Validation of Bias Correction Methods for RCM Scenarios for Hydrological Impact Modelling

Master-Arbeit

im Master-Studiengang (M.Sc.) **Climate Physics**

der Mathematisch-Naturwissenschaftlichen Fakultät

der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Nadine Mengis

908102

Erstgutachter: Prof. Dr. Douglas Maraun

Zweitgutachter: Prof. Dr. Katja Matthes

Kiel, am 24. Juli 2013

GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel

C A U Christian-Albrechts-Universität zu Kiel

Statkraft

Abstract

Hydrological impact modellers need bias corrected RCM time series of precipitation and temperature to assess future changes in various hydrological parameters. The objective of this thesis was a thorough validation of bias correction methods in order to bring out problems or advantages of these methods.

Therefore, two bias correction methods were applied within a perfect boundary setting in order to bias correct RCM precipitation data: One deterministic approach, the quantile mapping, and one probabilistic approach, the VGLM mixture model. These were applied to winter and summer time series and validated within a reference period.

The quantile mapping bias correction showed good results in reconstructing the statistical parameters of the observed time series as well as the accumulated precipitation values. However, the representation of the temporal structure of dry days is not addressed by this bias correction method.

The VGLM mixture model bias correction method was recently developed by Wong et al. [2013] and time series simulations from this method are thoroughly validated for the first time within this thesis. The simulations were able to reconstruct the statistical shape of the observed precipitation time series, but showed a better correspondence for the summer time series, because the variance within this season is lower. For the representation of the temporal structure of dry days, an overrepresentation of short dry spells is evident, indicating a lack of temporal structure within the logistic model of the VGLM mixture model application.

In order to assess problems of the quantile mapping bias correction, when applied to future time series, the method was validated within a pseudo-reality setting. Temperature and precipitation time series were bias corrected and used to force a hydrological model. At first, large deviations in the bias corrected time series compared the chosen pseudo-reality were found for mean values within the validation period. However, these could be explained by the fact that trends within one RCM time series remain unchanged even after the bias correction was applied. In addition, if the bias correction is applied using monthly data subsets, the temporal structure of the runoff time series, an output of the hydrological model, is adjusted to fit the one of the pseudo-reality.

It was found that the future runoff trends are strongly dependent on the trend in precipitation time series, if however the precipitation trend is low, a strong increase in mean temperatures can result in a deviation from the RCM trend. Differences in the mean

accumulated runoff values compared to the precipitation values within the validation period can be explained by an increased rate of evapotranspiration due to higher temperatures.

Zusammenfassung

Um zukünftige Änderungen in hydrographischen Parametern ein zu schätzen bzw. vorhersagen zu können, benötigen hydrologische Modellierer fehlerkorrigierte Zeitreihen von Temperatur- und Niederschlagswerten. Ziel dieser Arbeit war es, durch eine fundierte Evaluierung von Korrektur Methoden Probleme und Vorteile dieser Methoden herauszuarbeiten. Zu diesem Zweck wurden zwei Korrektur Methoden auf RCM Daten angewendet, welche mit Reanalyse Daten getrieben wurden. Eine deterministische Methode, das Quantile Mapping, und eine probabilistische Methode, das VGLM Mixture Model, wurden auf Winter- und Sommer-Zeitreihen innerhalb eines Referenz Zeitraums von 1960 bis 1999 angewendet und validiert. Das Quantile Mapping zeigte gute Ergebnisse, für die Rekonstruktion statistischer Parameter der beobachteten Zeitreihen, sowie für die Korrektur akkumulierter Niederschläge. Die zeitliche Struktur der Trockenperiode wird jedoch durch diese Methode nicht korrigiert.

Das VGLM Mixture Model zur Fehlerkorrektur wurde vor kurzem von Wong et al. [2013] entwickelt. Eine fundierte Validierung von simulierten Zeitreihen als Resultat dieser Korrektur Methode wird in dieser Studie zum ersten Mal durch geführt. Die statistischen Eigenschaften der beobachteten Zeitreihen wurden von den Simulationen zufriedenstellend rekonstruiert, es konnten jedoch bessere Ergebnisse für Sommer-Zeitreihen gefunden werden, da die Varianz in dieser Zeitreihe geringer ist. Bei der Repräsentation von Trockenperioden kann man eine Überrepräsentation kurzer Trockenperioden feststellen. Dies deutet darauf hin, dass das logistische Model des VGLM Mixture Models eine zu geringe zeitliche Abhängigkeit aufweist.

Für den Fall dass die Quantile Mapping Methode auf zukünftige Zeitreihen angewendet wird, war es wichtig mögliche Fehlerquellen zu identifizieren. Hierfür wurde das System in einer Pseudo-Realität getestet. Temperatur und Niederschlag Zeitreihen wurden fehlerkorrigiert und verwendet um ein hydrologisches Model an zu treiben. Bei der Auswertung konnten zunächst große Abweichungen in den Mittelwerten gefunden werden, wenn die Zeitreihen in dem Validierungszeitraum mit den Mittelwerten der Pseudo-Realitäten verglichen wurden. Die Ursache für diese Abweichungen konnte in den unveränderten Trends der RCM Zeitreihen gefunden werden: Durch das Quantile Mapping wird keine Veränderung in dem ursprünglichen Trend der Zeitreihe durchgeführt.

Darüber hinaus konnte festgestellt werden, dass falls die Korrektur auf monatliche Teil-Datensätze angewendet wird, die zeitliche Struktur des Wasserabflusses, ein Ausgabe Parameter des hydrologischen Models, sich der Struktur der verwendeten Pseudo-Realität

angepasste. Ebenso konnte gezeigt werden, dass Wasserabfluss Trends stark von den Trends im Niederschlag abhängig sind. Bei geringen Änderungen im Niederschlag kann ein starker Temperaturanstieg jedoch Änderungen des Wasserabfluss Trends hervorrufen. Die Differenzen der akkumulierten Abfluss Mittelwerte zu dem akkumulierten Niederschlag im untersuchten Validierungszeitraum können durch eine erhöhte Evapotranspiration aufgrund erhöhter Temperaturwerte erklärt werden.

Directory

1. Introduction

Impact studies, such as hydrological modelling, need a realistic assessment of the future climate change for a defined area on a local scale. With temperature and precipitation time series, hydrological impact models are driven to assess the possible changes in catchment runoff, evapotranspiration, soil moisture and other parameters. These are needed for example to assess the future rentability of catchments or the dimensions of the dams and whether they have to be reinforced to cope with future water amounts. In order to obtain future precipitation and temperature time series numerical model output has to be used.

General circulation models (GCMs) are essential to assess the large scale changes of the Earth‘s climate, however they are limited due to their coarse resolution of 100 km [Christensen et al., 2007] in representing local scale climate change. Since for most hydrological models the needed parameters are temperature and precipitation [Bronstert et al., 2007] and especially the latter occurs on a much finer scale, a higher model resolution has to be achieved to improve the representation of precipitation, with its high spacial and temporal variability [Maraun et al., 2010].

For regions of interest, such as Europe, models which are nested into a GCM, with a higher resolution and hence an improved representation of the atmospheric circulation are available [Christensen et al., 2007]: Dynamically downscaled regional climate models (RCMs). The RCM output has a much higher resolution, 50 km up to 10 km [Themeßl et al., 2010], but is still biased and needs to be corrected before used to force hydrological models [Christensen et al., 2008].

Another issue, not solved by using RCMs, is the scale discrepancy between the e.g. 25 km resolution of the RCM and the local scale on which precipitation is observed. This implies that bias correction methods do not only have to correct, but also to further downscale the RCM time series [Maraun et al., 2010]. The bias correction must deal with unexplained variability present in the local scale time series, but missing in the simulated time series on the grid box scale of the RCM [Maraun, 2013].

Until recently, only deterministic bias correction approaches were used, where the result of the correction is one single bias corrected RCM time series. For such approaches, the correction does not distinguish between explained variance simulated by the RCM on a grid box level and the unexplained variance on a small scale, that should be modelled as random noise [Wong et al., 2013]. Therefore, the bias correction will inflate the explained variance in the RCM time series to match the observed local variability, which might alter

other systematic features e.g. long term trends [Maraun, 2013]. This becomes especially important if the unexplained variability is as large as it is for precipitation [Wong et al., 2013].

In contrast to any deterministic bias correction, a probabilistic approach deals with unexplained variance in the time series [Wong et al., 2013]. Probabilistic approaches give a probabilistic distribution of precipitation intensities for each day. When sampled from these daily distributions in order to obtain a single bias corrected time series, this allows for unexplained variance. Thus a downscaling step is included in the bias correction [Wong et al., 2013]. Within this thesis the probabilistic bias correction approach is applied and validated for the first time.

Until recently, in order to run hydrological models deterministically bias corrected temperature and precipitation time series were used. To obtain time series used in the impact models, various approaches are applied by hydrologists.

One basic approach, used by e.g. Statkraft, is the delta change method: The difference (ratio) between the means of the future and the reference period of the RCM temperature (precipitation) time series is calculated and added (multiplied) to the observed temperature (precipitation) time series to gain a projected time series for a future period [E. Alfnes, personal communication; Yang et al., 2010]. This means the exact observed time series, with only a changed mean, is used to assess future changes in the Earth‘s climate.

One advantage of such an approach is that the trend present in the RCM time series is unaltered. This is a very important issue, when it comes to future predictions. However, problems of the delta change method are the fact, that no statistical moment other than the mean is corrected, hence the differences in the variance of the time series are disregarded. Furthermore by the application of this approach possible changes in circulation patterns, that would be present in a RCM future time series, are not assessed.

Other deterministic bias correction methods establish statistical relationships between the RCM and the observed time series, in order to correct RCM biases. This way not only the mean, but the overall statistical shape of the given variable is corrected. For these approaches a probability distribution has to be chosen, which is usually the normal distribution for temperature and the gamma distribution for precipitation [Yang et al., 2010]. These distributions are partly not precise enough to describe the complex nature of precipitation in particular. Therefore, more complex probability distributions have to be implemented. This will be discussed in further detail in Section 3.1.2.

In order to circumvent the problem connected to choosing a distribution the quantile mapping bias correction approach is not assuming a distribution, but corrects single

quantiles of the distribution empirically. For details of the quantile mapping bias correction method and its application see Section 3.1.1. The fact that the statistical shape of the parameter in the calibration period looks perfectly corrected, are reasons why this method is frequently used by hydrologists [Gudmundson et al., 2012; Teutschbein and Seibert, 2012; Themeßl et al., 2010].
For the probabilistic approach there are no references in the literature. The method was recently developed by Wong et al. [2013], where a validation of the approach is carried out based on the daily probability distribution. Whereas within this study, the bias corrected time series obtained from the model will be analysed. For details see Section 3.1.2.

This thesis consists of two parts: The first part focuses on the bias correction of precipitation time series within a reference period. In this context, one deterministic and one probabilistic approach is applied to winter and summer precipitation time series and validated against the correct representation of mean accumulated precipitation, the general shape of the precipitation intensities and dry spells. In addition to that, the question of how to best bias correct a catchment time series is addressed.
The second part of the thesis addresses the problems the deterministic approach might experience, when applied to predict future climate change. Precipitation and temperature time series are therefore bias corrected in a so called pseudo-reality setting. For more details see Section 3.2. Furthermore, the bias corrected time series are used to run the Nordic Hydrological Model (HBV, details see Section 3.3), to assess the impact climate change might have on the runoff of the catchment. The focus was set on runoff because it gives a good understanding of the dynamics of the catchment.
The effects of the applied bias correction on trends of the precipitation, temperature and the resulting runoff time series is investigated. Furthermore the annual cycles of the parameters, the dry days and the mean accumulated precipitation time series are investigated.
The objective of this thesis is to give hydrological impact modellers an improved understanding of the advantages and disadvantages of the two applied bias correction methods, through a profound validation.
In Section 2 the used data is explained in detail. The applied bias correction methods are described in Section 3, the results of the analysis are given in Section 4 and the conclusion which can be drawn as well as an outlook of what needs to be dealt with in future studies are given in Sections 5 and 6, respectively.

