obtained as:

$$T^* = \overline{T}_{obs} + \frac{\sigma(T_{obs})}{\sigma(T_{cal})}(T_{uncor} - \overline{T}_{obs}) + (\overline{T}_{obs} - \overline{T}_{cal}) \tag{3}$$

where $T_{uncor}$ is the uncorrected daily temperature from ERA15d or the reference scenario, $T_{obs}$ is the observed daily average temperature from the CHR data set, $T_{cal}$ is the corresponding basin average temperature, obtained from either the ERA15d data or the reference scenario. In this equation an overbar denotes the average over the considered period and $\sigma$ the standard deviation. This method was not appropriate for precipitation because it might cause negative values. The considered period is 17 years for the ERA15d data set and 35 years for the reference scenario. Again both statistics were determined for each 5-day block of the year separately, using the same 65-day windows as for the bias correction of daily precipitation.

# 4 Results bias correction

## 4.1 ERA15d data set

### 4.1.1 Precipitation

The method for calculation of the bias correction has been described in detail in Section 3.2. For each block of 5 days and for each subbasin a parameter $a$ and $b$ have been determined. The determined parameter $a$ is shown in Figure 6. The n$^{th}$ day of the year is plotted on the x-axis and the corresponding subbasins are plotted on the y-axis. The value of parameter $a$ is plotted on the z-axis. After correcting for the CV (determing parameter $b$), it can be concluded from Figure 6 that the parameter $a$ shows a seasonal pattern. The parameter value is higher in winter than in summer. This suggests that the observed averages in winter are higher than the ERA15d averages. In summer the parameter value is often smaller than one, which suggests that the ERA15d averages are too wet in the summer period. The area-weighted average of parameter $a$ is 0.78. This means that the ERA15d data set, when averaged over all subbasins and 5-day blocks, is too wet in comparison with the CHR observations.
The variation of parameter $b$ throughout the year and over the subbasins is shown in Figure 7. In this figure the seasonal pattern is less clear than the pattern of $a$ in Figure 6. It seems that the $b$ parameter is larger than one most of the time, especially in summer. Whether or not a value larger than one leads to an increased or decreased precipitation value, depends on the uncorrected daily precipitation amount. An uncorrected precipitation amount smaller than one results in a smaller value in combination with a parameter $b$ larger than one. A precipitation amount larger than one results in a larger value if $b > 1$.
A graph of the area-weighted average parameters $a$ and $b$ is shown in Figure 8. In both the left and the right figure, the area-weighted average and standard deviation are plotted. The shaded area in both plots represents the fluctuation of the area-weighted standard deviation around the area-weighted average. The area-weighted standard deviation of parameter $a$ is quite large when compared to that of parameter $b$. The area-weighted standard deviation of parameter $a$ is largest in winter. The area-weighted standard deviation of parameter $b$ is quite constant over time. From this figure it is clear that the effect of parameter $a$ is to adjust the ERA15d data set from wet to dry most of the time. The largest bias correction for both parameters occurs in the summer months. The last three months of the year have the smallest bias correction.
Figure 9 represents the ratio of the area-weighted average corrected precipitation, over the area-weighted averaged uncorrected precipitation. It can be seen that the ratio is close to one over almost the entire period and that the precipitation is corrected to a drier precipitation value most of the time. From day 230 until day 325 the correction is the other way around, from dry to wet.
In Figure 10 the most important statistics of the corrected and uncorrected ERA15d precipitation data are plotted versus the observed CHR precipitation data in five scatter plots. These statistics have been calculated over the entire period 1979-1995 and for each subbasin separately. This results in 134 data

points in each subplot. As mentioned before, the applied bias correction method corrects for the CV and the mean. The result of this is clearly visible in the plots of the mean, standard deviation and CV. The correlation coefficients of the bias-corrected ERA15d data are close to one for these three statistics. Although the method of bias correction was not developed for correcting the fraction of wet days and lag-1 autocorrelation, it seems that this method of bias correction has a positive effect on these statistics as well. Both correlation coefficients have increased significantly.

For each subbasin the average daily precipitation over the period 1979-1995 has been calculated. This is done for the CHR observed data and the uncorrected and corrected ERA15d data. The difference in daily precipitation between the CHR observations and the uncorrected ERA15d data is shown in Figure 11 (left panel). The difference between the two data sets varies between -2 and 2 mm/day. A positive value means that the ERA15d precipitation is wetter than the CHR precipitation. Especially in some subbasins in the Alps the ERA15d precipitation is larger than the CHR precipitation. The opposite is true for the Black Forest and some subbasins located in the southern part of Germany. Here the observations tend to be much wetter than the ERA15d data. The difference between the corrected ERA15d daily precipitation values and the CHR observations is shown in Figure 11 (right panel). The differences between the bias-corrected ERA15d data set and the observations are very small.

The next analysis is based on the monthly precipitation sums, averaged (area-weighted) over the period 1979-1995. Figure 12 represents the variation of the 17-year average monthly precipitation sum throughout the year. In this figure the CHR observations, the uncorrected and corrected ERA15d data are plotted. It is clear that the bias-corrected ERA15d data matches the observed CHR data better than the uncorrected ERA15d data. The largest bias correction is applied in the months May to November, which was already pointed out in Figure 8. The bias correction in September and October turns out to be the worst. The CHR observations show an increase in precipitation during the shift from August to September. The opposite is true for the ERA15d data. It turns out to be difficult to correct for these different shifts in precipitation. However, the bias-corrected ERA15d data performs better than the uncorrected ERA15d data.
A histogram has been made of the area-weighted 10-day precipitation sums over the entire period 1979-1995. This histogram is presented in Figure 13. The patterns of the uncorrected and corrected ERA15d data are quite similar. The classes between 10 and 50 mm are the classes in which most of the 10-day precipitation events occur. It seems that the pattern of the uncorrected ERA15d data matches the CHR observations better than the corrected ERA15d data does. This is explained by the fact that the ERA15d data needs to be corrected from wet to dry most of the time. Therefore the higher precipitation sums are shifted to the lower bin counters which results in increased frequencies in these bins.
The exceedance probability has been investigated for the 10-day winter precipitation sums. This is illustrated in Figure 14 for the CHR data and the uncorrected and corrected ERA15d data. A Generalized Pareto distribution is fitted through the data. The probability density function for the Generalized Pareto distribution is given by Equation 4. The parameters are calculated using the maximum likelihood estimation. The 10-day precipitation sums are area-weighted averages. Both the uncorrected and corrected ERA15d data match the CHR data well for exceedance probabilities $> 1/20$. However, the corrected ERA15d data performs better when compared with the observed CHR data for exceendance probabilities $> 1/20$. For exceedance probabilities $< 1/20$, the uncorrected ERA15d data is performing better.
The exceedance probability for 10-day summer precipitation sums has been investigated as well. These results are shown in Figure 15 and are quite satisfactory. A Generalized Pareto distribution is fitted through the data. The 10-day precipitation sums are area-weighted averages. For the summer season the corrected ERA15d data performs better for almost all exceendance probabilities. This is something one would not expect, because the bias correcton is not determined for 10-day precipitation sums, but for single day events.

$$y = f(x|k, \sigma, \theta) = \left(\frac{1}{\sigma}\right)\left(1 + k\frac{(x-\theta)}{\sigma}\right)^{-1-\frac{1}{k}} \tag{4}$$

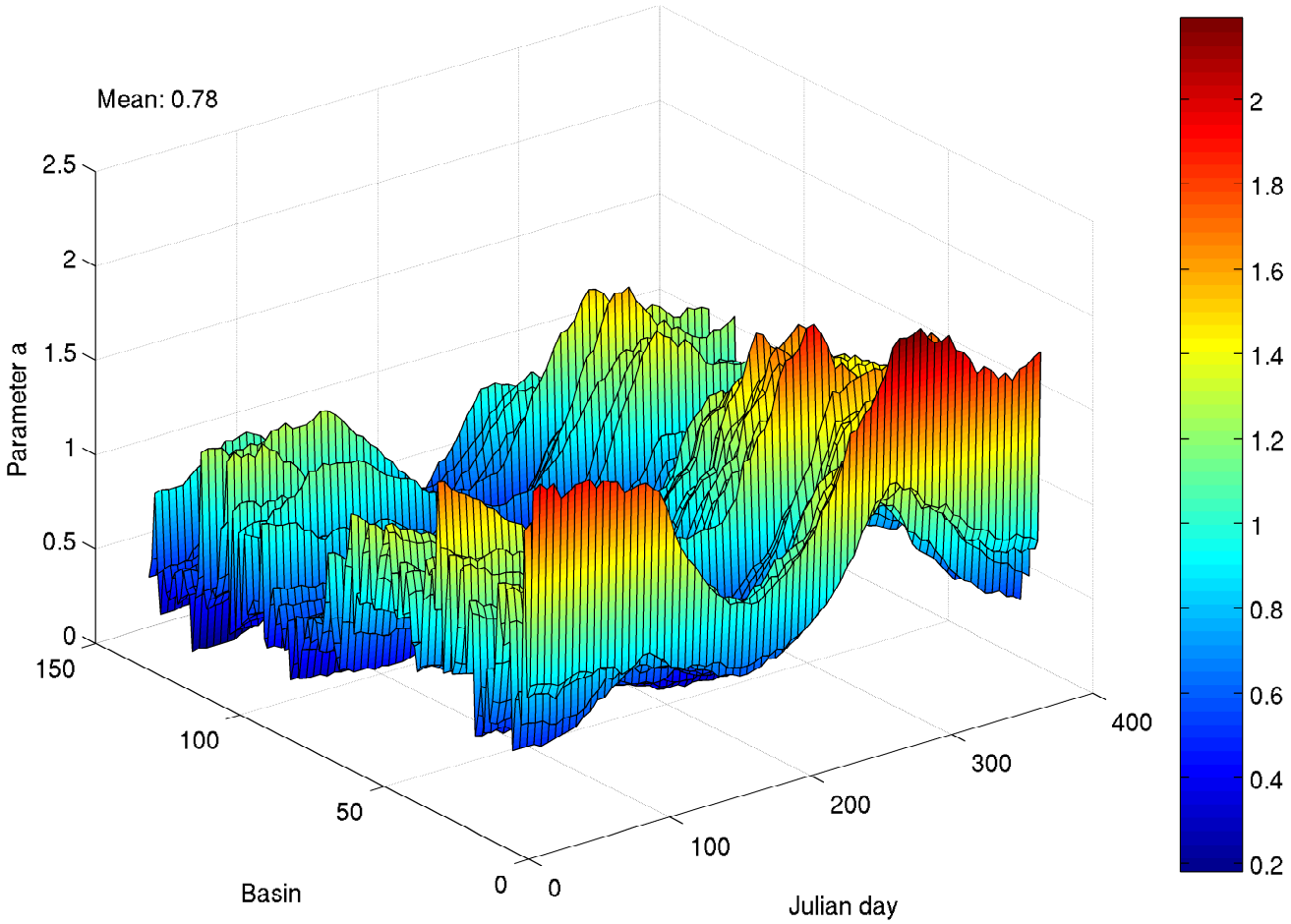



Figure 6: Derived parameter $a$ for each 5-day period and each subbasin.

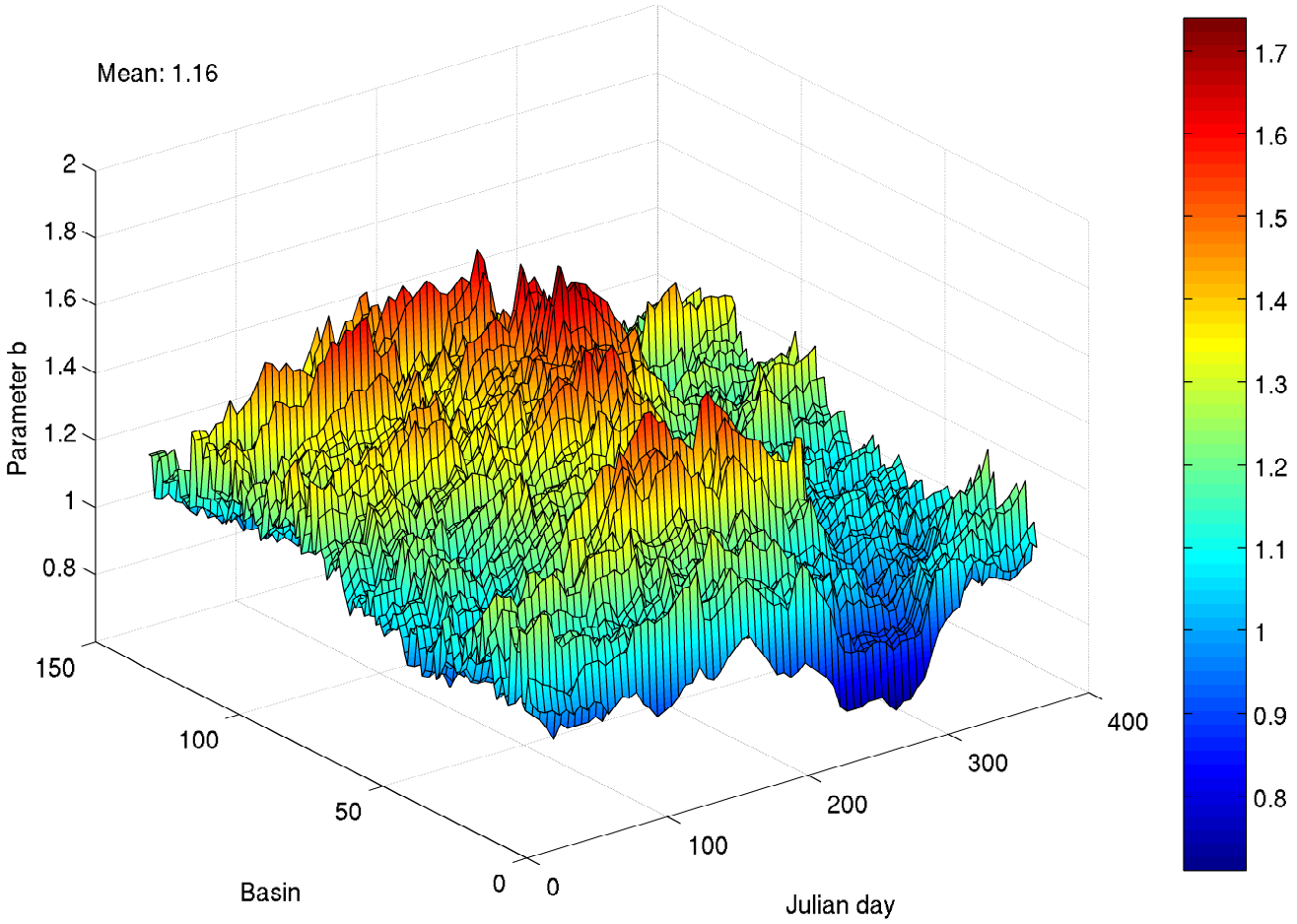



Figure 7: Derived parameter $b$ for each 5-day period and each subbasin.

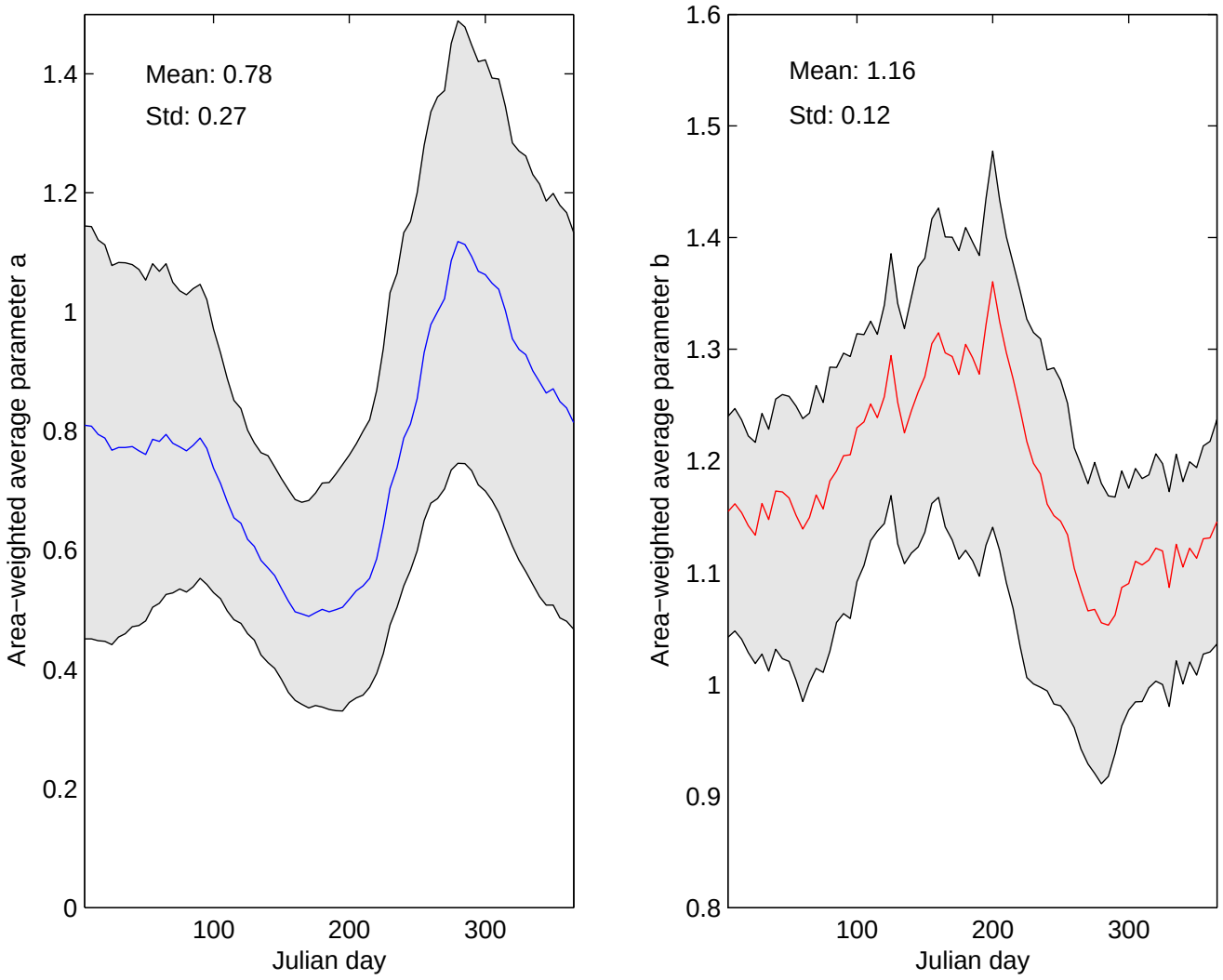



Figure 8: Area-weighted average parameters $a$ and $b$ for each 5-day period.