2. Data

2.1 Prevailing Climate of Norway and Reinsnosvatn

This work was carried out in cooperation with Statkraft, a norwegian hydropower supplier. The study area is therefore Norway, with a special focus on southern Norway.

Geographically, Norway lies on the same latitude bands like Northern Canada, Greenland and Iceland between 56 °N and 71 °N, but the temperatures are on average much higher. This is due to the influence of the North Atlantic Current, which transports warm surface water northeastward toward the European Continent [Met.no]. In addition, the prevailing southwesterly wind system transports warm and moist air across the shores to the inland regions, causing a more temperate climate in all of northern Europe, compared to Canada or Greenland.

Norway has an elongated shape, with a considerately long coast line and is dominated by a mountain ridge, dividing the country into a coastal and a more continental climate region, located to the west and the east of the mountain ridge, respectively, see Figure 1a).

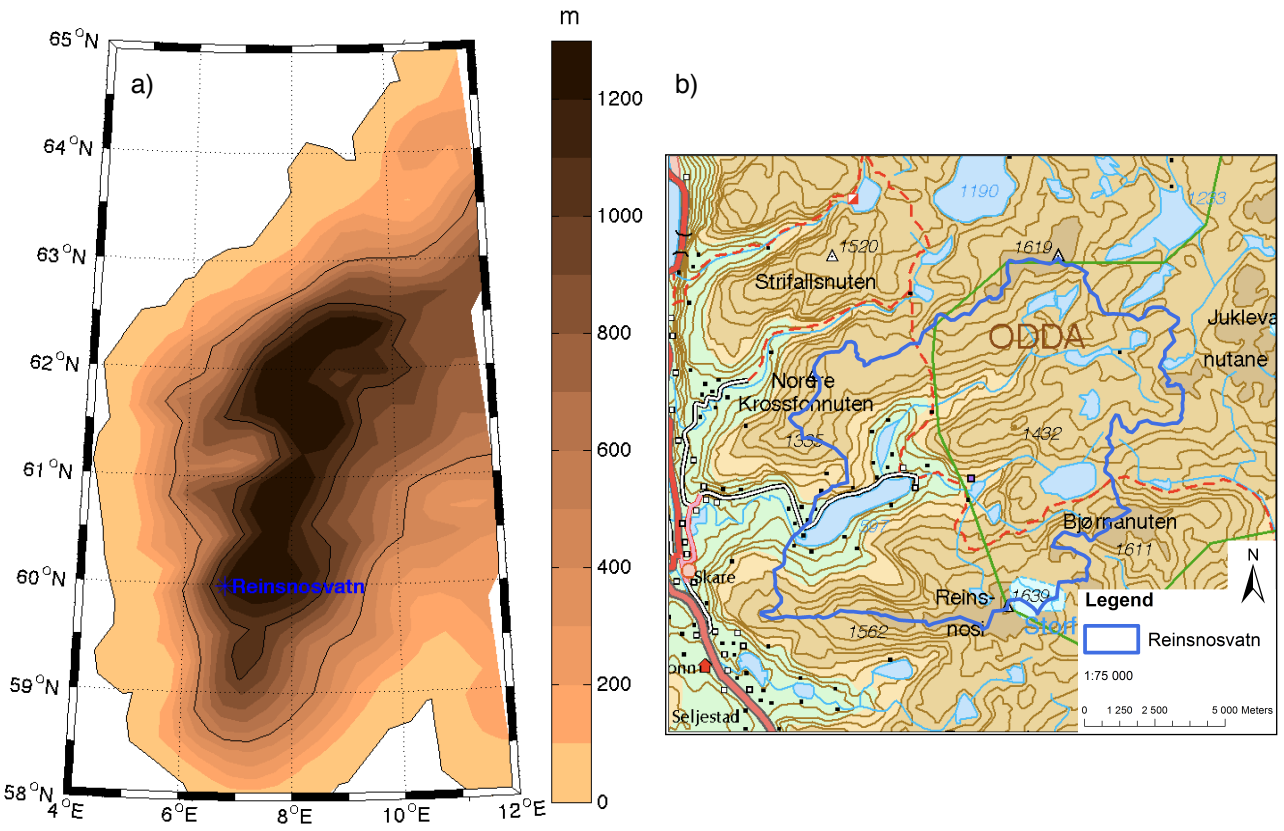


Figure 1: a) Map of southwestern part of Norway, with topographic height from the swedish RCM, the SMHI model. The blue star indicates the location of the catchment Reinsnosvatn.

b) A detailed map of the catchment Reinsnosvatn, with topography and boarders. Source: The Norwegian Mapping Authority - N500 Raster

The coastal climate regions experience heavy rainfalls when the moist air masses are forced to rise, approaching the mountains in the land‘s interior. Whereas, the more continental climate regions experience less rainfall and benefit from the temperate temperatures.

The catchment Reinsnosvatn, lies in the southwestern part of Norway, west of the mountain ridge, compare Figure 1a). It therefore experiences a coastal climate. Figure 1b) gives a detailed map of the catchment Reinsnosvatn, with borders indicated by the thick dark blue line. All rain falling within these borders is collected in the catchment. Furthermore the light blue lines show manmade connections between the single lakes to enlarge the size of the catchment.

The catchment Reinsnosvatn is selected for this study, and used as an exemplary catchment, for an area of high mean annual precipitation intensities, see Figure 2b). Figure 2 shows mean temperature and precipitation maps of the southwestern part of Norway.

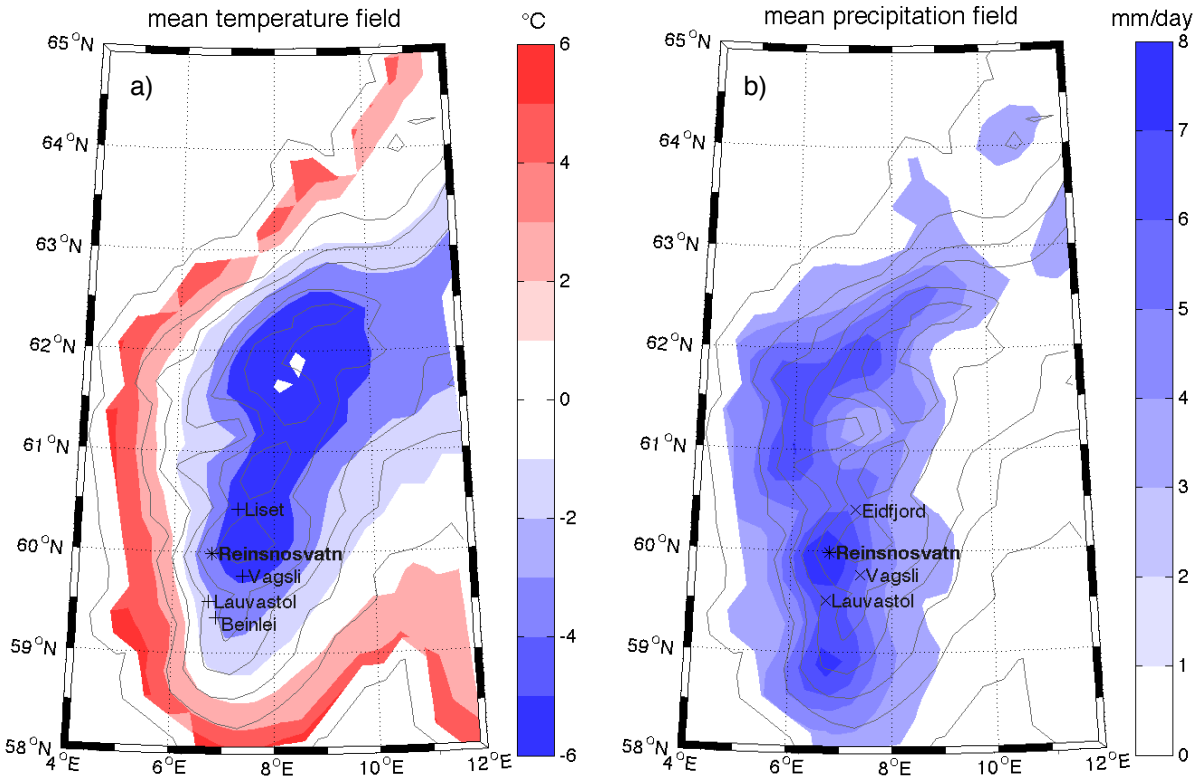


Figure 2: Map of a) mean surface temperature in °C and b) mean precipitation in mm/day in the southwestern part of Norway from the ERA40 forced swedish RCM, SMHI. The black pluses (crosses) indicate the location of the four (three) temperature (precipitation) measuring sites for the catchment Reinsnosvatn. The mean values are calculated in the time period from 1961 to 2000.

The location of the catchment is given, indicated by the star, as well as the location of the three precipitation stations (crosses) and the four temperature stations (pluses). A weighted average over the stations, a catchment time series, is used to force the Nordic Hydrological Model (HBV). For details on the model see Section 3.3.

The catchment has an area size of 121 km^2, of which 9.54% are covered by lake. As seen in Figure 2a), it is located in a region where mean annual surface temperatures are below zero. This is due to its high elevation ranging from 600 to 1620 m above sea level and explains why 1.24% of the catchment area is covered by glaciers. Figure 2b) shows the map of mean precipitation of the region, indicating that Reinsnosvatn and all its precipitation measuring sites lie within a band of high annual mean precipitation intensities. This makes the catchment an interesting site for the following investigations, addressing the question how well the bias correction methods deal with high precipitation intensities.

The three precipitation observational sites and the four temperature observational sites used to build the catchment area average time series are given in Table 1, additionally the coordinates, the elevation, and the weighting used by Statkraft for the catchment average are given.

a) Precipitation Station	Coordinates	Elevation	Weighting	Observed since
Dnmi-Fet.i.Eidfjord	60° 24.6 N 7° 16.8 E	735m	2,0	July 2005
Tokk-Vågsli	59° 46.2 N 7° 22.2 E	822m	3,0	December 1997
Ulla-Lauvastøl	59° 30.6 N 6° 42.6 E	627m	5,0	October 1997

b) Temperature Station	Coordinates	Elevation	Weighting	Observed since
Sima-Liset	60° 25.5 N 7° 16.5 E	748m	4,0	October 1997
Tokk-Vågsli	59° 46.2 N 7° 22.2 E	822m	1,0	December 1997
Ulla-Beinlei	59° 20.8 N 6° 51.1 E	1070m	1,0	October 1997
Ulla-Lauvastøl	59° 30.6 N 6° 42.6 E	627m	4,0	October 1997

Table 1: General information about the used a) precipitation and b) temperature time series. Given are the coordinates, the elevation, the weighting factor and the observational extend of the station time series.

Figure 3 shows the seasonal cycles of temperature and precipitation of the measuring site Vågsli, for geographical location see Figure 2. The precipitation intensities are highest during January, with average intensities of 5 mm/day, and lowest in the end of April, with average values below 2 mm/day. The seasonal cycle of temperature has its peak with an average of 12 °C in August, and its lowest values with -10 °C in January.

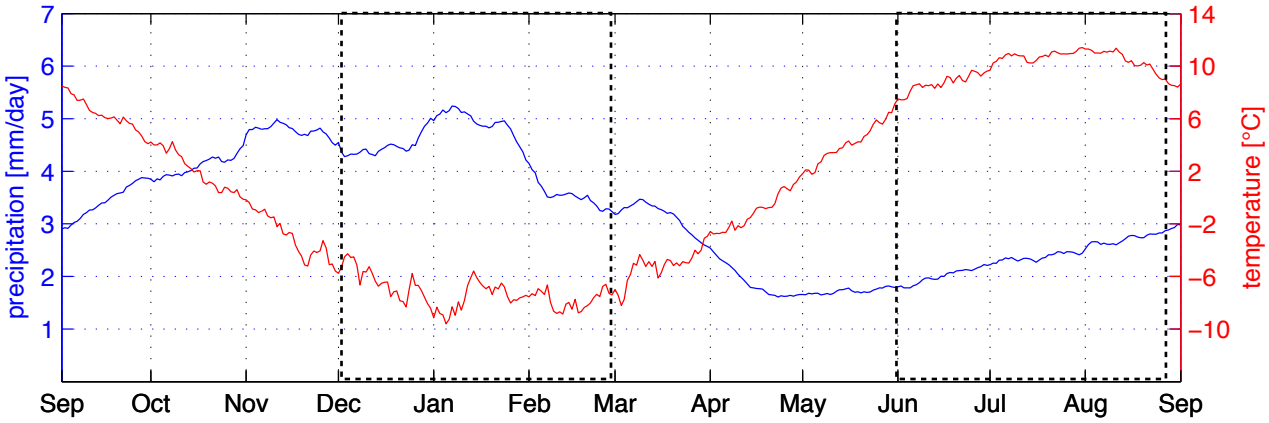


Figure 3: Average daily seasonal cycles of observed precipitation in mm/day (blue) and temperature in °C (red) for the observation site of Vågsli in a period from 1st September 1961 to 1st September 2010. The precipitation time series was additionally smoothed applying a 30 days running mean. The dashed black boxes correspond to the chosen winter and summer time periods.