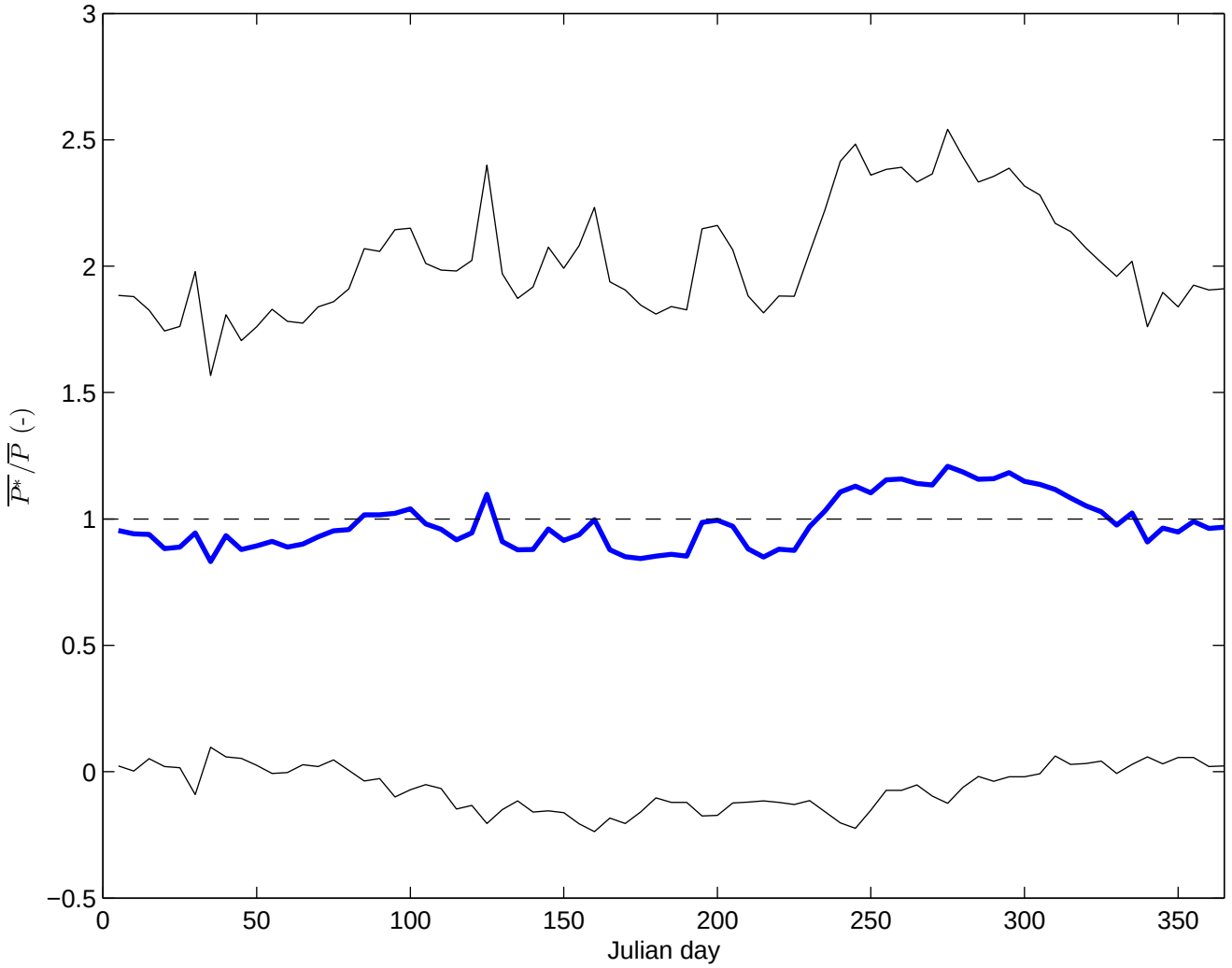



Figure 9: Ratio of area-weighted average corrected precipitation and the area-weighted average uncorrected precipitation (bold line). The thin black lines represent the mean plus or minus one standard deviation. The dashed black line is the ratio equal to one.

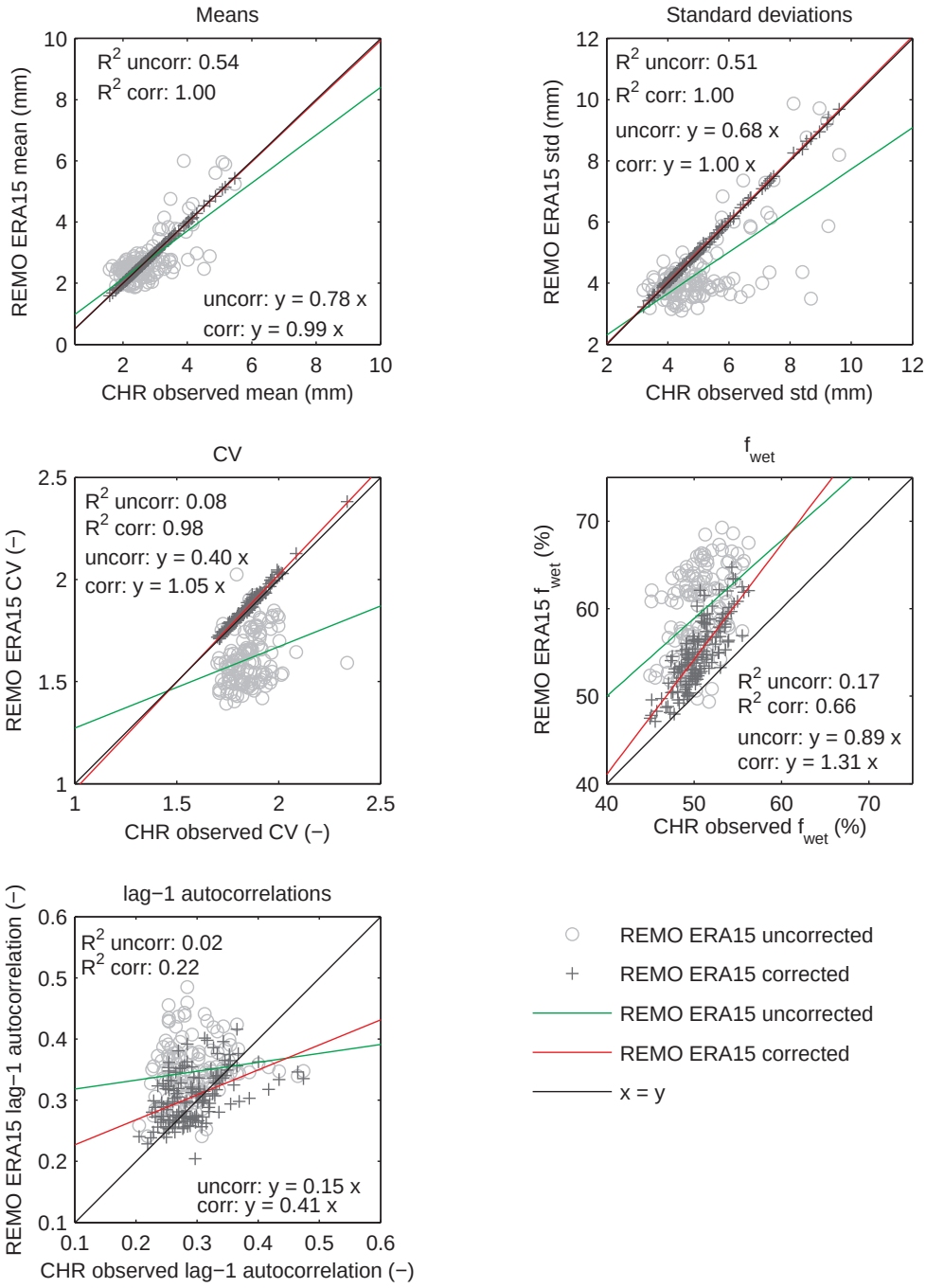



Figure 10: Scatter plots of the statistics of the observed CHR precipitation data versus the corrected and uncorrected ERA15d precipitation data. The statistics are calculated for each subbasin over the entire period 1979-1995. The fraction of wet days ($f_{wet}$) is the percentage of days where P>0.3 mm. In each subplot the correlation coefficient ($R^2$) and slope of the linear regression line are plotted.

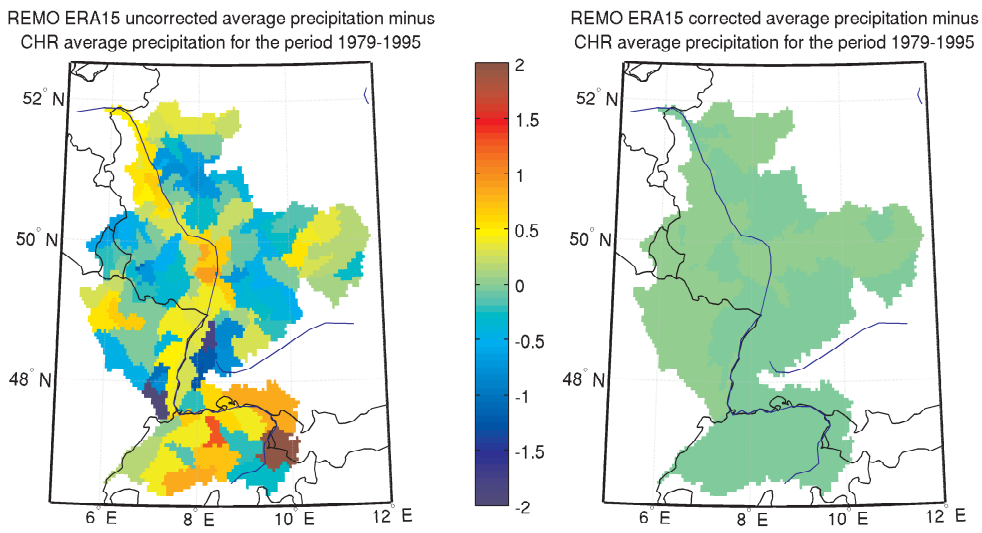



Figure 11: Difference in average daily precipitation between the CHR data and the uncorrected and corrected ERA15d data for the period 1979-1995.

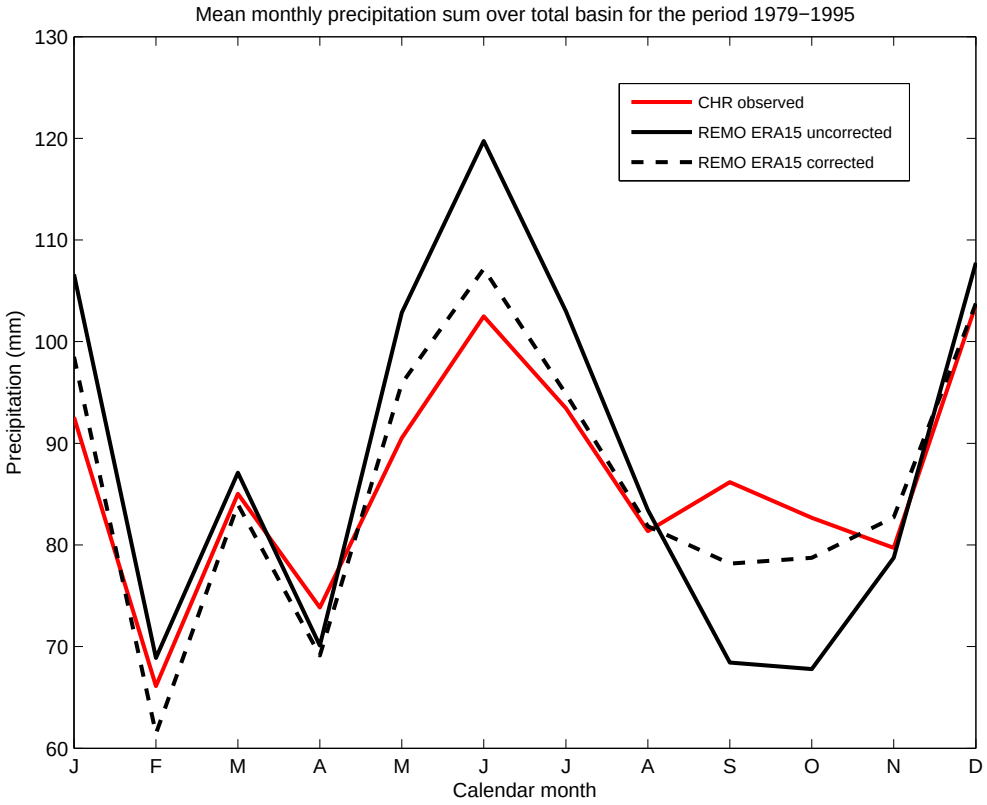



Figure 12: Area weighted monthly precipitation sum over the entire basin for the period 1979-1995.

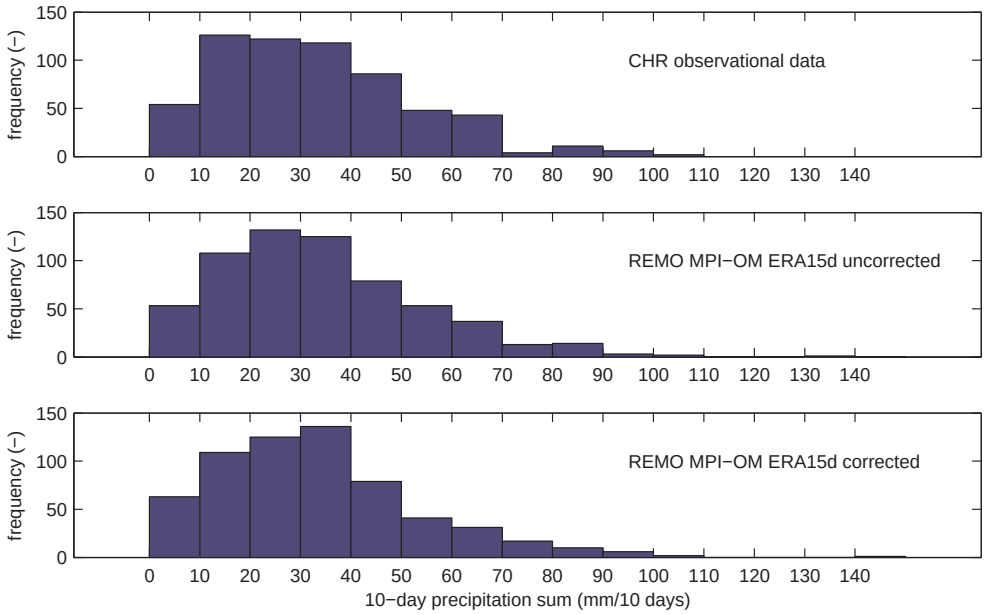



Figure 13: Histogram of the area-weighted 10-day precipitation sums for the period 1979-1995.