Since the winter and summer seasons are characterised by cold/wet and warm/dry climate conditions, see Figure 3, they are of particular interest to be analysed throughout this thesis. The black boxes in Figure 3 indicate the chosen time spans of the years.

2.2 Data for the Reference Period

The first part of the analysis concentrates on precipitation only, and is carried out for the reference period from 1st September 1961 to 1st September 1999. The 1st September corresponds to the beginning of the hydrological year as represented in the hydrological model used by Statkraft. The winter season of December, January and February, and the summer season of June, July and August, of each year are selected and bias corrected.

The observed daily precipitation sums of the stations are provided by Statkraft and the Norwegian Meteorological Institute, met.no. The station-based time series are not observed from 1961 onward, see Table 1, hence they were extended by Statkraft using linear regression with surrounding stations that date back until 1961 [personal communication E. Alfnes]. It was tested if the variance within the time series did change due to this extension and it was found that this is not the case. The three station-based precipitation time series are used to calculate a catchment time series using the corresponding weight factors, which was treated as an additional area time series. These observational time series will be used to bias correct the RCM time series.

The Regional Climate Model (RCM) output from the Swedish Meteorological and Hydrological Institute (SMHI) forced by NCEP/NCAR reanalysis data is used. It has been shown that this RCM performs well in reconstructing the weather patterns for Scandinavia [Landgren et al., 2012]. Any biases arising from a GCM forcing are avoided by using a reanalysis data forced RCM run, this is called a perfect boundary condition setting [Maraun, 2012]. The simulated daily precipitation sums, needed for the analysis, are provided by the ENSEMBLES project [http://ensemblesrt3.dmi.dk/], with a horizontal grid resolution of 25 km.

Station-based time series are calculated using inverse distance weighting of the four nearest grid points of the RCM data set. Additionally, a catchment-based time series was calculated from the four nearest grid points using the coordinates of the catchment, following Engen-Skaugen [2007].

2.3 Data for the Pseudo-reality

The second part of the analysis is carried out for the period from 1st September 1980 until 31st August 2099. Temperature and precipitation time series of the whole year are considered and bias corrected to run the Nordic Hydrological Model (HBV, details see Section 3.3).

For the pseudo-reality setting no observational data is needed, since always one RCM time series is considered as the observed time series and is then called pseudo-reality. Therefore, both temperature and precipitation data sets are taken from the RCM output of five different institutes, listed in Table 2. Information given is the driving GCM, the name of the RCM itself, the acronym, that will be used in the analyses from now on and the time period for which the data is available.

Institute / Contact	Driving GCM	RCM	Acronym	Period
DMI (Denmark Meteorological Institute) Ole B. Christensen	ECHAM5-r3	DMI-HIRHAM5	DMI-HIRHAM5	1951-2099
ICTP (International Centre for Theoretical Physics) Filippo Giorgi	ECHAM5-r3	RegCM	ICTP-REGCM3	1951-2100
KNMI (Royal Netherlands Meteorological Institute) Erik van Meijgaard	ECHAM5-r3	RACMO	KNMI-RACMO2	1950-2100
MPI (German Max-Plank-Institute) Daniela Jacob	ECHAM5-r3	REMO	MPI-M-REMO	1951-2100
SMHI (Swedish Meteorological and Hydrological Institute) Erik Kjellström	ECHAM5-r3	RCA	SMHIRCA	1951-2100

Table 2: Information about the used RCM runs for the pseudo-reality setting. Given are the driving GCM, the name of the RCM and the acronym, which will be used from now on in the analysis. In addition the time span for which the data is available is given. For detailed information about the models, see http://ensemblesrt3.dmi.dk/.

The driving GCM in all cases is the ECHAM5-r3 (developed by the German Max Plank Institute in Hamburg, as a modification of the models from the European Centre for Medium-Range Weather Forecasts, ECMWF), which follows the A1B future climate scenario.

All five RCM runs are provided by the ENSEMBLES project [http://ensemblesrt3.dmi.dk/] with a 25 km horizontal grid resolution.

Using the same GCM to drive different RCMs allows to discriminate the bias introduced by the GCM when comparing the RCM outputs, this corresponds to a perfect boundary condition setting [Maraun, 2013].

For both, temperature and precipitation time series, the station-based and catchment-based time series are calculated using inverse distance weighting of the four nearest grid points in the RCM data sets, following Engen-Skaugen [2007].

3. Bias Correction Methods and the Hydrological Model

3.1 Analyses within the Reference Period

In the reference period, the deterministic and the probabilistic bias correction approaches are applied to the same observed and RCM data sets and their performance is validated. This is all carried out in a perfect boundary condition setting using reanalyses driven RCM data. It is therefore expected, that the bias in the uncorrected RCM data arises only from the RCM and is low compared to GCM driven RCM output. At first the different methods are explained.

3.1.1 Quantile Mapping

The bias correction (BC) based on quantile mapping (QM) is recommended by various scientists, among them Gudmundsson et al. [2012] for Norway, Themeßl et al. [2010] for Austria and Teutschbein et al. [2012] for Sweden. Therefore this non-parametric method was chosen for this study as a deterministic approach to bias correct winter and summer precipitation time series.

Before applying the quantile mapping bias correction, a dry day correction was carried out. This forces the RCM time series to have the same probability of a dry day like the observations. The dry day threshold applied to the observational time series was chosen to be 1 mm/day following the approaches of Piani et al. [2010], Themeßl et al. [2010] and Wong et al. [2013]. From the given threshold the exact probability of an observed dry day could be estimated. The dry day threshold for the RCM time series was chosen accordingly in such a way, that the probability of a dry day in the RCM time series is equal to the observed probability of a dry day (Details on the effect on the dry day correction on the annual precipitation sums and the applied thresholds for the RCM station-based time series are given in the appendix).

For the quantile mapping bias correction the empirical quantiles of the time series from the observational, P_{obs}, and RCM time series, P_{rcm}, were calculated. The correction was then applied using the following equation (1):

$$P_{corr} = F_{obs}^{-1}\left(F_{rcm}\left(P_{rcm}\right)\right) \quad (1)$$

where $F_{rcm}(.)$ is the cumulative distribution function of the RCM time series and $F_{obs}^{-1}(.)$ is the inverse cumulative distribution function of the observational time series [Gudmundsson et al., 2012]. As an example, Figure 4 shows the concept for the measuring site Eidfjord, where a) shows the sorted uncorrected RCM winter precipitation intensities against the

sorted observed winter precipitation values. A negative bias of the uncorrected RCM time series is evident, giving too low precipitation intensities of about 17% for an observed intensity of 60 mm/day, compared to the observed ones. Figure 4b) shows how the concept of the correction itself works and c) shows the corrected RCM time series against the observed one.

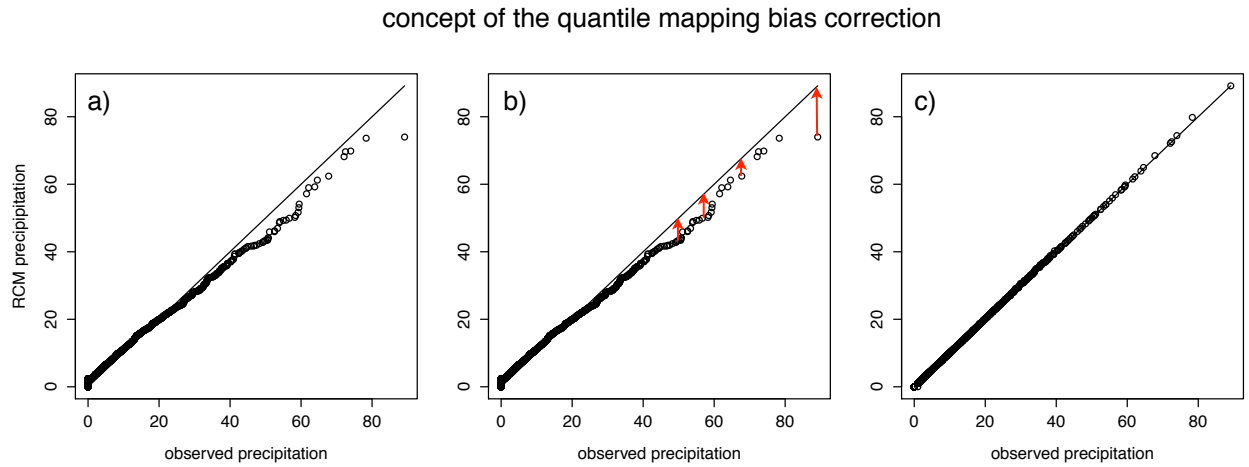


Figure 4: Sorted winter precipitation time series from the RCM against the sorted observed time series for Eidfjord: a) shows the uncorrected RCM values, b) explains the quantile mapping bias correction approach and c) shows the sorted time series after the bias correction was applied without cross validation.

The correction is replacing the RCMs precipitation intensity by the corresponding quantile of the observed precipitation time series. This is only applicable as long as the simulated precipitation intensities are within the observed range. For RCM precipitation values outside the range of the quantiles, calculated from observations, a linear fit for the upper five percentiles is estimated and used to extrapolate. This is especially important for future time series, which are likely to experience more extreme weather events [Wong et al., 2013].

For the reference period the time series are bias corrected applying a cross validation for a higher robustness of the analysis [Maraun et al., 2010]. Therefore, 29 seasons of the data set are used to fit the bias correction, which is then applied to the remaining 9 seasons of the data set.

3.1.2 VGLM Mixture Model - Implementation

When deterministic approaches are used to bias correct gridded RCM data with local scale time series there is one major disadvantage: No distinction between the explained variance simulated by the RCM on the grid box scale and the unexplained variance present in the local scale observed data is made [Maraun, 2013]. Probabilistic approaches

give a probabilistic distribution of precipitation intensities for each day or event, which allows for unexplained variance when sampling from the daily distributions in order to obtain a single time series [Wong et al., 2013].

The probabilistic bias correction method in this study was recently developed by Wong et al. [2013]. The method intends „to separate the explained variance from the total local scale variance and explicitly model the unexplained small-scale variance.“ [Wong et al., 2013]. In order to achieve this, a regression model is chosen and fitted to bias correct and downscale. Here, the observed precipitation is the predictand and the simulated precipitation is the predictor. The regression model chosen by Wong et al. [2013], is the Vector Generalized Linear Model (VGLM), because it is useful when a non-linear relationship between the dependent and independent variable has to be modelled [Wong et al., 2013]. This model framework allows the probabilistic distribution to vary temporally according to the predictor, in this case the simulated precipitation [Yee and Stephenson, 2007].

The bulk of the precipitation distribution is well reconstructed using a gamma distribution, but its tail is too light to model the extreme precipitation events. Hence an extreme value distribution is needed above a certain threshold. Therefore, the general shape of the probability distribution is constructed by combining a gamma distribution with a Generalized Pareto (GP) distribution, which is an extreme value distribution for the tail. The stationary model combining these two distributions was introduced by Vrac and Naveau [2007], the variant used here is from Frigessi et al. [2003] given in equation (2):

$$l_\beta(r) = c(\boldsymbol{\beta}) \times \{[(1 - w_{m,\tau}(r)) \times f_{\lambda,\gamma}(r)] + [w(r) \times g_{\xi,\sigma}]\}, \qquad \boldsymbol{\beta} = (\lambda, \gamma, \xi, \sigma, m, \tau) \tag{2}$$

where *r* is the observed precipitation, $f_{\lambda,\gamma}$ is the gamma distribution given in equation (3):

$$f_{\lambda,\gamma}(r) = \frac{1}{\lambda^\gamma \Gamma(\gamma)} r^{\gamma-1} \exp\left(\frac{-r}{\gamma}\right), \qquad \lambda, \gamma > 0 \tag{3}$$

with scale parameter, λ, and shape parameter, γ.