### 4.1.2 Temperature

The method for calculating the bias correction is described in detail in Section 3.3. The uncorrected daily temperature is corrected with the use of equation 3. The temperature in this equation is adjusted for the standard deviation and the mean. The ratio between $\sigma(T_{obs})$ and $\sigma(T_{cal})$ is shown in Figure 16 for each subbasin and 5-day block. Again a seasonal pattern can be seen. From January on, there is a downward trend until the start of summer. When the ratio between $\sigma(T_{obs})$ and $\sigma(T_{cal})$ is decreasing the $\sigma(T_{cal})$ is larger. During summer this ratio is approaching 1, and around mid-summer it is decreasing again. This means that $\sigma(T_{cal})$ is approaching $\sigma(T_{obs})$ in the first period of summer and is increasing again in the second summer period. The area-weighted average ratio of 0.96 suggests that the average spread in temperature for ERA15d is larger than the average spread in temperature for the observations. The difference between $\overline{T}_{cal}$ and $\overline{T}_{obs}$ is shown in Figure 17 for each subbasin and 5-day block. In this figure even a more pronounced seasonal pattern can be distinguished. Especially during summer the difference between $\overline{T}_{cal}$ and $\overline{T}_{obs}$ is larger. This difference tends to be negative, which means that the 17-year average of ERA15d is warmer than the 17-year average of the observations. There also seem to be two subbasins where the ERA15d 17-year average temperature in winter is lower than the observed temperature. These subbasins are located in the Alps, as is presented in Figure 18. Until now it is not clear why these areas are colder than the observed data set. These areas are located in mountainous regions with high altitudes. The ECHAM5- or REMO model may have some difficulties in predicting the correct temperature in these areas. These two basins are the same basins in which the ERA15d data is far too wet when compared with the observations (as we saw before in Figure 11). It might be the case that the ECHAM5 and/or REMO model simulates rainfall in these regions, which in fact has to be snow due to the high altitudes.
In Figure 19 the most important statistics of the corrected and uncorrected ERA15d temperature data are plotted versus the observed CHR temperature data in 4 scatter plots. These statistics are calculated over the entire period 1979-1995 and for each subbasin separately. This results in 134 data points in each subplot. As mentioned before, the chosen method of bias correction corrects for the mean and the standard deviation. This is clearly visible in the plots of the mean, standard deviation and CV. The correlation coefficients of the bias-corrected ERA15d data are close to 1 or even 1 for these statistics. Although it was not the intention to correct for the lag-1 autocorrelation, it seems that the applied bias correction has a positive effect on this as well. The correlation coefficient has increased from 0.17 to 0.44 for this statistic.

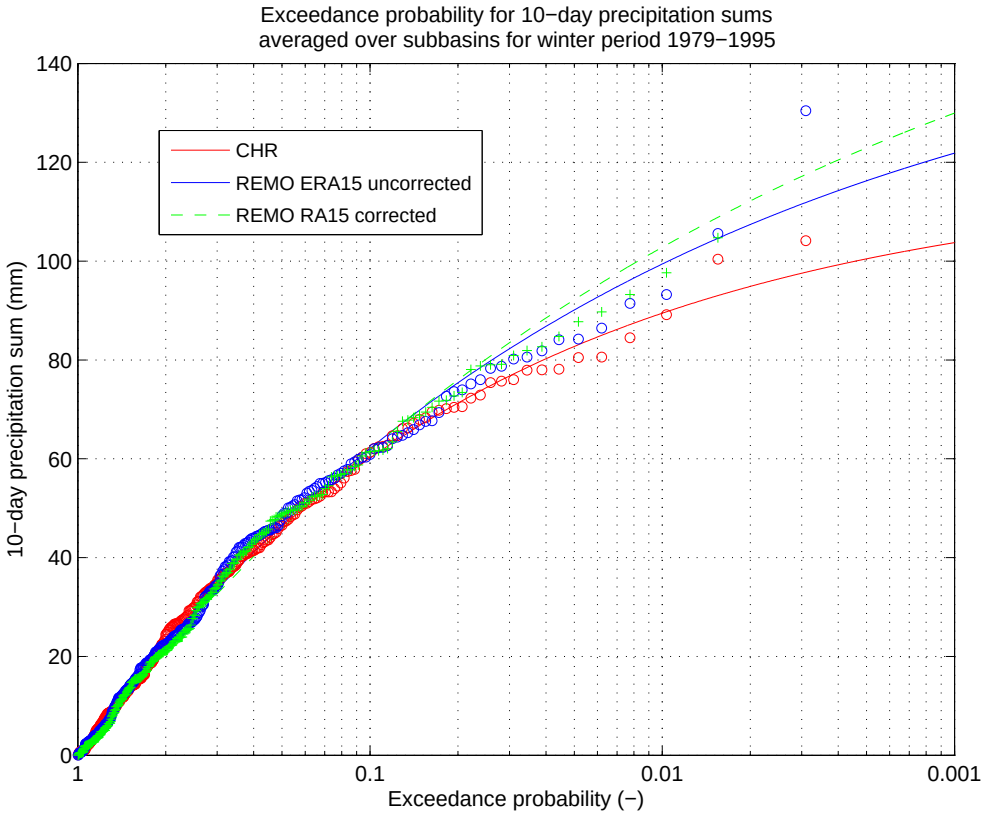



Figure 14: Exceedance probability of basin-averaged 10-day winter precipitation sums for the period 1979-1995. The 10-day precipitation sums are area-weighted averages.

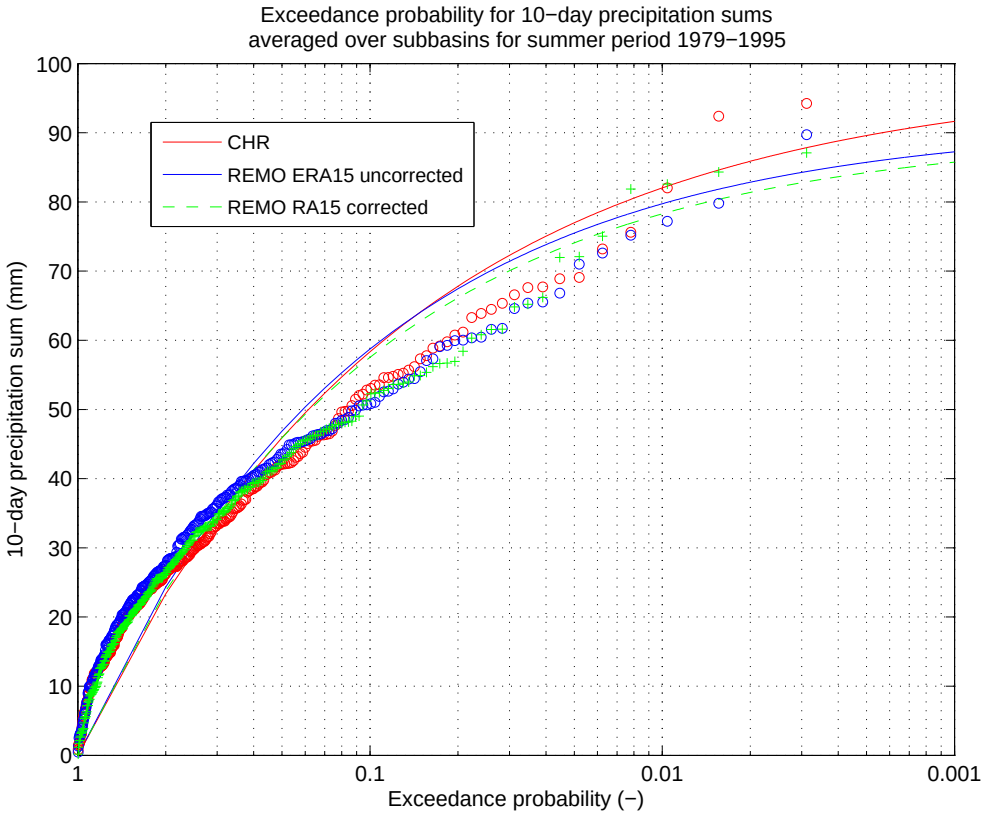



Figure 15: Exceedance probability of basin-averaged 10-day summer precipitation sums for the period 1979-1995. The 10-day precipitation sums are area-weighted averages.

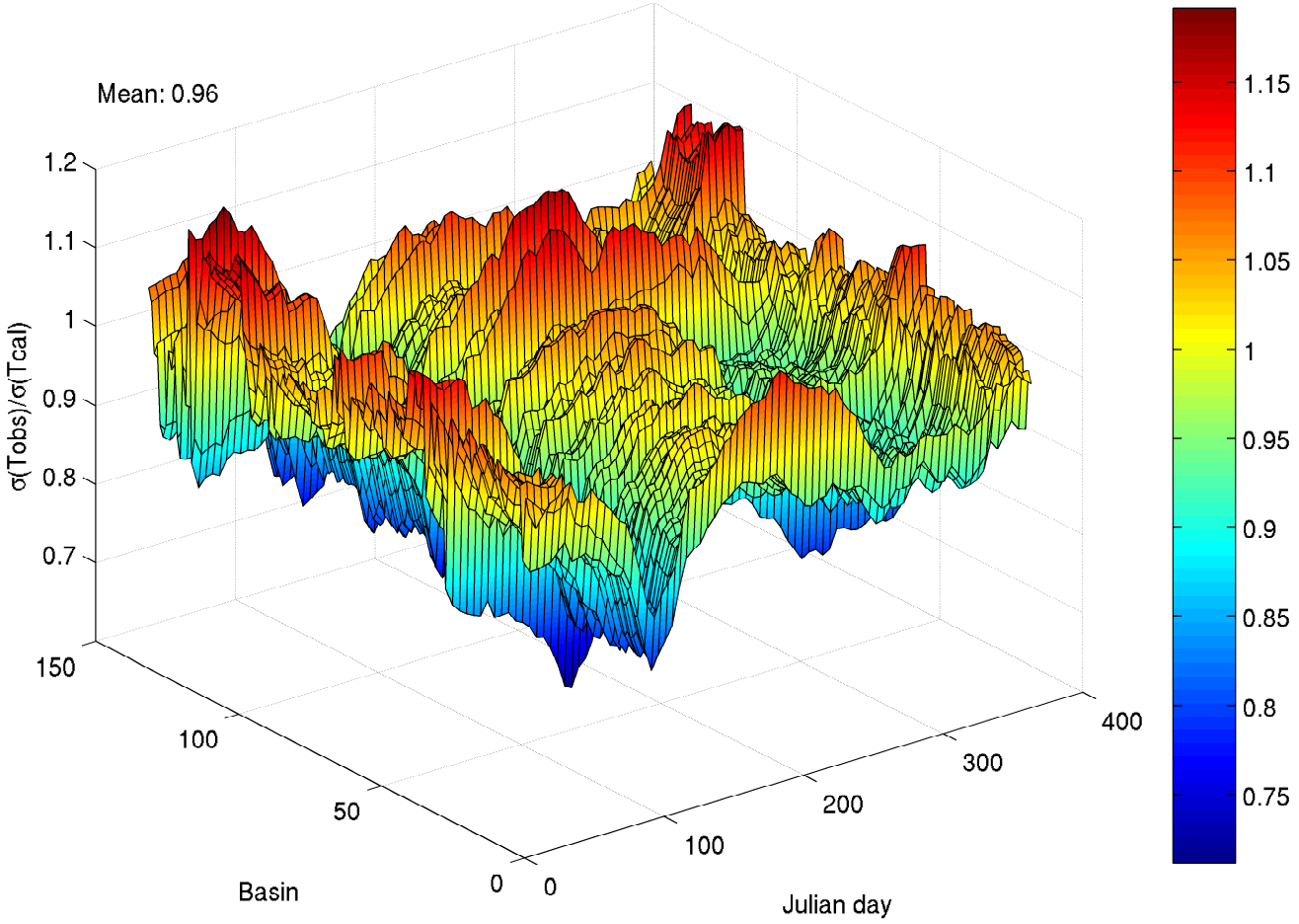



Figure 16: Ratio between $\sigma(T_{obs})$ and $\sigma(T_{cal})$ for each subbasin and 5-day period.

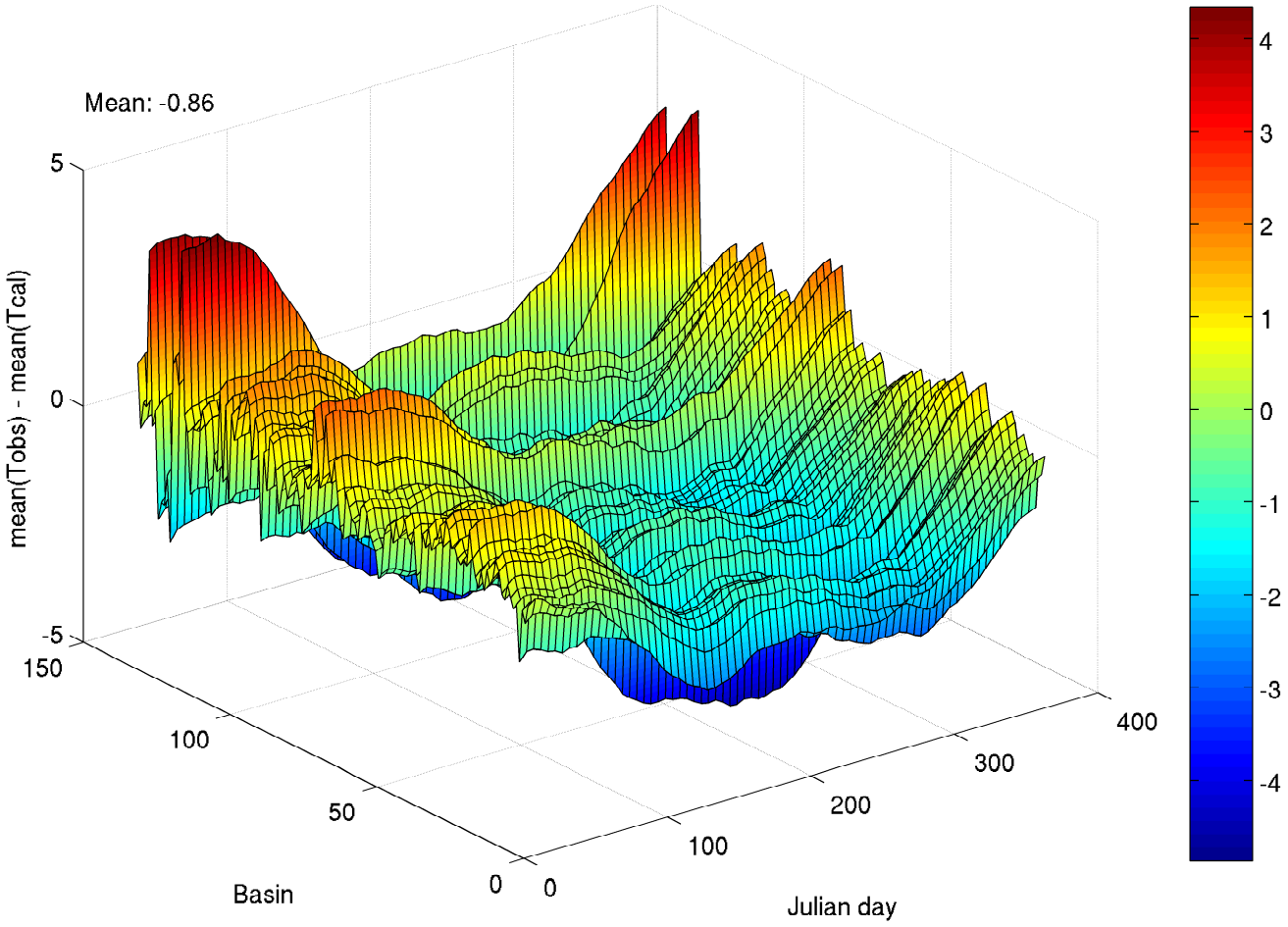



Figure 17: Difference between $\overline{T}_{cal}$ and $\overline{T}_{obs}$ for each subbasin and 5-day period.

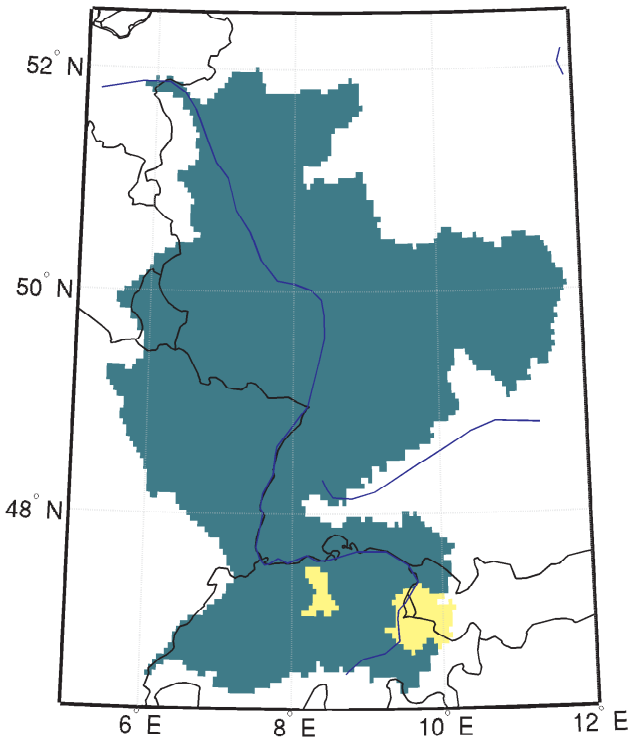



Figure 18: Two basins where the ERA15d data needs to be corrected from cold too warm in winter season. These basins are located in the Alps (yellow). These basins are known as the Rhein 2 and Limmat Reuss basin.

The next analysis is based on the average monthly temperature, averaged over the period 1979-1995. These averages are area-weighted. Figure 20 represents the variation of the 17-year average monthly temperatures throughout the year. In this figure the CHR observations, the uncorrected and the corrected ERA15d data are plotted. It is clear that the bias-corrected ERA15d data match the observed CHR data almost perfectly. The largest bias correction is applied during the months April until October.
The average basin temperature on a daily basis is plotted in Figure 21. The averages are area-weighted averages. The correlation between the corrected ERA15d daily temperature values and the observed CHR temperature values is satisfactory ($R^2$= 0.89). The uncorrected ERA15d daily temperature values also correlate well with the observed CHR temperature values. However, based on the narrower spread of the corrected ERA15d values around the $x = y$ line, it can be said that the corrected ERA15d values are more related to the CHR observation values than the uncorrected ones.
The difference between the CHR and ERA15d uncorrected average daily temperature over the period 1979-1995 is shown in Figure 22. A positive value corresponds to a higher temperature for the uncorrected ERA15d data set. The opposite is true for negative values. The differences in temperature vary between -1.5 and 3.5 °C. It can be concluded that the difference for the largest part of the Rhine basin is positive, which means that the uncorrected ERA15d temperature is warmer in that case. This positive difference is largest in the Alps. However, there are two subbasins in the Alps where the uncorrected ERA15d temperature is colder than the CHR temperature. These basins were mentioned before and are presented in Figure 18.
The difference between the CHR and ERA15d corrected average daily temperature over the period 1979-1995 is shown in Figure 22. It can be concluded that the bias correction for temperature leads to good results. The differences between the corrected ERA15d temperature values and the CHR temperature values have decreased enormously. The differences in temperature now vary between -0.4 and 0.4 °C, which can be seen as a significant improvement.

The relation between the daily average temperature and precipitation sum has been investigated for one subbasin. The chosen subbasin is the Rhein 2 basin (Figure 18, yellow basin on the right side). This

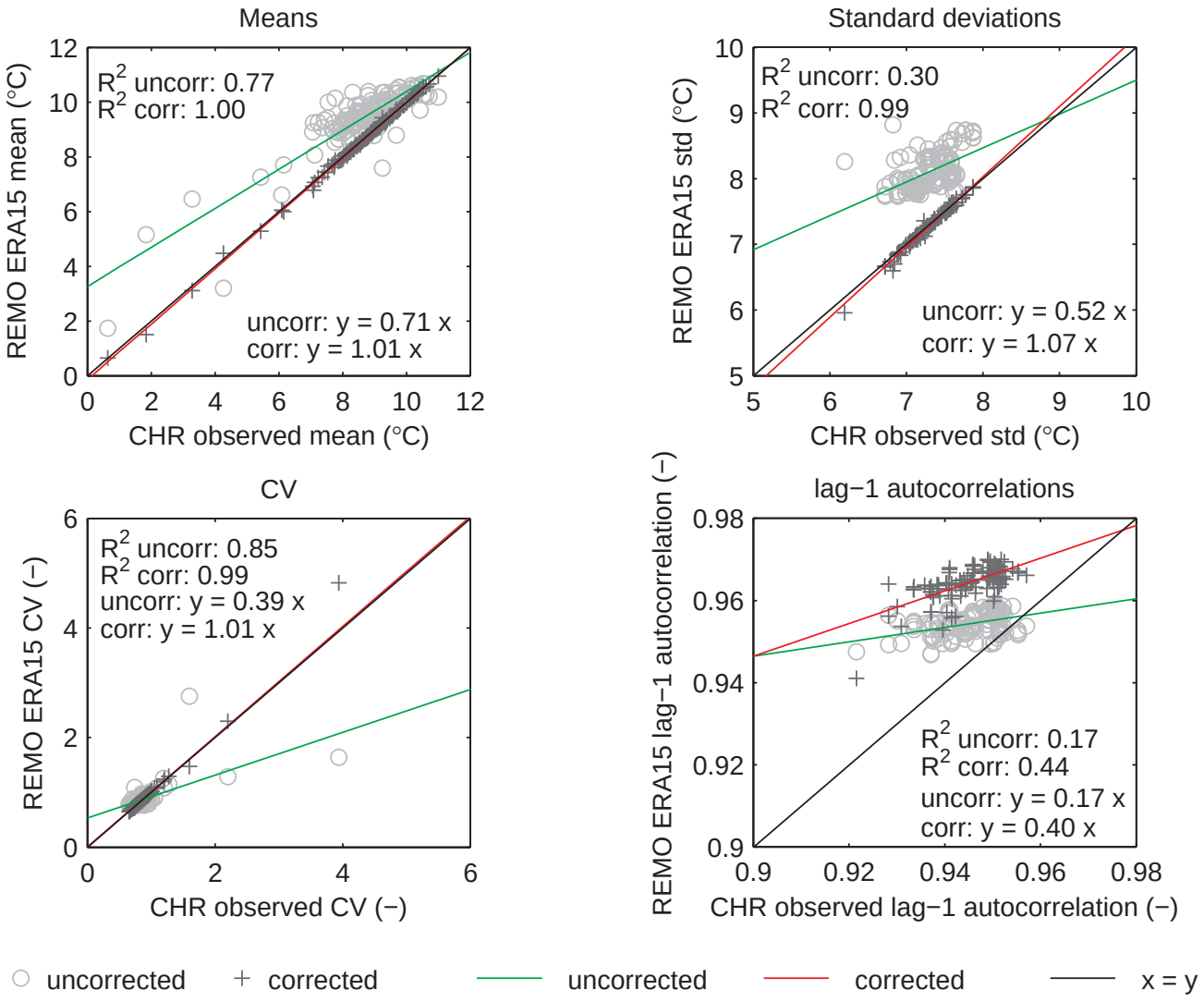



Figure 19: Scatter plots of the statistics of the observed CHR temperature data versus the corrected and uncorrected ERA15d temperature data. The statistics are calculated for each subbasin over the entire period 1979-1995. In each subplot the correlation coefficient ($R^2$) and slope of the linear regression line are plotted.

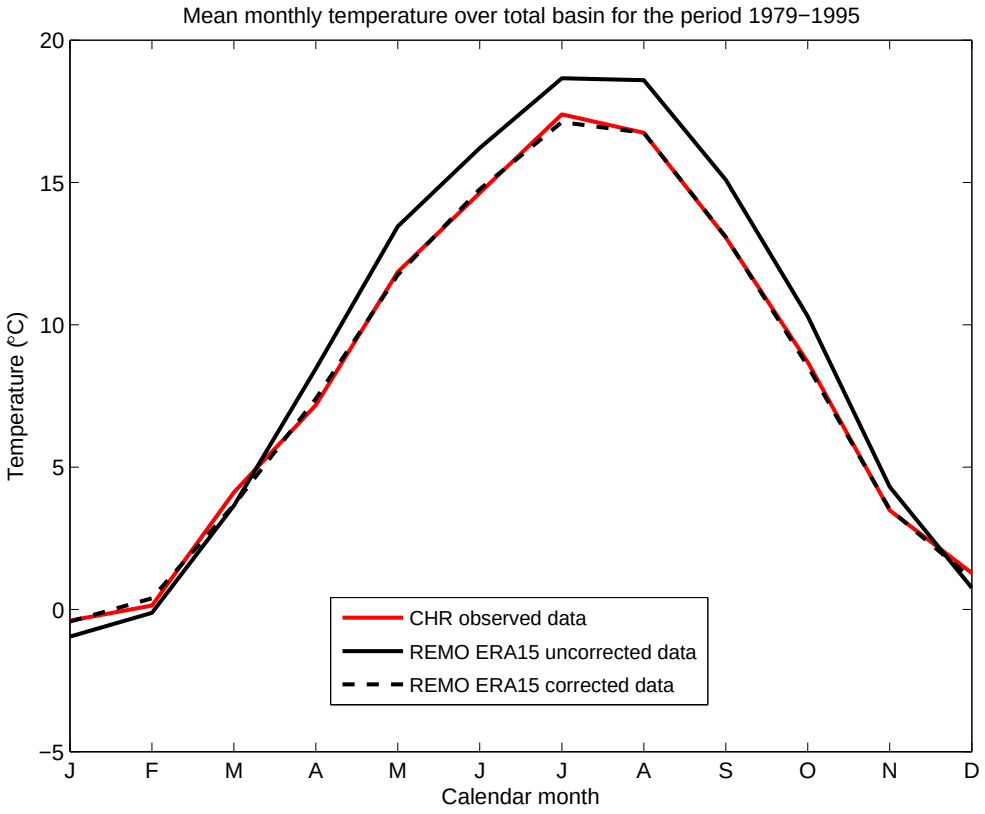



Figure 20: Area weighted monthly temperature over the entire basin for the period 1979-1995.

subbasin is chosen because the differences in temperature between the observed and ERA15d temperature before and after the bias-correction were largest for this basin. The correlation between these two variables is plotted in Figure 23. The goal of this analysis is to check whether or not any possibly existing relation between temperature and precipitation is disturbed after applying the bias correction. From Figure 23 it is clear that there is no direct relation between the daily average temperature and precipitation sum in all three cases. The distribution of the data points in the subplots is quite comparable.

## 4.2 Reference scenario

### 4.2.1 Precipitation

The same analysis which was done for the ERA15d data, has also been performed for the reference scenario. For each 5-day block and subbasin the parameters $a$ and $b$ have been determined. The computed parameter $a$ is shown in Figure 24. From Figure 24 it can be concluded that the parameter $a$ shows a seasonal pattern. The same was true in Figure 6 for the ERA15d data. The parameter value in winter is closer to 1 than in summer. This means that the bias in summer is larger than in winter. The largest part of the computed $a$ parameters have values smaller than one and the area-weighted average parameter $a$ has a value of 0.57. This suggests that, after correcting for the CV, the averages of the reference scenario are too wet when compared with the averages of the observed data.
The variation of parameter $b$ throughout the year and over the subbasins is shown in Figure 25. When Figure 25 is compared to Figure 24, the seasonal pattern of Figure 25 is less obvious. Most parameter values are larger than one, but this can lead to an increase as well as a decrease of precipitation, depending on the uncorrected daily precipitation value.
A graph of the area-weighted average parameters $a$ and $b$ is shown in Figure 26. In both the left and the right figure, the area-weighted average and standard deviation are plotted. The shaded area in both plots represents the fluctuation around the area-weighted average as measured by the area-weighted standard deviation. The area-weighted standard deviation of parameter $a$ is quite large when compared with that

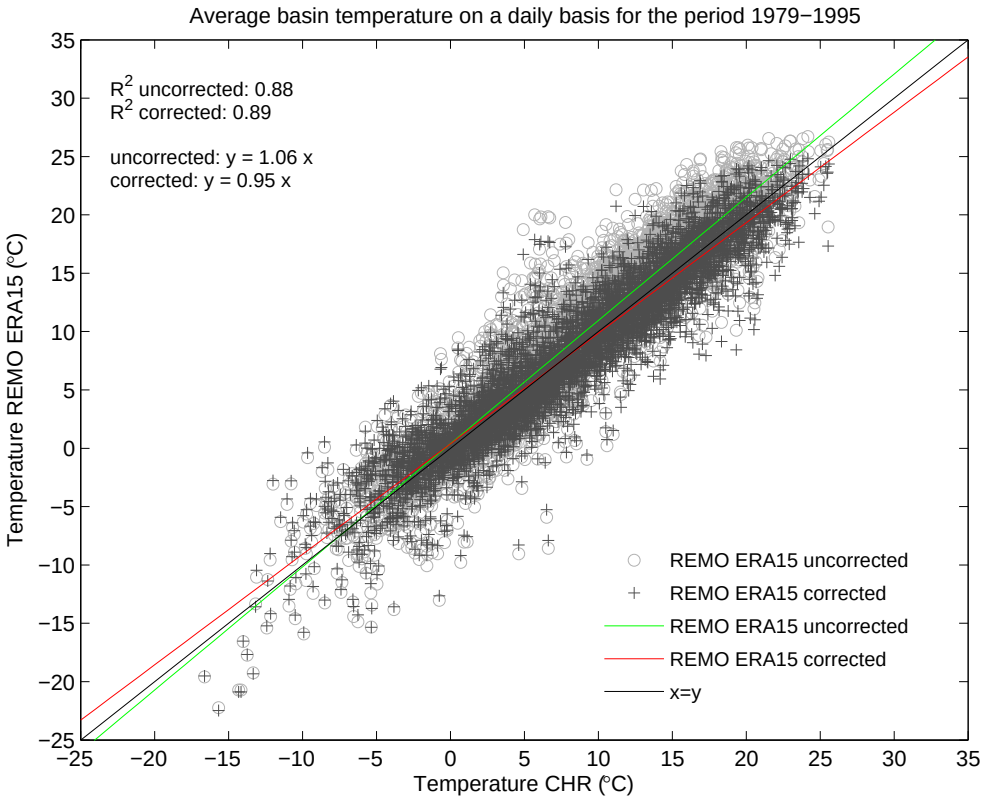



Figure 21: Scatter plot of average daily basin temperature. The CHR temperature is plotted against the ERA15d uncorrected and corrected temperature. The averages are area-weighted averages.

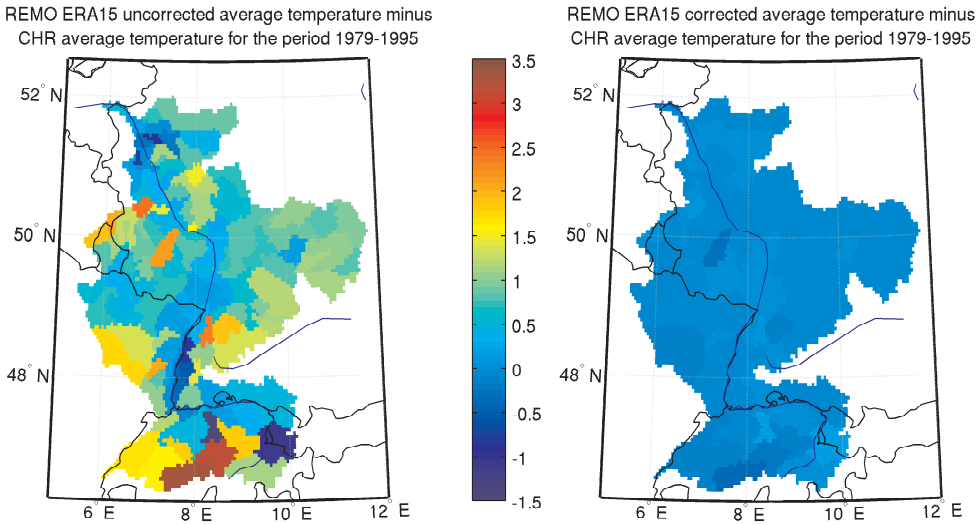



Figure 22: Difference in average daily temperature between the CHR data and the uncorrected ERA15d data for the period 1979-1995.

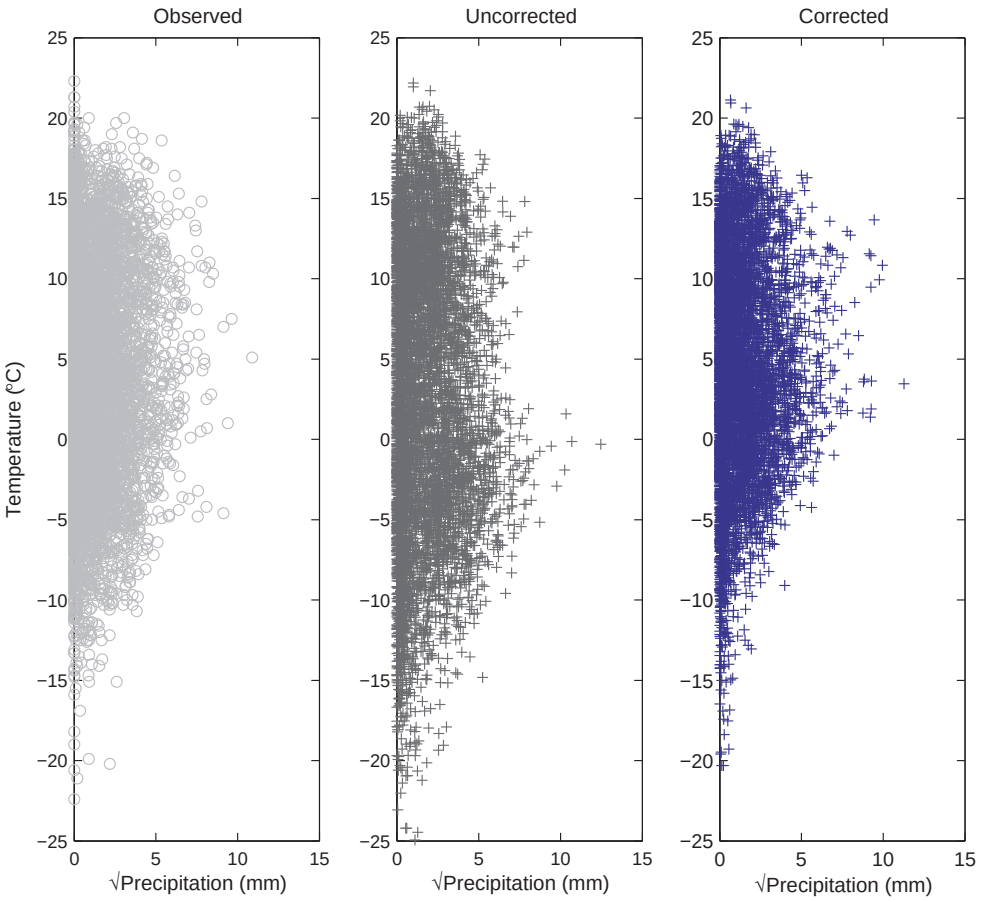



Figure 23: The correlation between the daily precipitation and temperature of the observed CHR data and uncorrected and corrected ERA15d data for the Rhein 2 subbasin for the period 1979-1995. The units on the x-axis are $[\mathrm{mm}^{-1/2}]$.