Furthermore $g_{\xi,\sigma}$ is the GP distribution explained in equation (4):

$$g_{\xi,\sigma}(r) = \frac{1}{\sigma}\left(1 + \frac{\xi(r-u)}{\sigma}\right)^{(-\frac{1}{\xi}-1)} \quad \text{when} \quad \xi \geq 0, x \geq u \tag{4}$$

with scale parameter, σ, and shape parameter, ξ. The weight function $w_{m,\tau}$ describes the transition between the two distributions and is given in equation (5):

$$w_{m,\tau}(r) = \frac{1}{2} + \frac{1}{\pi}\arctan\left(\frac{r-m}{\tau}\right), \qquad m, \tau > 0 \tag{5}$$

with the centre of the transition, m, and the rapidity of the transition, τ [Wong et al., 2013]. Hence, the mixture probability distribution, described by a set of six parameters, varies temporally for each day depending on the simulated precipitation. For these calculations only precipitation values greater than the dry day threshold of 1 mm/day were chosen. In order to circumvent problems with the gap of values between 0 and 1 mm/day, when fitting a gamma distribution, a value of 1 mm/day was subtracted from all values equal and larger than 1 mm/day. This value is added back again later, when values are sampled for the simulations.

Following the above preprocessing, the six parameters were calculated and fitted to the data, where the goodness-of-fit was rated according to the Akaike Information Criterion (AIC) [Vrac and Naveau, 2007; Akaike, 1974]. Based on the AIC, each station required a different model structure for each season (Information on the AIC and the results of the fitting analysis are given in the Appendix).

In addition to the mixture probability distribution, a logistic regression model is applied to model the probability of rainfall occurrence. The logistic regression model expresses a function of expectation $\mu = (\mu_1,...,\mu_n)^T$ of a random variable Y as a linear combination of a set of predictors x [McCullagh and Nelder 1989]. Its general form is given in equation (6):

$$g(\mu) = x\alpha, \tag{6}$$

where $g(.)$ is a monotonic function known as the link function, and $\alpha = (\alpha_1,...,\alpha_p)^T$ is a vector of coefficients. The probability p_i that a day i is wet is then modelled as a function of the simulated precipitation from the RCM, x, given in equation (7):

$$g(p_i) = \log\left(\frac{p_i}{1-p_i}\right) = x\alpha, \tag{7}$$

where $g(.)$ is the so-called “logit” link function and α is a vector of coefficients to be estimated. Hence, the probability p_i of a day i being wet can be expressed as:

$$p_i = \frac{\exp(x\alpha)}{1+\exp(x\alpha)} \tag{8}$$

For each day, there is a new predictor x and a new probability can be deduced from equation (8). From this the rainfall occurrence for each day can be simulated (wet or dry). If the day is dry, precipitation intensity is set to 0. Otherwise, the probability distribution of the

daily precipitation intensity is calculated and positive values are drawn from it [Wong et al., 2013].

Since hydrologists are interested in single bias corrected time series after the VGLM mixture model was fitted to the data, 100 simulations were carried out and will be analysed in the following.

As an example, Figure 5 shows the sorted values for chosen probabilities from the mixture distribution model against the sorted RCM precipitation values for the summer time series of Eidfjord. It shows the distribution for each of the RCM values, from which values are randomly drawn when simulating a wet day. Additionally, the resulting sorted time series of the quantile mapping bias correction is shown (grey line).

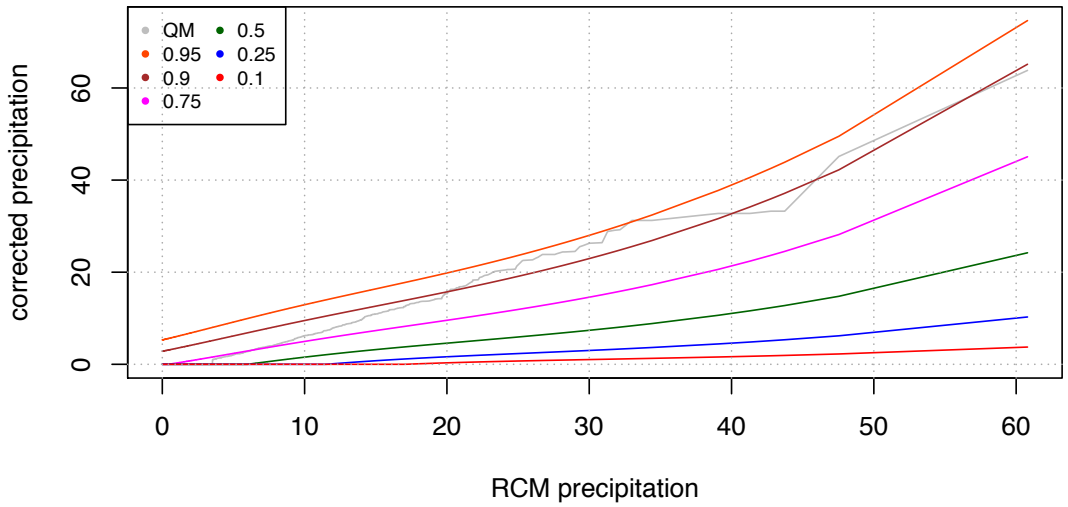


Figure 5: The probabilities of the Probabilistic Mixture Model against the swedish RCM, SMHI, precipitation intensities for the measuring site Eidfjord in summer in mm/day. The coloured lines give the values for the different probabilities which are calculated from the observed and RCM time series. The grey line gives the resulting time series from the quantile mapping bias correction.

For the sorted values of the 0.5 probability, a corrected precipitation value of zero is seen for simulated precipitation values up to 7 mm/day. Additionally, for highest simulated values of 60 mm/day the probability to model values beneath 25 mm/day is 50%. This shows that only a small part of the variance in the observed time series can be explained by the variance within the RCM time series.

However, when we assume that the quantile mapping bias corrected time series is close to the observed values, the range of the corrected precipitation values is reasonable and increases with increasing simulated precipitation intensities.

3.1.3 VGLM Mixture Model - Trouble shooting

When the method was first applied, the model structure that fitted a data set from the UK was used, with the intention of implementing the method for other areas easily. In addition a dry day threshold of 0.1 mm/day was used for the Norwegian data set. It was found that this did not satisfactory correct the data set for Norway, even after changing the dry day threshold to 1 mm/day, which has been used for the UK data set. Thereafter, it was decided, that the best model structure for the Norwegian data set had to be chosen using the AIC. This step of choosing the appropriate model structure for a given data set is recommended for whatever implementation.

Still after fitting the model to the data, the achieved correction was not satisfying. Parallel to my analysis the model was validated by Wong et al. [2013] and an error in the implementation was found, partly with the help of these analyses. It was found that the gamma distribution started fitting the distribution from values of 0 mm/day on, in the initial analysis, although they were excluded from the analysis. The code was corrected and the model structure was refitted. The correction achieved afterwards is analysed and discussed within this thesis.

3.2 Analyses within the Pseudo-reality Setting

In the second part of the analysis, possible problems of the quantile mapping bias correction will be addressed within a pseudo-reality setting, following Maraun [2012] and Räisänen and Räty [2012]. Therefore, 5 different RCM runs all driven by the same GCM are used, where one of the RCM runs is taken as a pseudo-reality and the remaining four RCM runs were bias corrected to the pseudo-reality within a calibration period.

For both precipitation and temperature, at first the correction was calculated within the calibration period, from 1st September 1980 until 1st September 2010. This would be a realistic calibration period, if the bias correction was fitted to an observed data set, instead of a pseudo-reality. After the bias correction was calculated within the calibration period, it was applied to the complete time series until 2099.

The bias correction was applied to monthly data sets, but it was fitted including the previous and following month: For example in order to fit the correction for January, December and February were included. This increases the amount of data used for the fitting and hence the robustness of the results.

For the precipitation time series the quantile mapping bias correction was chosen and applied as explained in Section 3.1.1.

Temperature values were bias corrected using the method following Piani et al. [2010]. This method was selected because it has shown good results in earlier studies. A linear regression to the differences in the distribution of the RCM temperature and the observed temperature was fitted against the RCM temperature values. This fit is used to correct the RCM temperature values.

In order to run the hydrological model of Statkraft the bias corrected time series of the different stations were averaged to calculated catchment area time series using the weighting factors given in Table 1.

3.3 The concept of the Nordic Hydrological Model - HBV

The Nordic Hydrological Model (HBV, Hydrologiska Byråns Vattenbalansavdelning) uses daily temperature and precipitation time series to calculate/model various hydrological parameters, such as glacier melt, snow coverage and volume, soil moisture, evapotranspiration, ground water content and finally runoff as a direct or indirect function of all the parameters. The model structure described here is the one used by Statkraft and within this thesis, but it is based on the model of Sælthun [1995].

Figure 6 shows a basic representation of the model structure as described by Lindström et al. [1997]. Input to the model are firstly parameters for the basic catchment structure, such as height distribution, lake and glacier percentage and an initial snow distribution. For details about the catchment Reinsnosvatn, see Section 2.1.

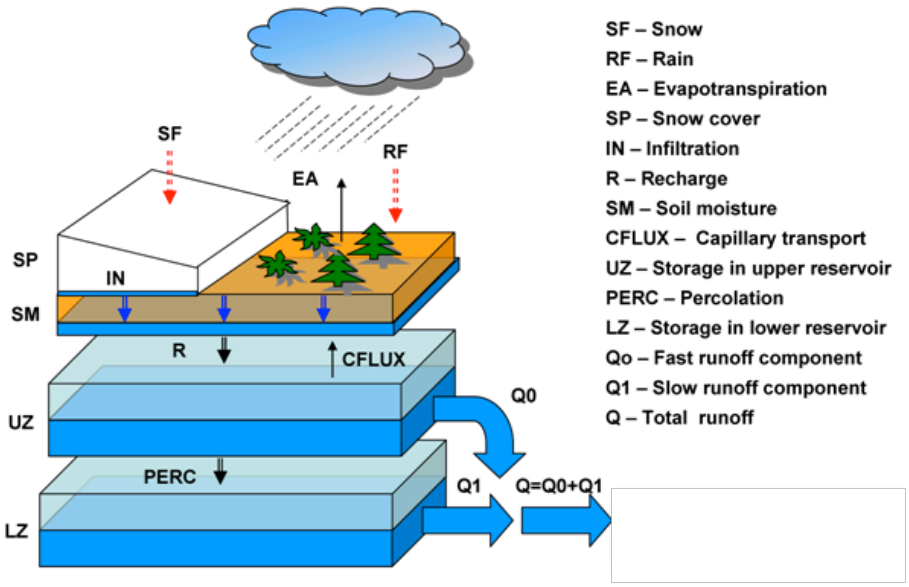


Figure 6: Schematic representation of the HBV-96 model with routines for snow, soil, and runoff response. Detailed description of the model can be found, see, e.g. Lindström et al., 1997.
Source: http://hikm.ihe.nl/floodsite/data/Task20/hbv.html

With the input of an area value for the catchment at first the snow sub-routine is driven to determine what kind of precipitation is falling, snow (SF) or rain (RF). Additionally, glacier/ snow melt/freezing and snow coverage (SP) are calculated and given as input to the next subroutine, the soil moisture zone. The soil moisture (SM) sub-routine calculates the evapotranspiration (EA) as a function of temperature and snow/glacier free area. Input to the soil moisture zone is not only melt from snow or glacier, but also rain on snow free areas and lakes. This unsaturated zone has no component to the runoff, but regulates the input to the upper ground water zone (UZ). The outflow (Q0) of this reservoir is dependent

on the amount of water within the zone, plus a constant percolation (PERC) to the lower ground water zone (LZ). The dynamical response sub-routine represents the lower ground water zone, input to this zone is the percolation from the upper ground water zone and the outflow (Q1) depends again on the amount of water within the zone. The total outflow, the runoff, (Q), is the sum of the fast runoff from the upper ground water zone and the slow runoff from the lower ground water zone.

For a detailed explanation of the HBV model see Sælthun [1995] or Lindström et al. [1997].

4. Results

4.1 Reference Period

For the reference period the corrected precipitation time series of the deterministic approach, the quantile mapping (QM) bias correction, and of the probabilistic approach, the vector generalised linear mixture model (VGLM.MM) bias correction, are compared to evaluate their performance in reconstructing certain features of the observed precipitation time series. Of particular interest are the representation of the statistical shape of the observed precipitation intensities, the mean accumulated precipitation and the dry spells.

4.1.1 Precipitation Intensities Distribution

The representation of the statistical shape of the observed precipitation intensities is examined by looking at histograms of the precipitation intensities from 1-2 mm/day to the 98th percentile of the corresponding observed time series. Shown in Figure 7 are the histograms of the dry day corrected observed, the uncorrected, the QM bias corrected and the VGLM.MM bias corrected precipitation time series. Based on the 100 simulations of the VGLM.MM bias correction the spread of the distribution is given using a box plot.

In particular the shape of the distribution of the lower precipitation intensities can be discussed from these figures.

The bias present in the uncorrected RCM data of the SMHI model is discussed first: In winter, Figure 7a), a positive bias in the amount of occurrences for the station Eidfjord is evident for precipitation intensities lower than 13 mm/day. This indicates an overrepresentation of small precipitation intensities for the uncorrected RCM data. The same feature is evident for Vågsli, however for this station there is an overall overrepresentation in all of the precipitation intensities up to 25 mm/day. The time series of Lauvastøl and Reinsnosvatn show a good representation of the statistical shape of the precipitation intensities in the uncorrected RCM time series, which can be attributed to the fact that the RCM is driven by reanalysis data.

For summer, Figure 7b), an overestimation of precipitation occurrences especially for precipitation intensities below 20 mm/day is present in all stations. Again the station Vågsli shows the largest bias in the uncorrected RCM time series for all precipitation intensities shown.

These are the biases of the uncorrected RCM precipitation values that have to be addressed by the bias correction methods.

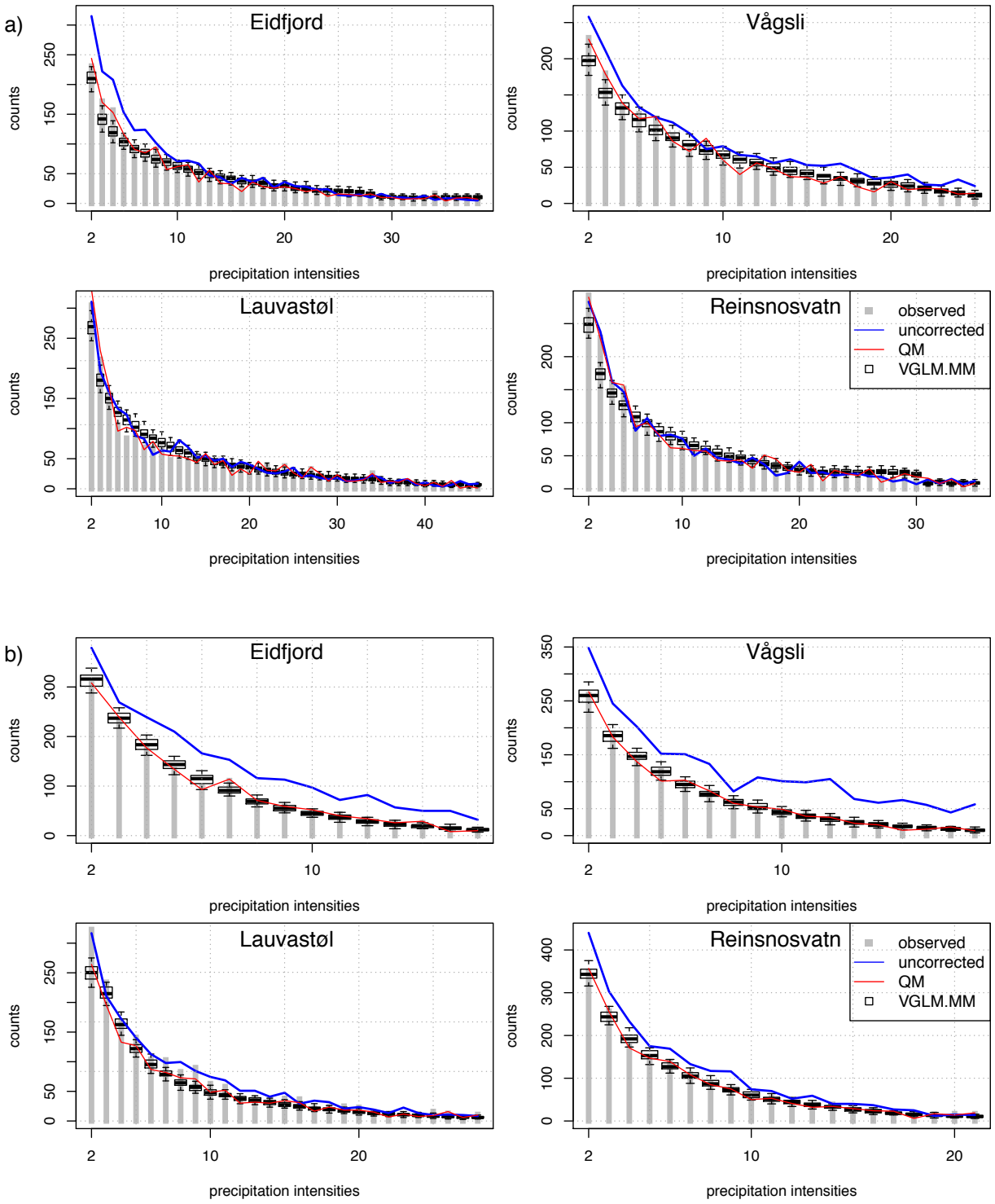


Figure 7: Histograms of a) winter and b) summer precipitation intensities starting from 1-2 mm/day to the 98th percentile of the observations. Shown are the three precipitation stations and the catchment time series of the observed (grey bars), the uncorrected SMHI model (blue line), the quantile mapping (QM) bias corrected (red line) and the VGLM mixture model (VGLM.MM) bias corrected (black boxes) time series. The box gives the median of the 100 simulations from the VGLM.MM bias correction as well as the upper and lower quartile (25 percentile and 75 percentile) and the 5th and 95th percentile.

The QM bias corrected time series show for both, winter and summer, a good correspondence to the observed histograms of precipitation intensities. This is expected given how the time series are constructed, compare Figure 4.

For the VGLM.MM bias corrected precipitation intensities in winter, shown in Figure 7a), the shape of the median values shows a good correspondence to the observed shape of the precipitation intensities: In general, the distributions are more continuous than the observed precipitation histograms, as it is expected from the construction of the time series, which follow a gamma distribution. For intensities between 1-2 and 2-3 mm/day for all the stations, the largest deviation from the shape of observed precipitation occurrences is found where the median of the 100 simulations underestimates the occurrence of these intensities by 20 to 50 counts. Even the 95th percentile calculated based on the VGLM.MM does not reach the high number of occurrences counted in the observed time series. In addition, for Lauvastøl, an overestimation of precipitation intensities occurrences is found for intensities between 4 to 13 mm/day. This bias might arise from the fact that for Lauvastøl the occurrences of the precipitation intensities of 1-2 mm/day is very high with over 300 counts. The simple gamma distribution has to cope with this feature and therefore the occurrences for the transition precipitation intensities are overestimated. Furthermore, the transition between the gamma and the Generalised Pareto distribution is visible, for the catchment Reinsnosvatn it can be seen in the area of 30 mm/day.

For summer, Figure 7b), the distributions of the 100 simulations from the VGLM.MM bias correction of precipitation intensities follow the shape of the observed histogram well. Compared to winter, the difference in the median of the VGLM.MM bias corrected and the observed time series for precipitation intensities between 1-2 mm/day is smaller. Only for Lauvastøl the difference between the median of the VGLM.MM simulations and the observed number of occurrences for precipitation intensities of 1-2 mm/day is as high as 70 counts, which is similar to the bias present in winter. It can be concluded that in general, summer time series are better reconstructed by the VGLM.MM bias correction.

4.1.2 Statistical Shape of the Precipitation Time Series

In order to assess the basic correction achieved by the bias correction methods the quantiles of the different time series are looked at. Figure 8 shows the quantiles of the uncorrected, the QM bias corrected and the VGLM.MM bias corrected RCM time series plotted against the quantiles of the observed time series.

The figure shows in particular the behaviour of the time series concerning the higher quantiles.

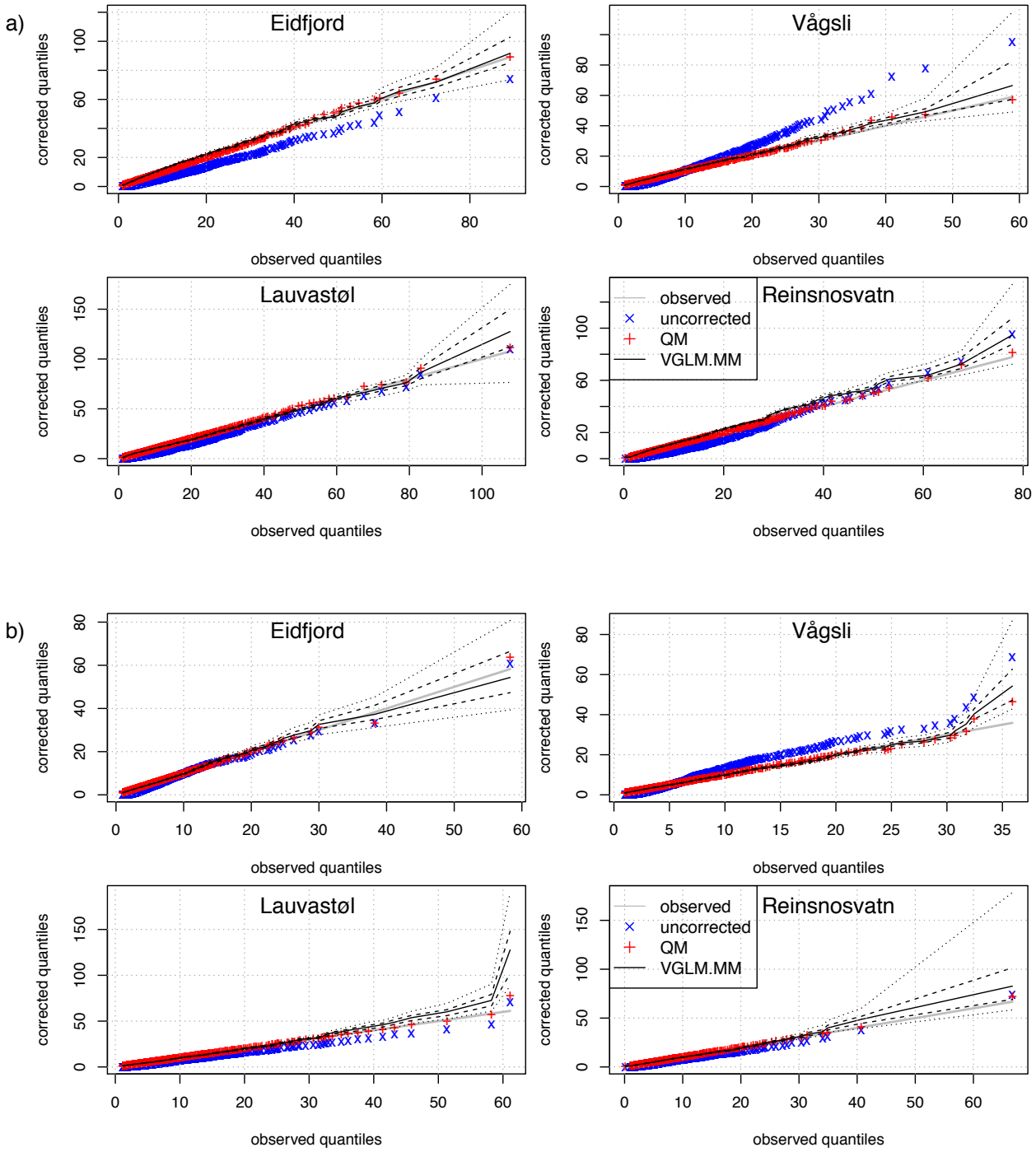


Figure 8: Quantiles of winter a) and summer b) uncorrected (blue crosses), QM bias corrected (red pluses) and 100 simulations from the VGLM.MM bias corrected (black line) time series against the observed ones (grey line) for the three stations and the catchment time series. The thick black line indicates the median of the 100 simulations, the dashed lines gives the 1st and 3rd quartile and the dotted lines give the 5th and 95th percentile of the VGLM.MM bias correction results. Only non-zero values were used to calculate the quantiles.