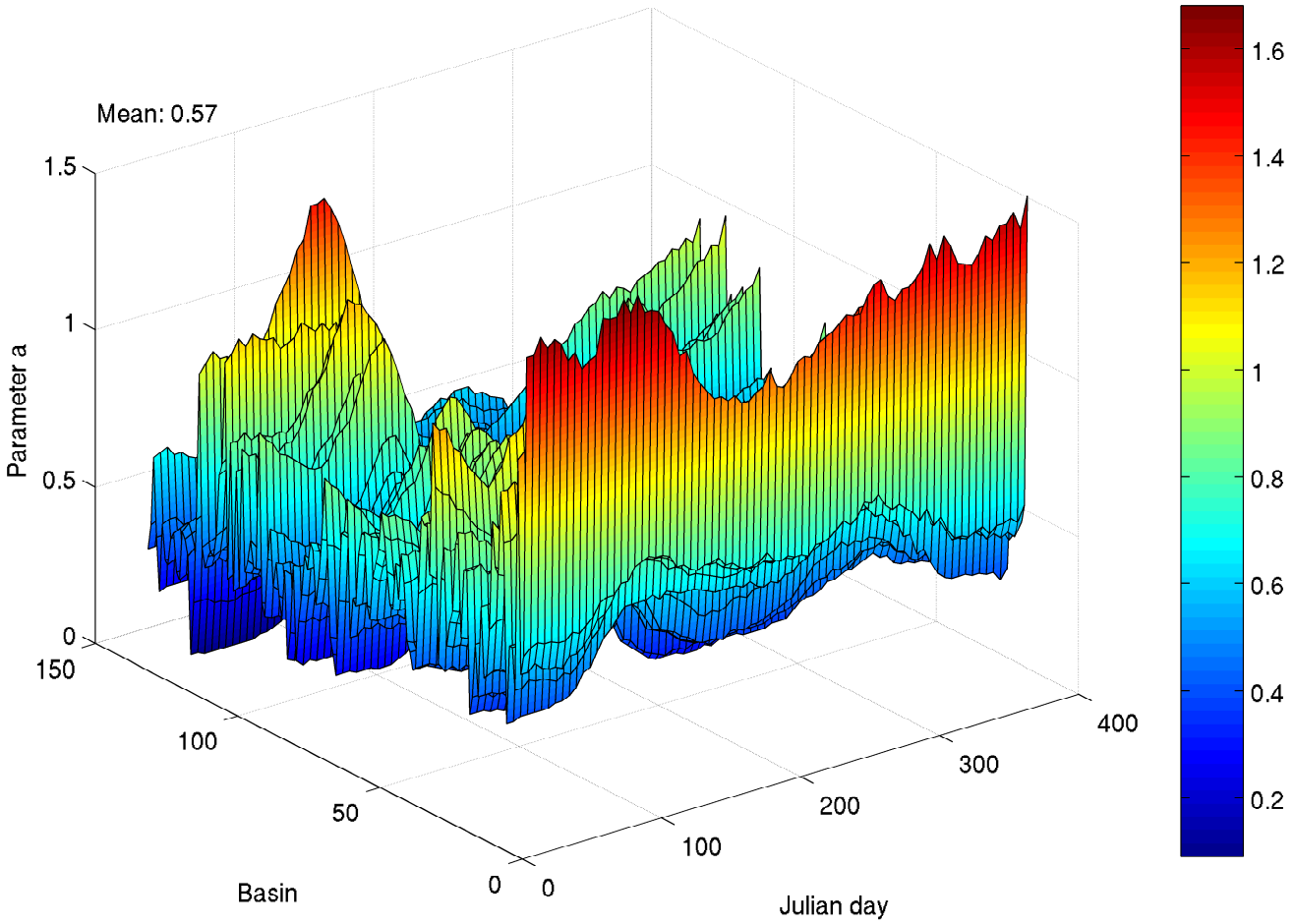



Figure 24: Derived parameter $a$ for each 5-day period and each subbasin.

of parameter $b$. The area-weighted standard deviation of parameter $a$ is quite constant over time. The area-weighted standard deviation of parameter $b$ is increasing after April. From this figure it is clear that the parameter $a$ has the effect to correct the reference scenario from wet to dry most of the time. The largest bias correction for both parameters is during the summer months. The months March and April have the smallest bias correction. After summer, the parameter $b$ remains quite high when compared to Figure 8. In addition, parameter $a$ is not approaching one, as was seen in Figure 8, but reaches only a value of 0.60.

Figure 27 represents the ratio of the area-weighted average corrected precipitation over the area-weighted average uncorrected precipitation. The black lines in this plot represent the fluctuation around the mean as measured by the area-weighted standard deviation. The dashed black line is the ratio equal to one. It can be seen that the ratio is around 0.80 over the entire period and that the precipitation is corrected to a drier value most of the time.
In Figure 28 the most important statistics of the corrected and uncorrected reference precipitation data are plotted versus the observed CHR precipitation data in 5 scatter plots. These statistics are calculated over the entire period 1961-1995 and for each subbasin separately. This results in 134 data points in each subplot. As seen before with the ERA15d data, the correction for the mean and CV also works fine for the reference scenario. The correlation coefficients are close to one and the corrected data points are close to the $x = y$ line. The correction seems to have a positive effect on the fraction of wet days too. The effect of the bias correction on the lag-1 autocorrelation is small. This was not the case with the ERA15d data where the effect had a positive influence on this statistic.

For each subbasin the average daily precipitation over the period 1961-1995 has been calculated. This

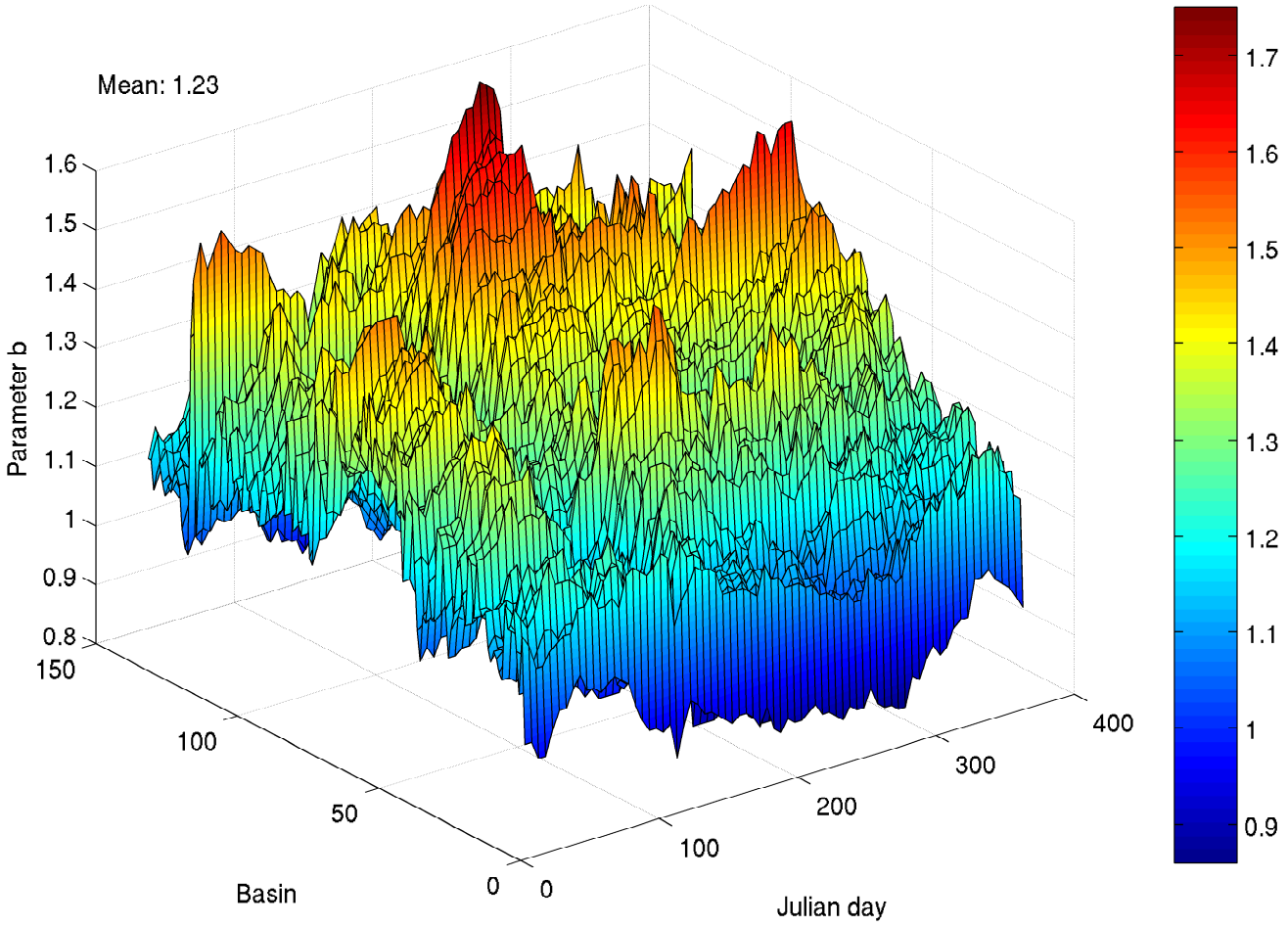



Figure 25: Derived parameter $b$ for each 5-day period and each subbasin.

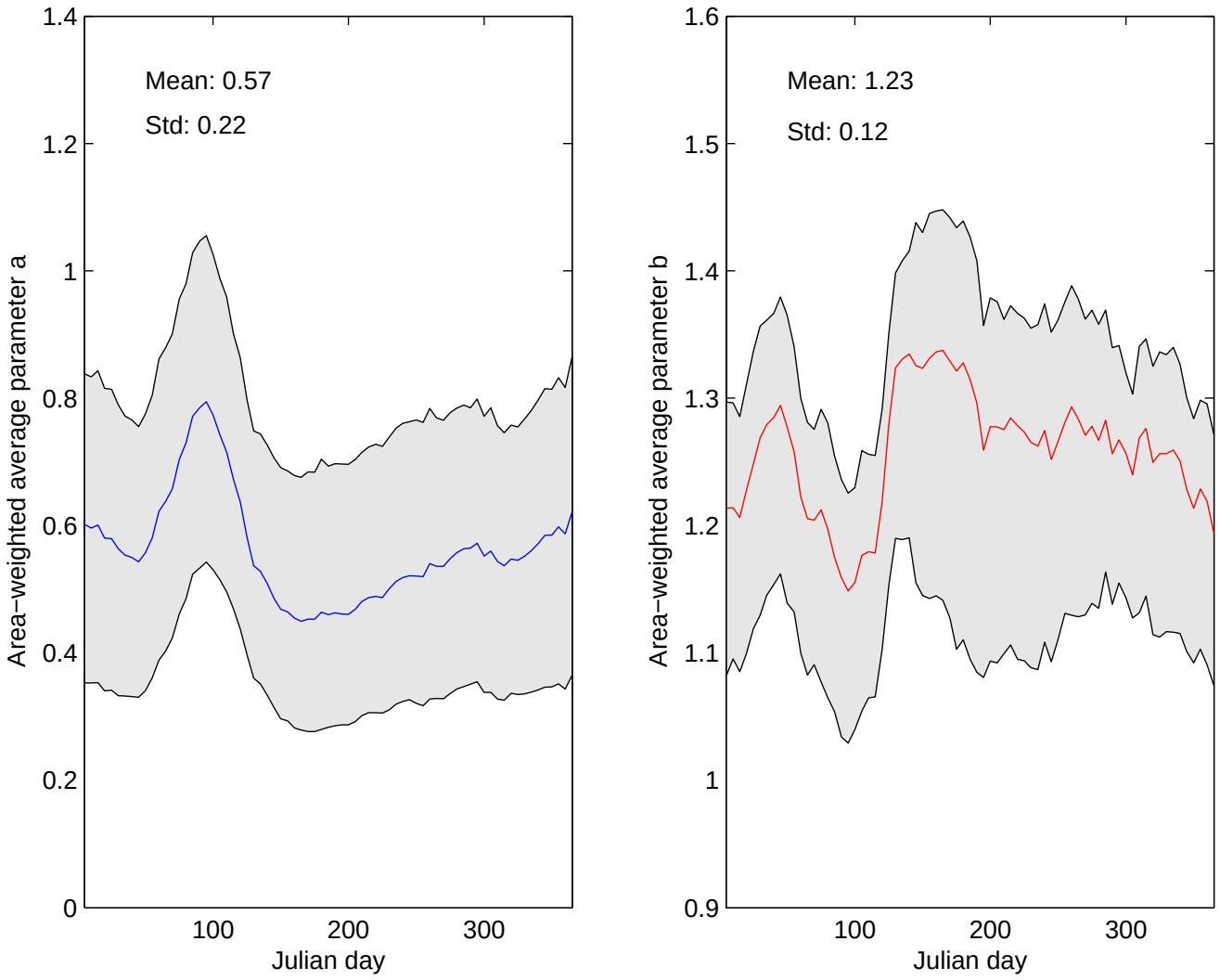



Figure 26: Area weighted average parameters $a$ and $b$ for each 5-day period.

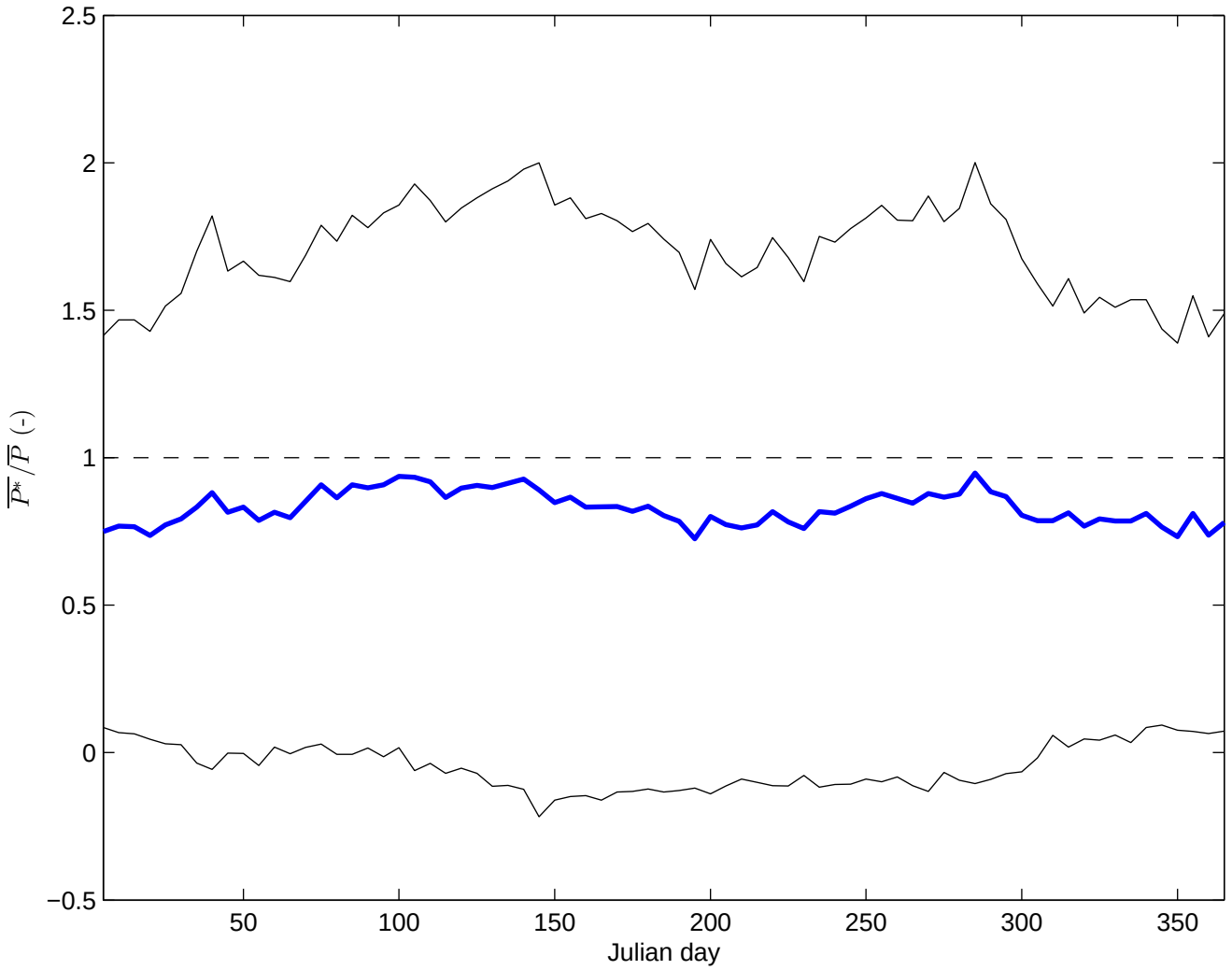



Figure 27: Ratio of area-weighted average corrected precipitation over the area-weighted average uncorrected precipitation (bold line). The thin black lines represent the mean plus or minus one standard deviation.