The quantiles of the uncorrected RCM indicate what kind of bias is present for the different time series.

For winter, Figure 8a), the uncorrected RCM values of Eidfjord, Lauvastøl and Reinsnosvatn show a negative bias for almost all observed precipitation intensities. For Reinsnosvatn the bias becomes positive for intensities larger than 35 mm/day. For Vågsli in winter the uncorrected RCM time series shows a negative bias for precipitation intensities of up to 10 mm/day. For higher intensities a general overestimation of the precipitation intensities in the uncorrected RCM time series is present, e.g. for observed precipitation intensities of 30 mm/day the bias accounts for 40% of the observed value and increases even more for higher quantiles.

In summer, Figure 8b) the quantiles of the uncorrected RCM time series do not show as large a bias as in winter. For Eidfjord, Lauvastøl and Reinsnosvatn a small negative bias is present in the uncorrected RCM time series, but for the highest quantile the bias for all three stations becomes positive and accounts for about 10% of the observed values. The largest bias in the uncorrected RCM time series is present for Vågsli. For precipitation intensities of up to 5 mm/day a negative bias is seen in the time series. However, for larger precipitation intensities a strong positive bias is present, which indicates an overestimation of precipitation intensities compared to the observed precipitation values of up to 90% for the highest quantile.

These very different biases should all be addressed by the bias correction methods.

In general for both seasons and all stations the QM bias correction performs well in correcting the RCM values to fit the quantiles of the observations. This is expected by how the bias corrected time series are constructed applying the QM method, compare Figure 4. The QM bias correction method corrects the bias for the main part of the precipitation distribution well, however if a strong bias is evident in the higher quantiles, as for Vågsli in summer, Figure 8b), the correction is not able to correct these biases to a satisfactory state. In this case a bias of 20% is still present in the bias corrected time series compared to the observed precipitation intensity. Furthermore, for Eidfjord and Lauvastøl in summer, after applying the QM bias correction the highest quantile shows a larger bias than before the QM bias correction was applied. This is a very undesirable behaviour, which might arise from the cross validation or the linear extrapolation, as explained in Section 3.1.1.

For both winter and summer, the quantiles from the 100 simulations of the VGLM.MM bias corrected time series show a good correspondence for the lower quantiles. The largest deviation from the lower observed quantiles is evident for Reinsnosvatn in winter, Figure 8a), where a slight positive bias is seen for precipitation intensities larger 15 mm/day, e.g. a bias of about 15% for observed precipitation intensities of 35 mm/day is evident.

However, for the highest quantile a large spread within the 100 simulations from the VGLM.MM is evident in all of the stations and especially in summer. In winter for Eidfjord, Figure 8a), the median of the 100 simulations for the highest quantile shows a slight positive bias of about 5% compared to the observations. Furthermore for the interquartile range of the simulations the remaining bias ranges between -2% to 15% compared to the observations. Although there is a slight overcorrection of the RCM value, this is considered a satisfactory correction for the VGLM.MM bias correction method. For Vågsli, the median value for the highest quantile shows a higher positive bias of about 10%, but the observed quantile still lies within interquartile range of the 100 simulations from the VGLM.MM and an improvement compared to the uncorrected RCM time series is evident. This is not the case for Lauvastøl, where the median value for the highest quantile has a positive bias of 15% compared to the observations and therefore is larger than the bias present in the uncorrected RCM time series. Furthermore for Lauvastøl and Reinsnosvatn the highest observed quantile does not lie within the interquartile range of the VGLM.MM simulations.

In summer for Eidfjord, Figure 8b), the achieved correction for the highest quantile is satisfactory, although an slight overcorrection is evident. The median shows a slight negative bias of about 5% and the observed value lies within the interquartile range of the simulations. This is not the case for any of the other stations in summer, where the observed value of the highest quantile always lies outside the interquartile range of the VGLM.MM simulations. However, for Vågsli the achieved correction is still an improvement compared to the uncorrected RCM time series. The positive bias was reduced from 90% to a value of 40% for the median of the 100 simulations. For Lauvastøl and Reinsnosvatn the value of the highest quantile from the VGLM.MM bias correction has a larger bias than the uncorrected RCM time series, with a remaining positive bias as high as 18% for Reinsnosvatn and 100% for Lauvastøl.

The large interquartile ranges of the simulations for the upper quantiles is an expected behaviour, since for the upper part of the distribution the VGLM.MM samples from a distribution with a higher variance, compare Figure 5. In an ideal case, like for Eidfjord in winter and summer, the distribution of the interquartile range should be normal around the observed value for the largest quantile. For Vågsli, Lauvastøl and Reinsnosvatn such a behaviour is not seen for both winter and summer. Especially for Lauvastøl and Vågsli in summer the spread is very high and not centred around the observed value, but rather located around higher precipitation intensities, Anyhow, for Vågsli an improvement compared to the uncorrected RCM time series is seen. This indicates an overestimation of the high precipitation intensities within the VGLM mixture model.

4.1.3 Mean Accumulated Precipitation

The mean accumulated precipitation is an important feature for impact modellers, since it gives information about the amount of water available as input to the impact model and it should correspond well to the observed value. For the accumulated winter (summer) precipitation the sum of all daily precipitation intensities from 1st December to 28th/29th February (from 1st June to 31st August) was calculated. This was carried out for all 38 seasons and the mean of these 38 values was calculated for all time series.

Table 3 gives the corresponding values of winter and summer observed time series, as well as the values for the uncorrected RCM and the values for the corrected RCM time series. For the probabilistic approach the mean of the 100 simulations mean accumulated precipitation is given with one standard deviation.

a) Station / Bias Correction		mean accumulated winter precipitation [mm/year]	relative to observed [%]
Eidfjord	Observed	516,7	100
	RCM	571,3	110,6
	QM	518,0	100,3
	VGLM.MM	572,9 ± 15,4	110,9 ± 3,0
Vågsli	Observed	367,0	100
	RCM	770,7	210,0
	QM	375,9	102,4
	VGLM.MM	404,8 ± 9,3	110,3 ± 2,5
Lauvastøl	Observed	716,3	100
	RCM	765,9	106,9
	QM	725,9	101,3
	VGLM.MM	723,6 ± 18,2	101,0 ± 2,5
Reinsnosvatn	Observed	568,9	100
	RCM	605,9	106,5
	QM	574,7	101,0
	VGLM.MM	651,2 ± 15,8	114,5 ± 2,8

b) Station / Bias Correction		mean accumulated summer precipitation [mm/year]	relative to observed [%]
Eidfjord	Observed	220,1	100
	RCM	408,2	185,4
	QM	220,6	100,2
	VGLM.MM	219,4 ± 6,1	99,7 ± 2,8
Vågsli	Observed	207,6	100
	RCM	617,8	297,6
	QM	207,9	100,1
	VGLM.MM	207,9 ± 6,3	100,1 ± 3,0
Lauvastøl	Observed	401,0	100
	RCM	499,8	124,7
	QM	402,3	100,3
	VGLM.MM	402,8 ± 10,9	100,5 ± 2,7
Reinsnosvatn	Observed	304,9	100
	RCM	393,0	128,9
	QM	305,9	100,3
	VGLM.MM	306,9 ± 8,0	100,7 ± 2,6

Table 3: Mean accumulated precipitation in mm/year for a) winter and b) summer. The values are calculated from the dry day corrected observed (Observed), the uncorrected SMHI model (RCM), the quantile mapped bias corrected (QM) and the VGLM mixture model bias corrected time series (VGLM.MM). In addition, values relative to the dry day corrected observed mean accumulated precipitation value are given in %.

For both winter and summer, the uncorrected RCM time series overestimate the mean accumulated precipitation. The worst case is Vågsli, where the accumulated precipitation for the uncorrected RCM time series accounts for 210% of the observed value in winter and for almost 300% in summer. Generally in winter the overestimation of the mean accumulated precipitation within the uncorrected RCM time series compared to the observed is not as strong as in summer. For winter Eidfjord, Lauvastøl and Reinsnosvatn show a bias in the accumulated precipitation of 7-11%, whereas in summer the bias for the three stations is higher ranging between 25-85%.

The QM bias correction is able to correct the value of the mean accumulated precipitation of the RCM for both winter and summer. The resulting values after the application of the QM bias correction range from 0-2% larger than the observed for winter and is the exact observed value for summer.

For the VGLM.MM bias correction it can differentiated between the two seasons: For summer, Table 3b), the correction of the mean accumulated precipitation is satisfactory with mean values ranging between 0-1% larger than the observed values. However, for winter, Table 3a), the mean accumulated precipitation remains biased after the bias correction was applied, with a range in the mean values between 1-15% larger than the observed value. Only for Lauvastøl and Vågsli an improvement of the mean accumulated precipitation compared to the uncorrected RCM time series is achieved, whereas for Eidfjord the correction did not show any effect on the mean accumulated precipitation, while for Reinsnosvatn the value with 114.5% is worse compared to the uncorrected RCM time series, with 106.5%.

This becomes problematic, if the bias corrected time series were used to run the HBV model, since the total amount of precipitation in the model would on average be increased compared to what was observed and hence will result in too much runoff.

4.1.4 Dry Spells

For a thorough analysis of the time series in the reference period the representation of dry spells is regarded in the following section. A dry day is defined as a day with precipitation intensities smaller than 1 mm/day. First, the total number of dry days within the 38 seasons and the corresponding dry day probability are calculated and given in Table 4.

a) Station	Observed	Uncorrected	QM bias corrected	VGLM.MM bias corrected
Eidfjord	1699 (0.50)	1283 (0.37)	1693 (0.49)	1700 ± 23 (0.50 ± 0.01)
Vågsli	1827 (0.53)	1173 (0.34)	1826 (0.53)	1825 ± 22 (0.53 ± 0.01)
Lauvastøl	1370 (0.40)	1302 (0.38)	1367 (0.40)	1369 ± 23 (0.40 ± 0.01)
Reinsnosvatn	1422 (0.41)	1412 (0.41)	1425 (0.42)	1420 ± 23 (0.41 ± 0.01)

b) Station	Observed	Uncorrected	QM bias corrected	VGLM.MM bias corrected
Eidfjord	2039 (0.58)	1264 (0.36)	2035 (0.58)	2034 ± 28 (0.58 ± 0.01)
Vågsli	2232 (0.64)	1068 (0.31)	2231 (0.64)	2229 ± 24 (0.64 ± 0.01)
Lauvastøl	1686 (0.48)	1265 (0.36)	1693 (0.48)	1690 ± 25 (0.48 ± 0.01)
Reinsnosvatn	1754 (0.50)	1282 (0.31)	1751 (0.50)	1754 ± 27 (0.50 ± 0.01)

Table 4: Total number of dry days for the observed, uncorrected, QM bias corrected and VGLM.MM bias corrected winter (a) and summer (b) time series for the three stations and the catchment time series. For the 100 simulations from the VGLM.MM bias correction the mean and one standard deviation is given. In addition the corresponding dry day probabilities are given in brackets.

All uncorrected RCM time series, for winter and summer, have a too low number of total dry days, which is in general well corrected by both of the bias correction methods, the QM and the VGLM.MM bias correction approach.

To gain a better understanding of the distribution of dry days, the number of dry spells of certain lengths, a dry spell spectrum, was calculated and is displayed as a histogram, shown in Figure 9.