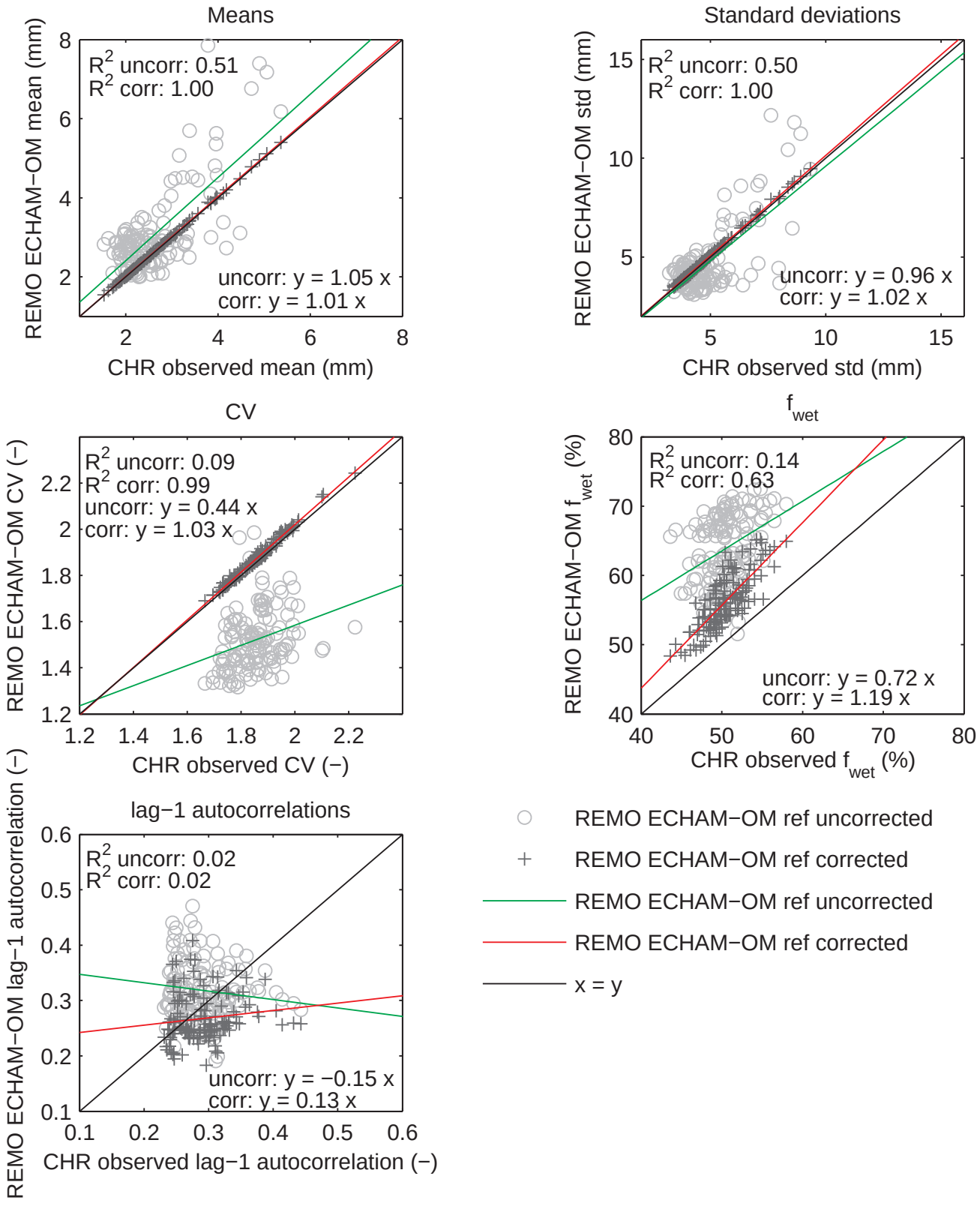



Figure 28: Scatter plots of the statistics of the observed CHR precipitation data versus the corrected and uncorrected reference precipitation data. The statistics are calculated for each subbasin over the entire period 1961-1995. The fraction of wet days ($f_{wet}$) is the percentage of days where P>0.3 mm. In each subplot the correlation coefficient ($R^2$) and slope of the linear regression equation are plotted.

has been done for the CHR observed data and the uncorrected and corrected reference scenario. The difference in daily precipitation between the CHR observations and the uncorrected reference scenario is shown in Figure 29. The difference between the two data sets varies between -1.5 and 4 mm/day. A positive value means that the reference scenario precipitation is wetter than the CHR precipitation. The difference is positive in the largest part of the Rhine basin, which means that the reference scenario is wetter in those areas. The reference scenario is especially wetter in Switzerland and the basins located near the Rhine. However, there are certain areas where the observations are wetter than the reference scenario. This is the case for the Black Forest and some basins located in the southern part of Germany. The difference in daily precipitation between the corrected reference scenario and the CHR observations is shown in Figure 29. The differences between the bias-corrected reference scenario and the observations turn out to be very small. The bias correction seems to have a positive effect on the average daily values.

The next analysis is based on the monthly precipitation sums, averaged (area-weighted) over the period 1961-1995. Figure 30 represents the variation of the 35-year averaged monthly precipitation sum throughout the year. In this figure the CHR observations, the uncorrected and the corrected reference scenario are plotted. It is clear that the uncorrected reference scenario precipitation is too wet throughout the entire year, except for March. However, it seems that the bias-corrected reference scenario precipitation matches the observed CHR precipitation quite well. The bias correction has some difficulties in correcting the precipitation for March. The CHR observations show an increase in precipitation during the shift from February to March. The opposite is true for the reference scenario. It turns out to be difficult to correct for these different shifts in precipitation.

A histogram has been made of the area-weighted 10-day precipitation sums over the entire period 1961-1995. This histogram is presented in Figure 31. It is clear that the histogram of the CHR observed data is almost similar to that of the corrected reference scenario data. Large 10-day precipitation frequencies occur within the 10-40 mm range. The uncorrected reference scenario data has a longer tail to the right. It looks like the large values are filtered out in the corrected reference scenario.
The exceedance probability has been investigated for the 10-day winter precipitation sums. This is illustrated in Figure 32 for the CHR data and the uncorrected and corrected reference scenario. A Generalized Pareto distribution is fitted through the data. The 10-day precipitation sums are area-weighted averages. The corrected reference scenario matches the CHR data very well. The corrected reference scenario matches the observed CHR data better for all exceedance probabilities than the uncorrected reference scenario does.
The exceedance probability for 10-day summer precipitation sums has been investigated as well. These results are shown in Figure 33. A Generalized Pareto distribution is fitted through the data. The 10-day precipitation sums are area-weighted averages. Again the corrected reference scenario is a better match to the observed CHR data for all exceedance probabilities.

### 4.2.2 Temperature

The method for calculating the bias correction is described in detail in Section 3.3. The uncorrected daily temperature is corrected with the use of equation 3. The temperature in this equation is adjusted for the standard deviation and the mean. The ratio between $\sigma(T_{obs})$ and $\sigma(T_{cal})$ is shown in Figure 34 for each subbasin and 5-day block. A clear seasonal pattern is visible in this figure. In winter $\sigma(T_{obs})$ seems to be higher than $\sigma(T_{cal})$ in most subbasins. During summer, the ratio between $\sigma(T_{obs})$ and $\sigma(T_{cal})$ is close to one, which means that the correction for standard deviation is largest in winter. The area-weighted average ratio between $\sigma(T_{obs})$ and $\sigma(T_{cal})$ is 1.09. This suggests that the average spread in temperature for the observations is larger than the average spread in temperature for the reference scenario.
The difference between $\overline{T}_{cal}$ and $\overline{T}_{obs}$ is shown in Figure 35 for each subbasin and 5-day block. In this figure an even better seasonal pattern can be distinguished. The difference between $\overline{T}_{cal}$ and $\overline{T}_{obs}$ is increasing from February until April. An increasing negative difference means that the 35-year average of the reference scenario is warmer than the-35 year average of the observations. The difference is decreasing from April until September, which means that the 35-year average of the reference scenario is approaching the 35-year average of the observations. From September until the end of the year the

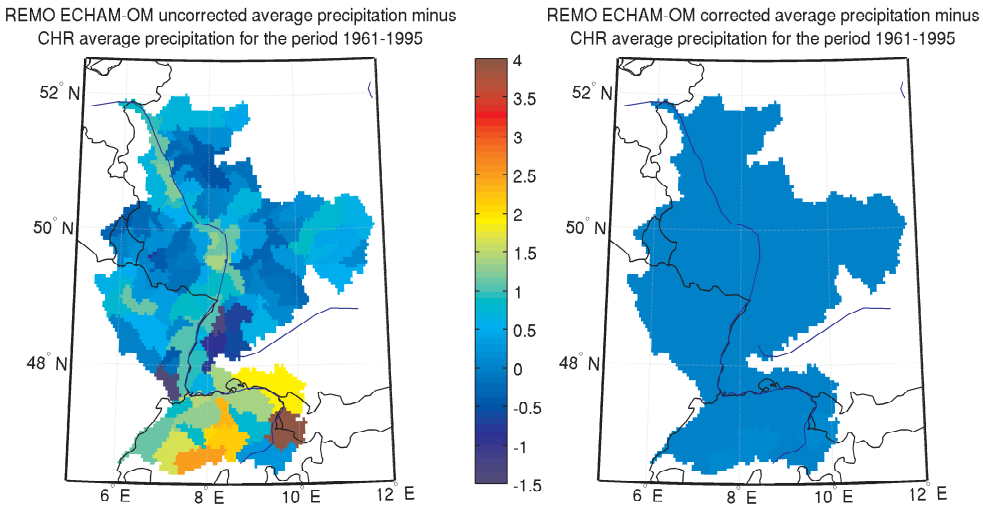



Figure 29: Difference in average daily precipitation between the CHR data and the corrected and uncorrected reference scenario for the period 1961-1995.

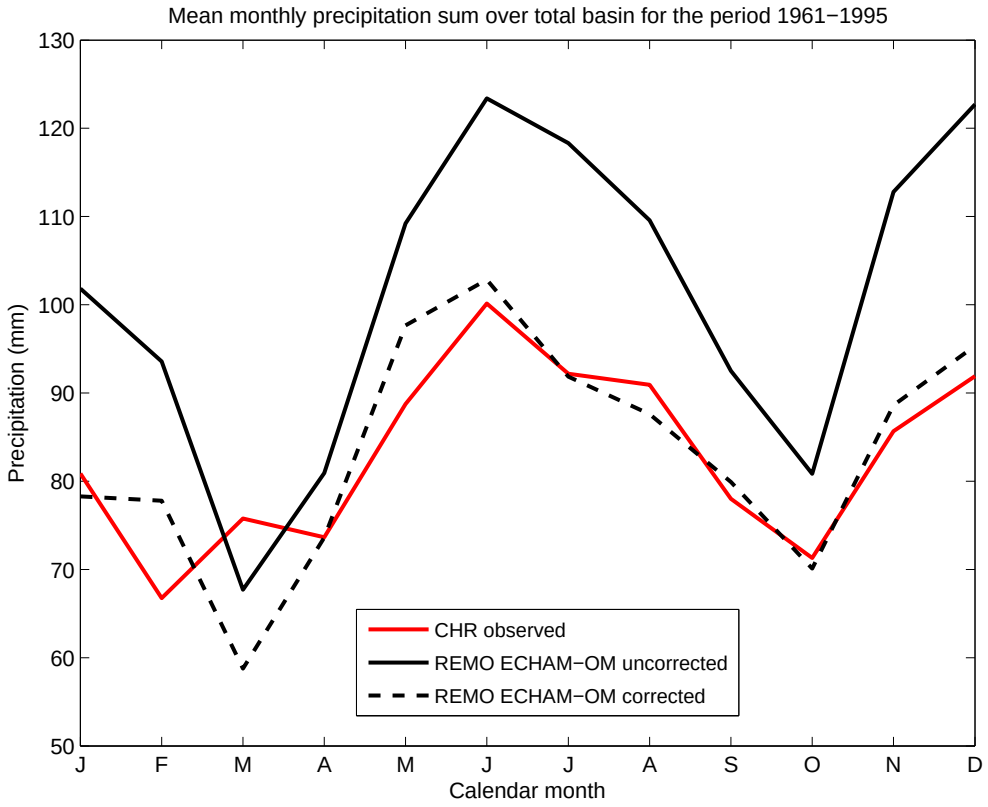



Figure 30: Area-weighted monthly precipitation sum over the entire basin for the period 1961-1995.

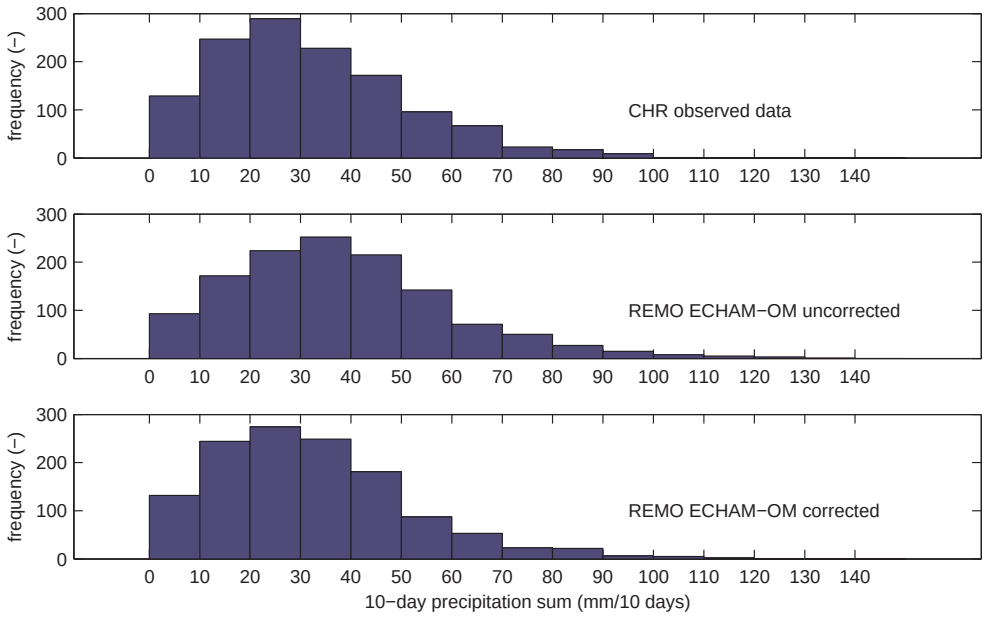



Figure 31: Histogram of the area-weighted 10-day precipitation sums for the period 1961-1995.

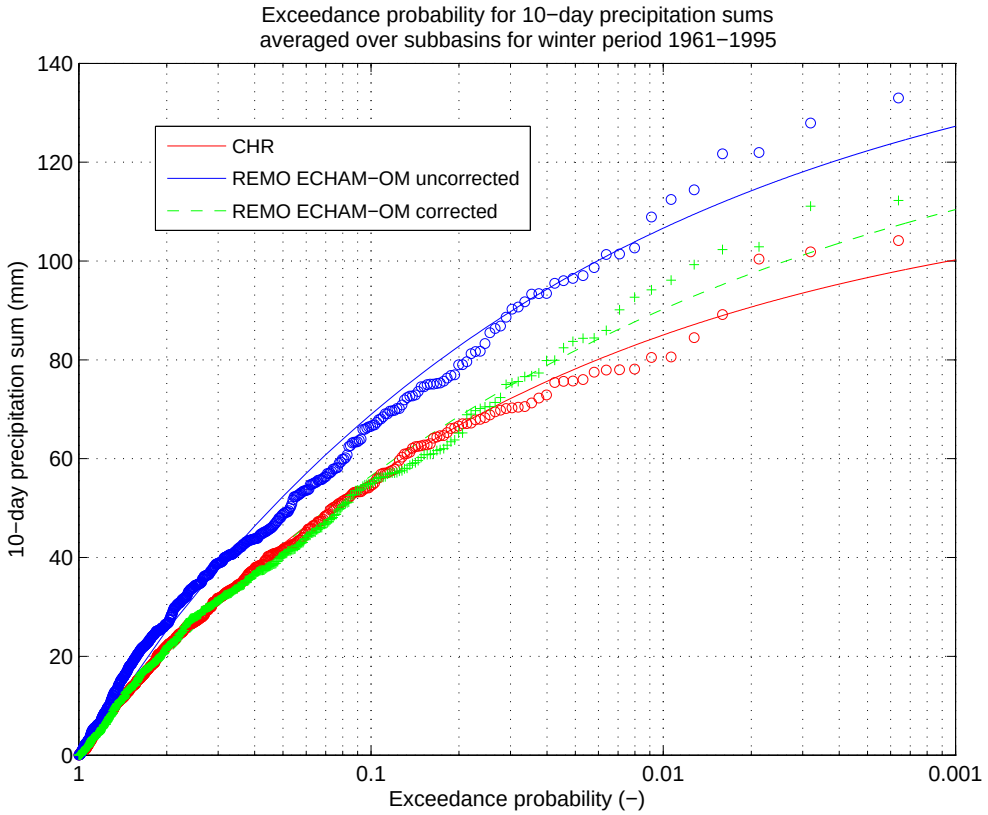



Figure 32: Exceedance probability of basin-averaged 10-day winter precipitation sums for the period 1961-1995. The 10-day precipitation sums are area-weighted averages.

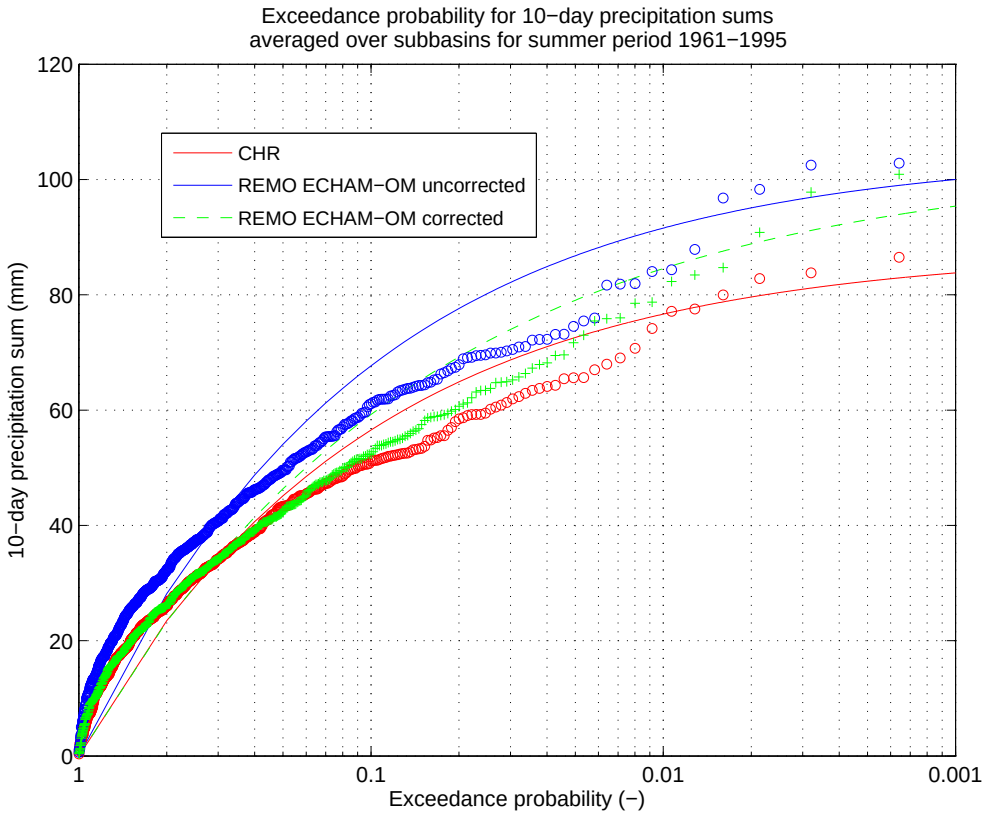



Figure 33: Exceedance probability of basin-averaged 10-day summer precipitation sums for the period 1961-1995. The 10-day precipitation sums are area-weighted averages.

difference is increasing again. The overall average between $\overline{T}_{cal}$ and $\overline{T}_{obs}$ is -0.76 °C. This suggests that the 35-year basin average temperature of the reference scenario is too warm when compared to the observed average temperature.

In Figure 36 the most important statistics of the corrected and uncorrected reference temperature data are plotted versus the observed CHR temperature data in four scatter plots. These statistics are calculated over the entire period 1961-1995 and for each subbasin separately. This results in 134 data points in each subplot. As mentioned before, the applied bias correction method corrects for the mean and the standard deviation. This is clearly visible in the plots of the mean, standard deviation and CV. The correlation coefficients of the bias-corrected reference data are close to one or even one for these statistics. Although it was not the intention to correct for the lag-1 autocorrelation, it seems that the applied bias correction has a positive effect on this as well. The correlation coefficient has increased from 0.06 to 0.46 for this statistic.

The next analysis is based on the average monthly temperature, averaged over the period 1961-1995. These averages are area-weighted. Figure 37 presents the variation of the 35-year average monthly temperature throughout the year. In this figure the CHR observations, the uncorrected and the corrected reference scenario are plotted. The bias-corrected reference scenario temperatures match the observed CHR temperatures very well. The largest bias correction is applied in winter, which is in agreement with Figure 34 and Figure 35.

The average basin temperature on a daily basis is plotted in Figure 38. Again the average is a an area-weighted average. There is hardly any difference in correlation between the corrected and uncorrected reference scenario temperatures and the observed temperatures. The spread of the corrected reference scenario values around the $x = y$ line seems to be a bit narrower when compared to the uncorrected ones. However, the $R^2$ of the uncorrected temperature values is higher than the $R^2$ of the corrected temperature values. The correlation between the ERA15d daily temperature values and the observations, as was seen in Figure 21, is higher than the correlation between the reference scenario daily temperatures and the

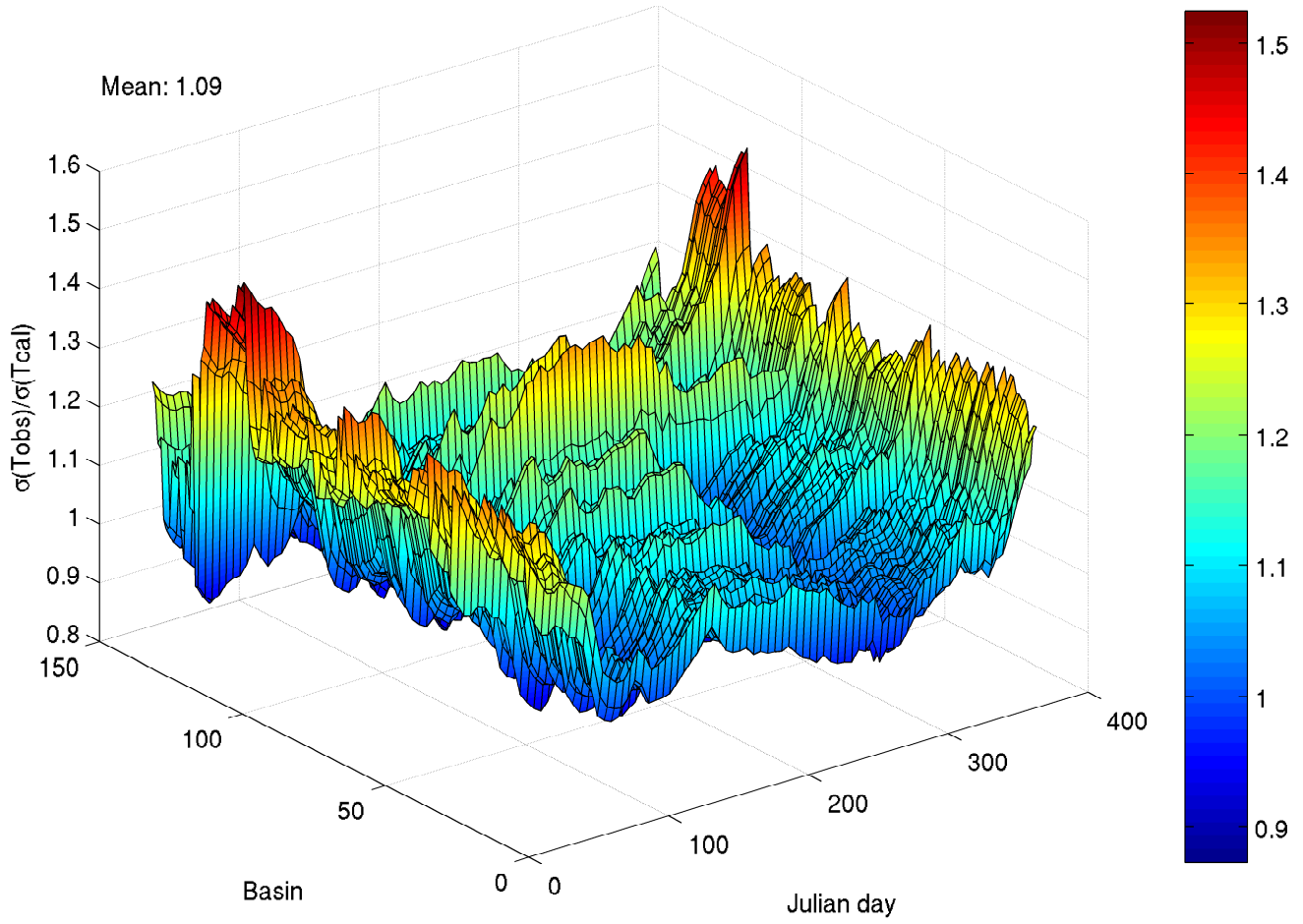



Figure 34: Ratio between $\sigma(T_{obs})$ and $\sigma(T_{cal})$ for each subbasin and 5-day period.

observations.
The difference between the CHR and uncorrected reference scenario average daily temperature over the period 1961-1995 is shown in Figure 39. A negative value corresponds to a higher temperature for the CHR observed temperature. The opposite is true for positive values. The differences in temperature vary between -1.5 and 3 °C. It can be concluded that the difference for the largest part of the Rhine basin is positive, which means that the uncorrected reference scenario temperature is warmer in that case. This was also the case for the ERA15d data set. The largest positive differences can again be found in the Alps. There are two subbasins in the Alps where the uncorrected reference scenario temperature is colder than the CHR temperature. These two subbasins (Figure 18) were also colder in the ERA15d data set.
The difference between the CHR and corrected reference scenario average daily temperature over the period 1961-1995 is shown in Figure 39. It can be concluded that the bias correction for temperature leads to good results. The differences between the corrected reference scenario temperature values and the CHR temperature values have decreased enormously. The differences in temperature now vary between -0.3 and 0.5 °C, which can be seen as a significant improvement.

As before with the ERA15d data, the relation between the daily average temperature and precipitation sum of the reference data has been investigated for one subbasin as well. The correlation between these two variables is plotted in Figure 40. The goal of this analysis is to check whether or not any possibly existing relation between temperature and precipitation is disturbed after applying the bias correction. From Figure 40 it is clear that there is no relation between the daily average temperature and precipitation sum in all three cases. The distribution of the data points in the subplots is quite comparable.

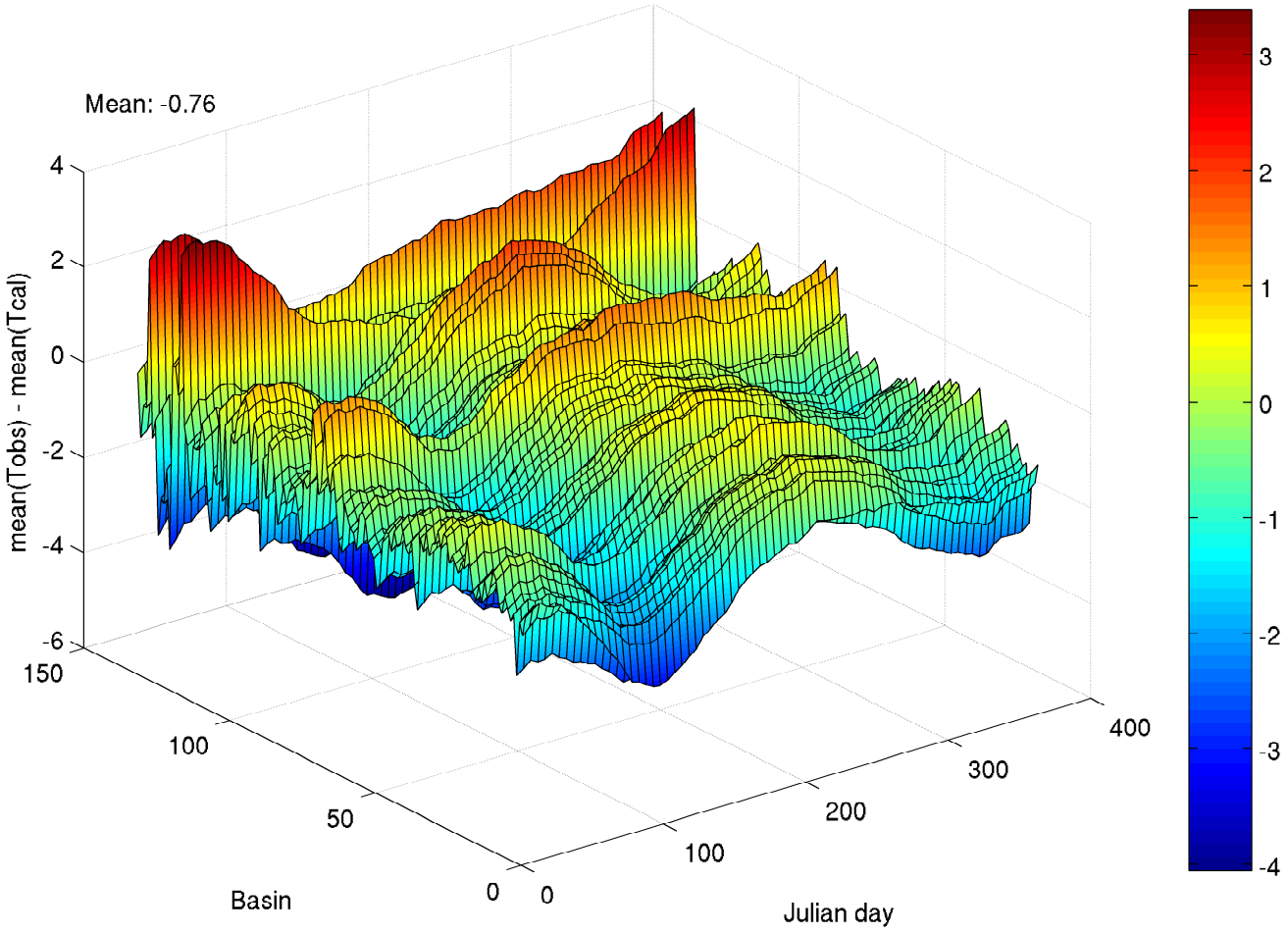



Figure 35: Difference between $\overline{T}_{cal}$ and $\overline{T}_{obs}$ for each subbasin and 5-day period.