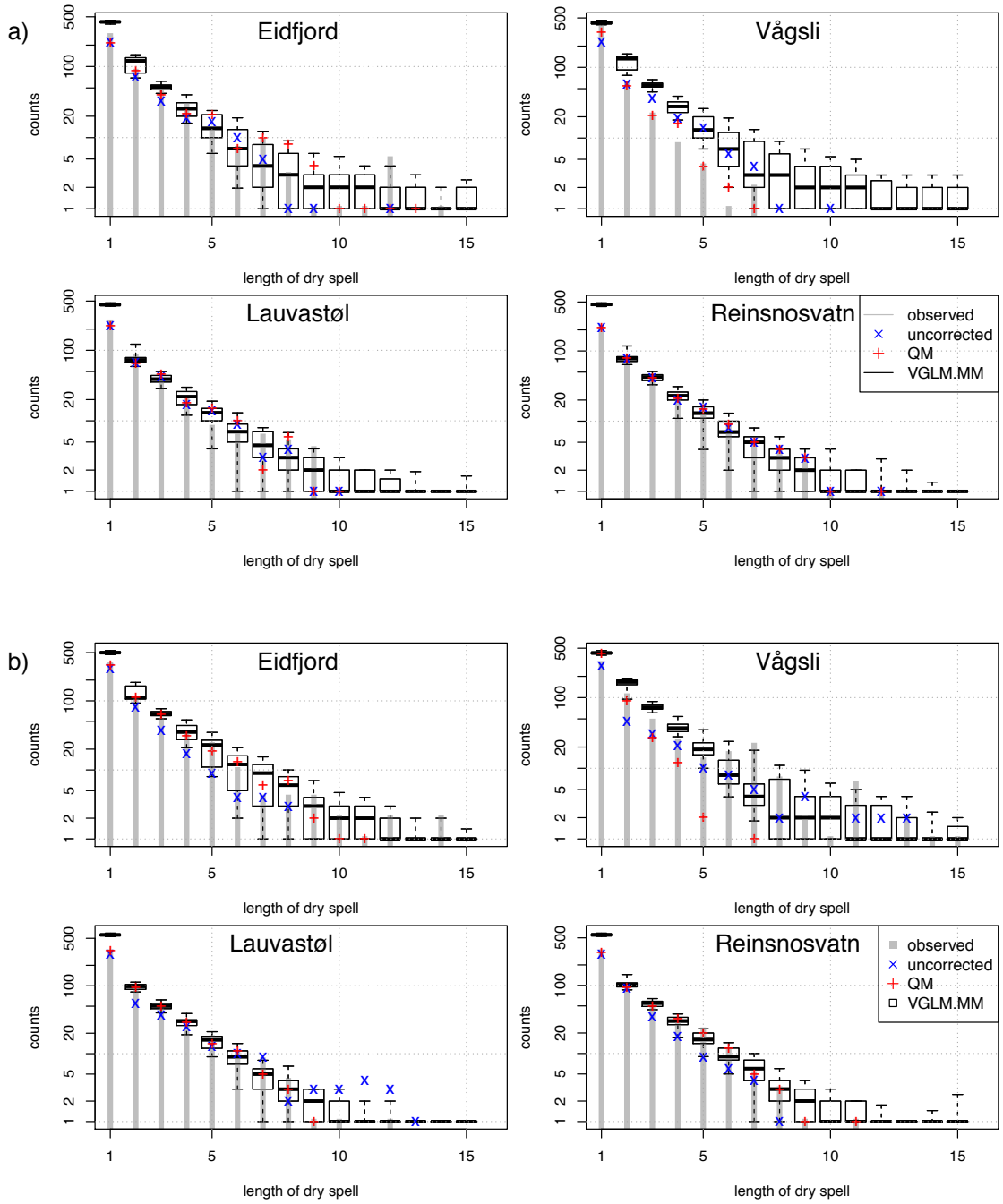


Figure 9: Histogram of dry spells for winter (a) and summer (b) observed (grey bars), uncorrected (blue crosses), QM (red pluses) and the VGLM.MM (black boxes) bias corrected time series are shown. The box gives the median of the 100 simulations from the VGLM.MM bias correction as well as the upper and lower quartile (25 percentile and 75 percentile) and the 5th and 95th percentile. Note, the logarithmic y-axis.

Comparing the shape of dry spells spectra makes sense, because the RCM is driven by reanalysis data and hence the overall temporal structure is similar to the observed. For hydrological modelling, the correction of dry spells is important, because otherwise the dynamical part of the model may be altered: If e.g. there are more short dry spells than observed and thereby less long dry spells, the ground water reservoir does not have the same amount of time to empty, which might alter the temporal characteristics of the runoff. Moreover, for a lower number of dry days or shorter dry spells the evapotranspiration might increase, due to a higher amount of available soil moisture in the unsaturated zone of the HBV model. This would result in a slightly reduced runoff.

From the observed dry spells, shown in Figure 9, it becomes clear that this parameter is not continuous: There are sporadic peaks such as the peak in the twelve days dry spell for Eidfjord in winter, or the peaks in seven and eleven days dry spells for Vågsli in summer.

The uncorrected RCM time series show large biases in the representation of dry spells, which are not systematic. In winter, Figure 9a), Eidfjord and Lauvastøl show an underrepresentation of short dry spells up to 4 consecutive dry days and no distinctive pattern for longer dry spells, there are large over and under representations of various dry spells. For Vågsli, the uncorrected RCM time series overestimates the amount of dry spells for all the different lengths except for one and two days, where it underestimates the number. Only for Reinsnosvatn, the uncorrected RCM time series shows a very good representation of the dry day spectrum of the observed time series.

In summer, Figure 9b), the dry spell spectrum of the uncorrected RCM time series for Eidfjord shows a similar behaviour as for winter. For Lauvastøl, the short dry spells are again underrepresented, whereas dry spells of lengths 9 to 13 consecutive dry days are strongly overrepresented. The dry spell spectrum for Vågsli, shows in contrast to winter an underrepresentation of almost all the dry spells, except for 9 and 12 day long dry spells. For Reinsnosvatn the representation is again good, however a slight underestimation of dry spells longer than 3 days is evident in the uncorrected RCM time series.

In winter, Figure 9a), for QM bias corrected time series, the number of dry spells of one and two days is still slightly underestimated for all stations. For Eidfjord, Vågsli and Lauvastøl the dry spell spectra of the observations are well represented in the QM bias corrected time series, especially for the shorter dry spells up to 3 or 4 consecutive dry days. However, for the longer dry spells there are many outliers in both positive and negative direction. This make it seem as if the hits are more coincidently. It becomes clear that this parameter is not directly addressed by the correction. Although a dry days correction is carried out, the occurrence is strongly dependent on the temporal structure of

the uncorrected RCM time series. For Reinsnosvatn the QM bias corrected time series shows a good representation of dry spells for different lengths. This can be explained by the good representation of the dry spells of the uncorrected RCM time series. However, it is seen that the bias correction does not alter this good representation.

For summer, Figure 9b), the dry spells seem to be better represented in the QM bias corrected time series. For Lauvastøl and Reinsnosvatn they are very close to the observed number of occurrence, as it is in the case for Eidfjord for shorter dry spells up to 6 consecutive dry days. For Vågsli the representation after the bias correction has not improved and became for some parts even worse when dry spells lengths of 4-7 days are considered. Furthermore for Eidfjord, Lauvastøl and especially for Vågsli, the long dry spells are missing in the QM bias corrected time series compared to the observed.

For the VGLM.MM bias correction it has to be noted, that for one station within one season all 100 simulations follow the same logistic regression model. They are therefore expected to have a small range of values for the dry days. As a general remark before looking at the details, it is seen that the structure of the dry spells spectrum is more continuous for the VGLM.MM bias corrected time series. This is expected, since it is based on 100 simulations. Only twice this continuous behaviour is disrupted. For Eidfjord in winter and for Vågsli in summer there is a peak in the 15 days long dry spell. This in general continuous structure lacks the ability to represent the arbitrary structure of dry spell occurrences from the observed time series.

As a first note, it is seen that the amount of dry spells of one day length is overrepresented for all stations and both seasons except for Vågsli in winter. This indicates a lack of temporal structure in the logistic regression model, identified by the fact that dry days occur more often as a single event rather than as a sequence of dry days.

In winter, Figure 9a), for Eidfjord, Lauvastøl and Reinsnosvatn the shape of the dry spells spectrum matches the one from the observations, although it is an overrepresentation of dry spells longer than ten days compared to the observations for Lauvastøl and Reinsnosvatn. The station Vågsli however, shows a general overrepresentation of the number of dry spells for all lengths, especially for dry spells longer than eight days.

In summer, Figure 9b), the VGLM.MM time series show a good resemblance to the observed dry spells spectrum, disregarding the sporadic peaks. Again the station Vågsli shows the largest differences, with overrepresentation of dry spells of lengths between two and five days. Furthermore, in summer, there is a representation of the long dry spells in the VGLM.MM bias corrected time series , which is again not seen in the observations. This is especially the case for Lauvastøl and Reinsnosvatn.

4.1.5 Heidke Skill Score - Testing the Logistic Model

In the following section the ability to correct or reconstruct the day to day variability will be analysed. Since the RCM is forced with reanalysis data the daily variability can be compared to the observed one. The concept for the Heidke Skill Score (HSS) was originally introduced by Doolittle [1888] but it is nowadays known as the Heidke Skill Score [Heidke, 1926]. This skill score is a tool from the weather forecast verification and tests the skill to forecast an event on the basis of a 2x2 contingency table. In this case the skill of the corrections to reconstruct the observed temporal structure is tested, based on the test statistic of a wet day. The application of the HSS is taken from Wilks [2005]. The contingency table applied here is given in Table 5.

$n=a+b+c+d$	observed wet day	observed dry day	
simulated wet day	a	b	marginal probability of a wet day forecast
simulated dry day	c	d	marginal probability of a dry day forecast
	marginal probability of a wet day observation	marginal probability of a dry day observation	

Table 5: The concept of the contingency table for the Heidke Skill Score, testing the forecast performance of a wet / dry event in the bias corrected RCM time series.

The amount a states how often an event was forecasted and actually occurred on that particular day and is referred to as a hit, b stands for the event was forecasted but did not occur, which means a false alarm. c stands for the event was not forecasted but did occur, which is a miss and d stands for the event was not forecasted and did not occur, which is a correct no-forecast.

The HSS score tests if the amount of correct forecasts is higher than the amount of correct forecasts that would be expected by random forecasts that are statistically independent of the observations, so random hits by mere chance.

The probability of a correct wet day forecast by chance is given in equation (9), the probability of a correct dry day forecast by chance is given by equation (10):

$$P_{yes} = [\,(a+b)/n\,] / [\,(a+c)/n\,] = (a+b) * (a+c) / n^2 \quad (9)$$

$$P_{no} = [\,(b+d)/n\,] / [\,(c+d)/n\,] = (b+d) * (c+d) / n^2 \quad (10)$$

The HSS compares the perfect correct (PC) forecasts, i.e. the sum of a and d, with the sum of P_{yes} and P_{no}, as given in equation (11):

$$\begin{aligned} HSS &= PC / [\, P_{yes} + P_{no} \,] \\ &= [\, (a+d)/n \,] / [\, P_{yes} + P_{no} \,] \\ &= 2 * (a*d - b*c) / [\, (a+c)*(c+d) + (a+b)*(b+d) \,] \end{aligned} \quad (11)$$

The HSS was applied to test how accurate the dry / wet events are represented in the bias corrected time series on a day to day basis. The skill score might range from -∞ to 1, where negative values indicate that the forecast is worse than the reference forecast. If the HSS=0 the forecast shows no skill in representing the reference forecast, and for a value of 1 the forecast is perfect. In this case this means that the corrected time series do show the same day to day wet/dry behaviour as the observations. The results of the HSS calculation for winter and summer are given in Table 6. The HSS values are given for the uncorrected RCM time series, the QM bias corrected time series and VGLM.MM bias correction, where the mean and one standard deviation of the score for the 100 simulations is given.

a) DJF HSS	uncorrected	QM	VGLM.MM
Eidfjord	0,16	0,59	0,39 ± 0,01
Vågsli	0,13	0,51	0,33 ± 0,01
Lauvastøl	0,24	0,59	0,38 ± 0,01
Reinsnosvatn	0,30	0,99	0,48 ± 0,01

b) JJA HSS	uncorrected	QM	VGLM.MM
Eidfjord	0,02	0,37	0,18 ± 0,01
Vågsli	0,02	0,39	0,18 ± 0,02
Lauvastøl	0,03	0,47	0,23 ± 0,01
Reinsnosvatn	0,03	0,99	0,37 ± 0,02

Table 6: Results from the HSS score calculations for winter a) and summer b) of the three stations and the catchment time series for the uncorrected and quantile mapping bias corrected time series and the mean and one standard deviation of the 100 simulations of the VGLM mixture model bias correction.