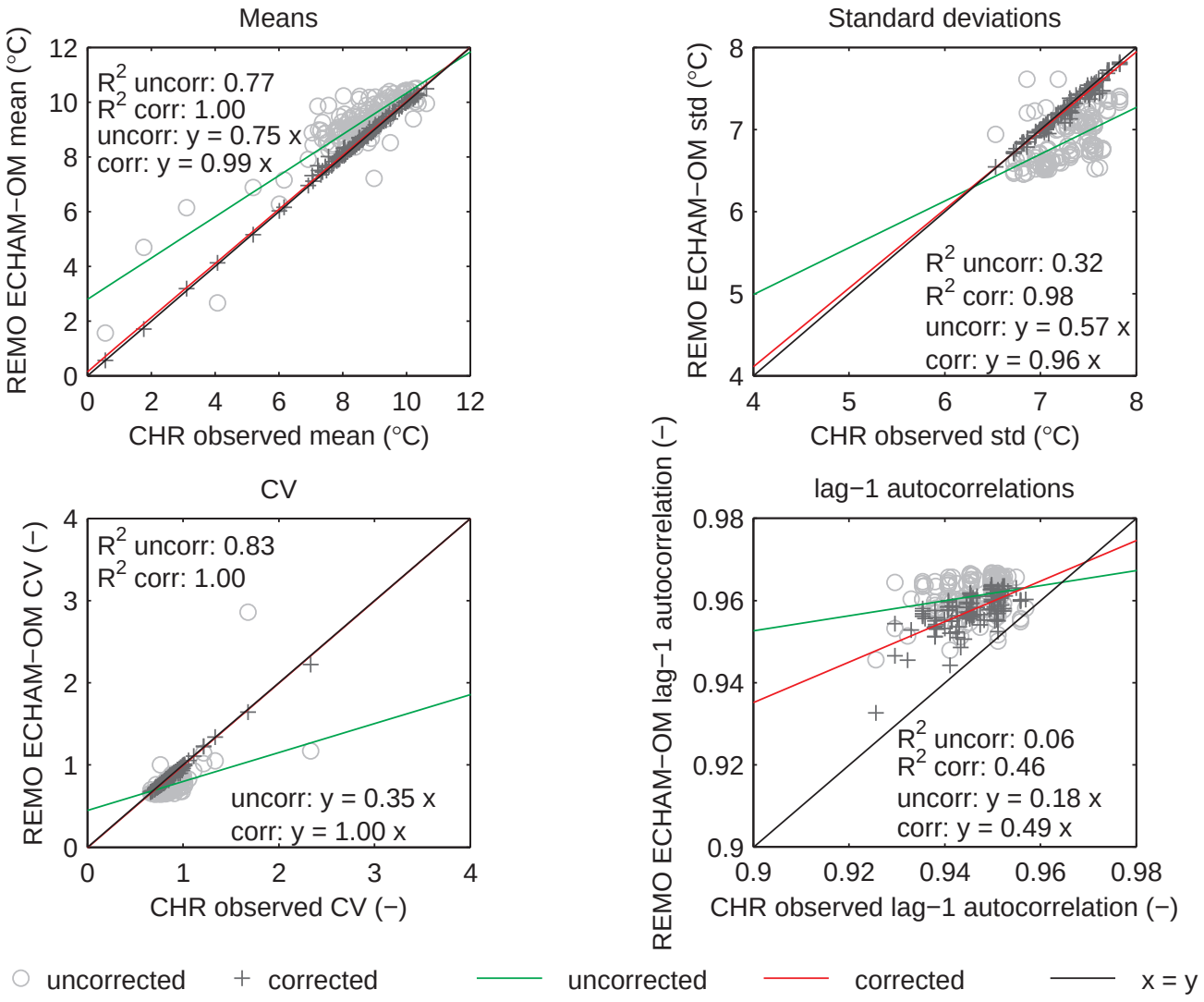



Figure 36: Scatter plots of the statistics of the observed CHR temperature data versus the corrected and uncorrected reference temperature data. The statistics are calculated for each subbasin over the entire period 1961-1995. In each subplot the correlation coefficient ($R^2$) and slope of the linear regression equation are plotted.

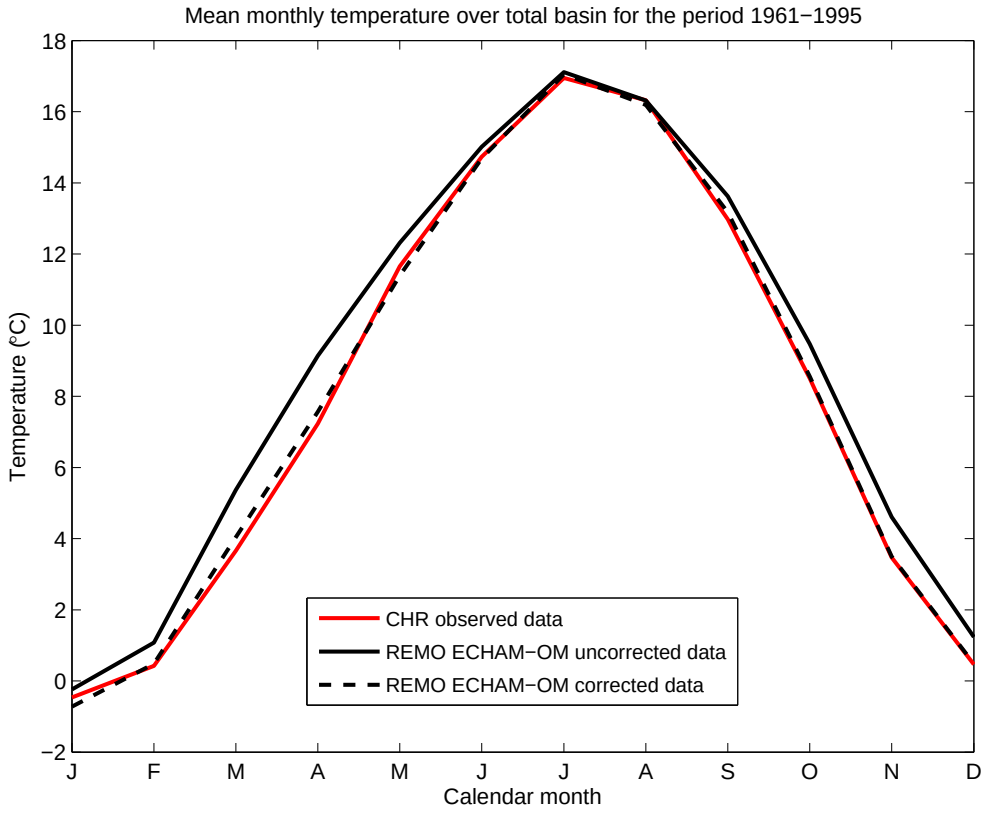



Figure 37: Area-weighted monthly temperature over the entire basin for the period 1961-1995.

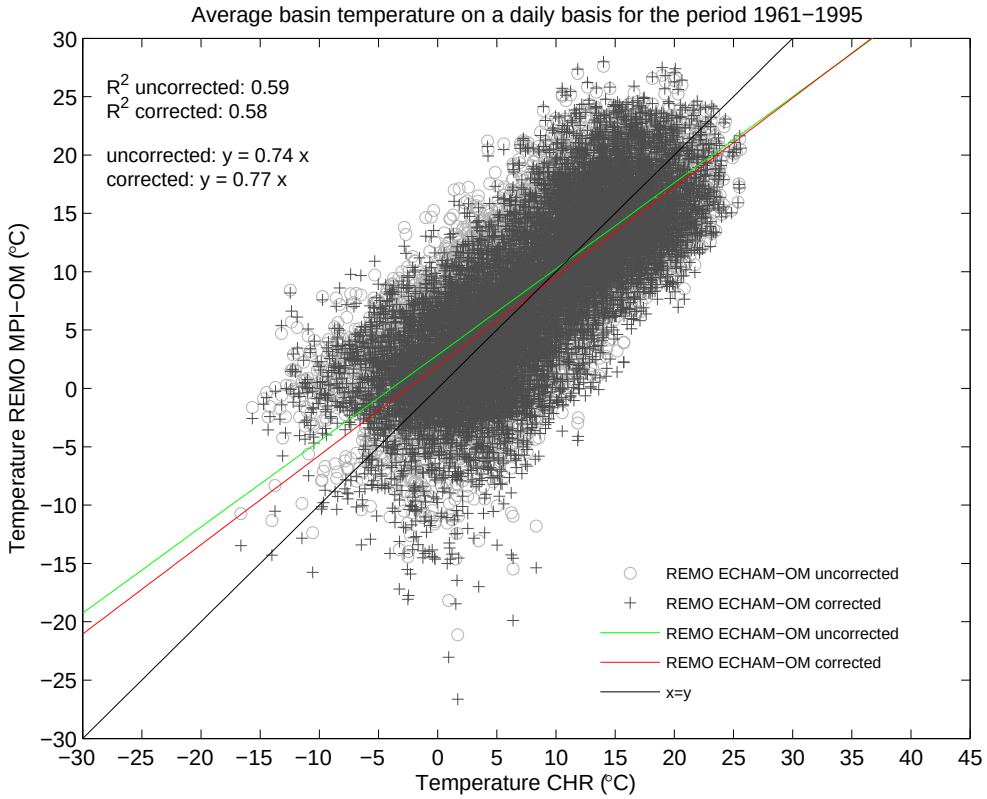



Figure 38: Scatter plot of average daily basin temperature. The CHR temperature is plotted against the uncorrected and corrected reference scenario temperature. The averages are area-weighted averages.

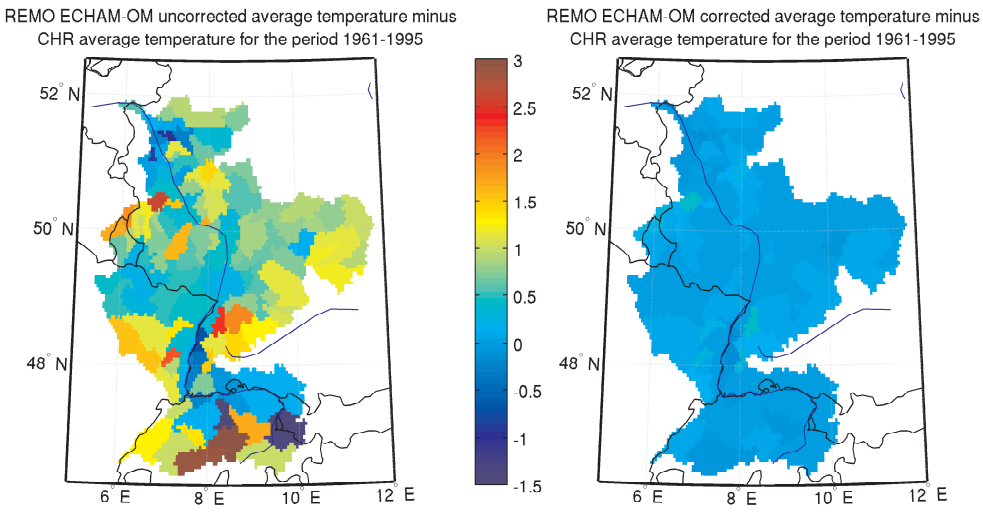



Figure 39: Difference in average daily temperature between the CHR data and the uncorrected reference scenario for the period 1961-1995.

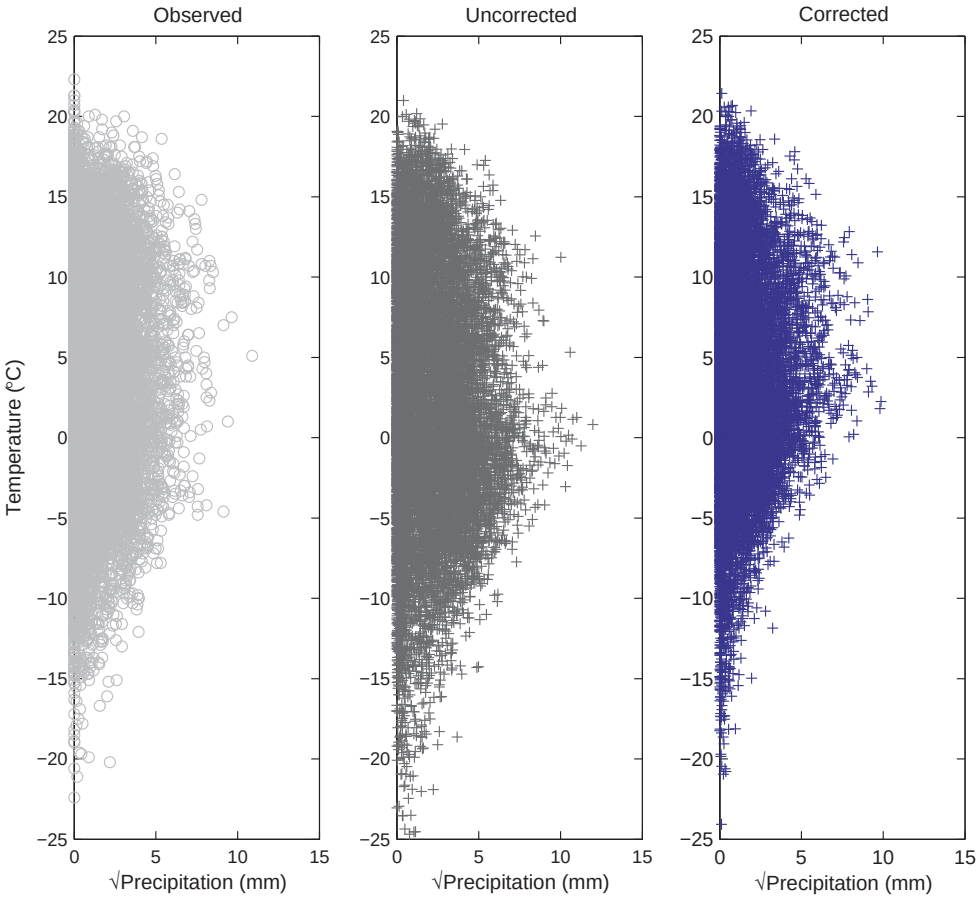



Figure 40: The correlation between the daily precipitation and temperature of the observed CHR data and uncorrected and corrected reference data for the Rhein 2 subbasin for the period 1961-1995. The units on the x-axis are [$\mathrm{mm}^{-1/2}$].

# 5 Conclusions and recommendations

## 5.1 Conclusions

In the previous sections the observed CHR, ERA15d and reference scenario data sets have been analyzed. Both the ERA15d and reference scenario data set have been corrected for a bias in temperature and precipitation. The results of the applied bias correction for temperature and precipitation were quite satisfactory.
Before applying the bias correction, the ERA15d data was somewhat too wet and warm for the largest part of the Rhine basin when compared with the observations. The difference between the bias-corrected ERA15d data and the observations turned out to be very small. The employed bias correction method corrects for the mean and CV of the precipitation data. Besides this, it seems that it corrects other statistics as well. The correlation between the fraction of wet days of the observed CHR data and the ERA15d data increased significantly after bias correction. The same is true for the lag-1 autocorrelation. The bias correction also has a positive effect on the exceedance probabilities of the 10-day precipitation sums for summer. The same is true for winter, but for exceedances $< 1/20$, the uncorrected ERA15d data matches the CHR observations better. The bias correction applies very well for the ERA15d temperature data. The monthly average temperature of the corrected ERA15d data is almost similar to that of the CHR observed monthly average temperature. Again the bias correction has a postitive effect on the lag-1 autocorrelation, while it only corrects for the mean and standard deviation.
The reference scenario was too warm and wet for the largest part of the Rhine basin before applying the bias correction. The method of bias correction seems to have a positive effect on both the temperature and precipitation data. The spatial differences between the observed and bias-corrected precipitation data have decreased significantly. The bias correction method had a positive effect on the fraction of wet days and lag-1 autocorrelation for the ERA15d data. The same is true for the reference scenario. Both the 10-day precipitation sums for winter and summer of the corrected reference scenario match the observed CHR data better than the uncorrected scenario does. Even very small exceedance probabilities match well with the corrected reference scenario. The spatial difference between the corrected and observed temperature data has decreased a lot. The average monthly temperature of the corrected reference temperature almost matches the observed average temperature. Although the correction for temperature only applies to the mean and the standard deviation, it also seems to have a positive effect on the lag-1 autocorrelation.

## 5.2 Recommendations

The bias correction is determined for daily subbasin values. The VIC model is run for grid cells with a resolution of 0.05 x 0.05 degrees at a 3-hourly timestep. Therefore, it might not be fully correct to correct the 3-hourly gridcell data based on a bias determined for daily subbasin values. Unfortunately there was no observed data available on this temporal and spatial scale. Further research concerning this matter is recommended. Unfortunately precipitation and temperature data is not the only input the VIC model needs. The VIC model also needs wind speed, incoming short- and longwave radiation, vapour pressure and specific humidity. There are no measured data available for these parameters. Therefore these parameters are left uncorrected. Further research concerning this topic is recommended.

The bias correction can now be used to correct the daily precipitation and temperature data. The corrected ERA15d precipitation and temperature data will be used to re-calibrate the VIC model. The determined bias correction of the reference scenario will be used to correct the precipitation and temperature data of this scenario, but will also be used to correct the precipitation and temperature data of the three IPCC emission scenarios B1, A1B and A2. Otherwise comparison of the hydrological impact between the future and present day climate remains impossible.

# References

Beersma, J. (2002), Rainfall generator for the rhine basin description of 1000-year simulations, *KNMI Publication*, *186-V*.

Hurkmans, R., H. de Moel, J. Aerts, and P. Troch (2008), Water balance versus land surface model in the simulation of rhine river discharges, *Water Resources Research*, *44*, doi:10.1029/2007WR006168.

Leander, R., and T. Buishand (2007), Resampling of regional climate model output for the simulation of extreme river flows, *Journal of Hydrology*, *332*, 487–496, doi:10.1016/j.jhydrol.2006.08.006.