In winter, Table 6a), for the uncorrected RCM time series the values are positive and range from 0.13 for Vågsli to 0.30 for Reinsnosvatn. This indicates, that the day to day variability given within the uncorrected RCM time series resembles the observational time series, but is low considering, that the RCM is driven by reanalysis data.

For the QM bias corrected winter time series, the HSS values are higher than for the uncorrected RCM time series, ranging from 0.51 for Vågsli to 0.99 for Reinsnosvatn. This indicates, that the dry days correction and the applied dry day threshold for the different RCM time series affected days with low precipitation intensities, which were regarded as a dry day in the observation time series. The high HSS value for Reinsnosvatn indicates an almost perfect correspondence between the bias corrected time series and the observations. This is also evident in the spectrum of the winter dry spells.

For the 100 simulations for winter of the VGLM.MM bias correction, the mean HSS values range from 0.33 for Vågsli to 0.48 for Reinsnosvatn. This indicates an improvement of the forecast performance for a wet event compared to the uncorrected RCM time series. The HSS only regards the performance of the logistic part of the VGLM.MM and shows that the representation of the temporal structure has problems in reconstructing the observed one. This can party be explained by that the information for the temporal structure to build the logistic model is taken from the RCM time series.

In summer, Table 6b), for the uncorrected time series, the correspondence of the wet / dry day to day variability is poor, the values range from 0.02 for Eidfjord and Vågsli to 0.03 for Lauvastøl and Reinsnosvatn. The skill to forecast the observed wet events within the reanalysis forced RCM is almost not present.

For the QM bias corrected time series the temporal structure of the wet days is much improved compared to the uncorrected RCM time series. The HSS values range from 0.37 for Eidfjord to 0.99 for Reinsnosvatn. These values show that the performance in reconstructing the temporal structure of the observational time series are dependent on the station. Again for Reinsnosvatn an almost perfect forecast is achieved by the quantile mapping bias correction.

For the 100 simulations of the VGLM.MM bias correction for summer, the mean HSS values range between 0.18 for Eidfjord and Vågsli to 0.37 for Reinsnosvatn. An improvement compared to the uncorrected RCM time series is seen, but again the logistic model is not able to reconstruct the observed temporal structure.

In general, all bias corrected time series have the highest HSS value for Reinsnosvatn. This might be the case because, the observed time series of Reinsnosvatn is an area time series, as is the uncorrected RCM time series. The variability within these time series is therefore more likely to resemble each other.

4.1.6 How to build a Bias Corrected Catchment Time Series?

When applying any bias correction method within the context of hydrological modelling, the question arises, whether to first bias correct station time series and average them to a catchment time series after the correction was applied, or if to first built catchment time series from the RCM and then directly bias correct them is a better approach to force the hydrological model?

For both bias correction approaches one directly bias corrected catchment time series and one catchment time series from the bias corrected stations were calculated. Figure 10 shows the sorted precipitation intensities of the uncorrected, the QM bias corrected and VGLM.MM bias corrected time series plotted against the observed intensities.

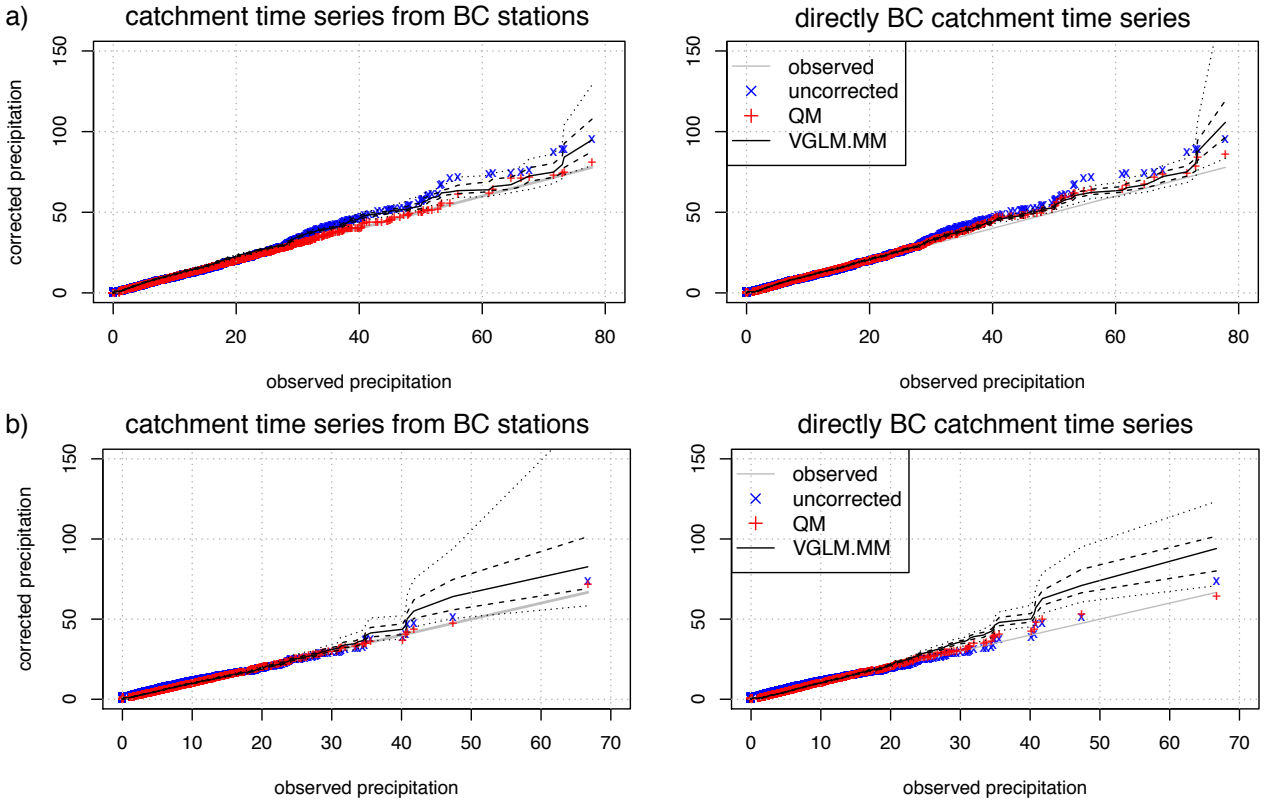


Figure 10: Sorted catchment time series of the winter (a) and summer (b) for the uncorrected (blue crosses), the QM bias corrected (red pluses) and the range of the 100 simulations from the VGLM.MM (black lines) against the sorted observed time series. The solid black line indicated the median, the dashed ones the 1st and 3rd Quartile and the dotted the 5th and 95th percentile of the 100 simulations. The left side shows directly bias corrected catchment time series, the right side shows the results, if the catchment time series is calculated as an average of bias corrected stations.

For the winter precipitation intensities, Figure 10a), the uncorrected RCM does not show a strong bias for values of up to 30 mm/day, however for higher precipitation intensities the

bias becomes positive and amounts to an increase of 20% compared to the observed values.

In winter, Figure 10a), the direct approach of the QM bias correction shows a close fit to the observed time series up to precipitation intensities of 55 mm/day. For higher values there is a slight positive bias present. This bias might be explained by a lower number of observed high precipitation intensities due to the cross-validation, therefore the uncertainty of reproducing the correct high precipitation intensity increases. This is a feature of the bias correction method that is also expected to be observed if the QM is applied for future predictions. Furthermore, for the range of the 100 simulations from the VGLM.MM it becomes evident that the bias in the directly bias corrected catchment time series reduces the bias present in the uncorrected RCM time series. For precipitation intensities up to 30 mm/day the corrected time series follow the observed one well. For higher intensities up to 50 mm/day, the simulations seem to follow the behaviour of the uncorrected RCM time series, whereas for intensities larger than 60 mm/day the bias of the uncorrected RCM time series is again reduced.

Winter catchment time series from bias corrected stations show for both bias correction methods a satisfactory correction for precipitation intensities up to 30 mm/day. However, for higher precipitation intensities a larger positive bias is evident for both the QM and the VGLM.MM bias correction. Especially for the QM bias correction, there is a stronger deviation from the observations, whereas for the VGLM.MM the structure is only changed for the highest quantile.

This indicates that the approach to directly bias correct a catchment time series would give better results than an average calculate from bias corrected stations.

For the summer catchment time series a similar result can be drawn: The directly bias corrected time series lie closer to the observed precipitation intensities than the catchment time series calculated from the bias corrected stations, for both of the methods.

It should be mentioned that in summer the uncorrected RCM time series corresponds really well with the observed one and it seems as if at least for the station average there is an additional bias introduced for bias corrected values larger than 35 mm/day for both methods.

4.2 Pseudo-reality

What is ought to be addressed in this section is the ability of the quantile mapping (QM) bias correction to perform in future scenarios. Therefore the RCM runs will be bias corrected to a pseudo-reality, another RCM run driven by the same GCM. Therefore the bias arising from the GCM is present in the same manner for all the RCM runs and hence can be disregarded.

The bias corrected time series will be compared to the pseudo-reality they were corrected to, with regard to mean accumulated precipitation, annual cycles and the dry day probability. However, trends in the bias corrected time series are compared to the uncorrected RCM time series, in order to see if changes are introduced due to the QM bias correction.

In addition to temperature and precipitation time series, also runoff from the HBV model is investigated within this section. It is calculated from station-based bias corrected temperature and precipitation time series. The time series were used on a station basis, although it was found to be an improvement to directly bias correct a catchment time series, because the implementation of the HBV model used in this study demands for station time series.

4.2.1 Temperature, Precipitation and Runoff in the Validation Period

In order to validate the basic correction achieved by the QM bias correction the mean annual mean temperatures, annually accumulated precipitation and the resulting annually accumulated runoff for the validation period are studied. It is expected that the bias corrected time series show good correspondence to the pseudo-reality they were corrected to in these features.

Therefore, Figure 11 shows the differences of the bias corrected time series to the chosen pseudo-reality for mean annual mean temperatures, mean annually accumulated precipitation and mean annually accumulated runoff in the validation period, from 2069 to 2099. The x-axis shows from which uncorrected RCM the time series are taken, and the numbers on the y-axis give the RCM (identified by number) used as pseudo-reality to bias correct the uncorrected RCM time series to.

For temperature deviations, shown in Figure 11a), the range between the pseudo-realities and the corrected time series lies between -0.4 °C for the DMI-HIRHAM5 bias corrected time series, when corrected to the SMHIRCA time series and 0.5 °C for the time series of SMHIRCA, when bias corrected to the DMI-HIRHAM5 as the chosen pseudo-reality. A difference of 0.4 °C accounts for 16% of the trends present in the temperature time series,

compare Table 7. In general the time series of the DMI-HIRHAM5 (first column on the left) show negative differences larger than 0.2 °C after bias corrected to any pseudo-reality. Also evident is a positive bias present in all RCM time series bias corrected to the DMI-HIRHAM5 (top row). The time series bias corrected to the DMI-HIRHAM5 hence show larger mean annual mean temperatures in the validation period than the pseudo-reality. The opposite is true for the SMHIRCA bias corrected time series. All time series of this RCM show slightly positive temperature values of up to 0.4 °C, when compared to any of the pseudo-realities and lower temperature values are evident in the bias corrected times series if the time series were bias corrected to the SMHIRCA as the chosen pseudo-reality.

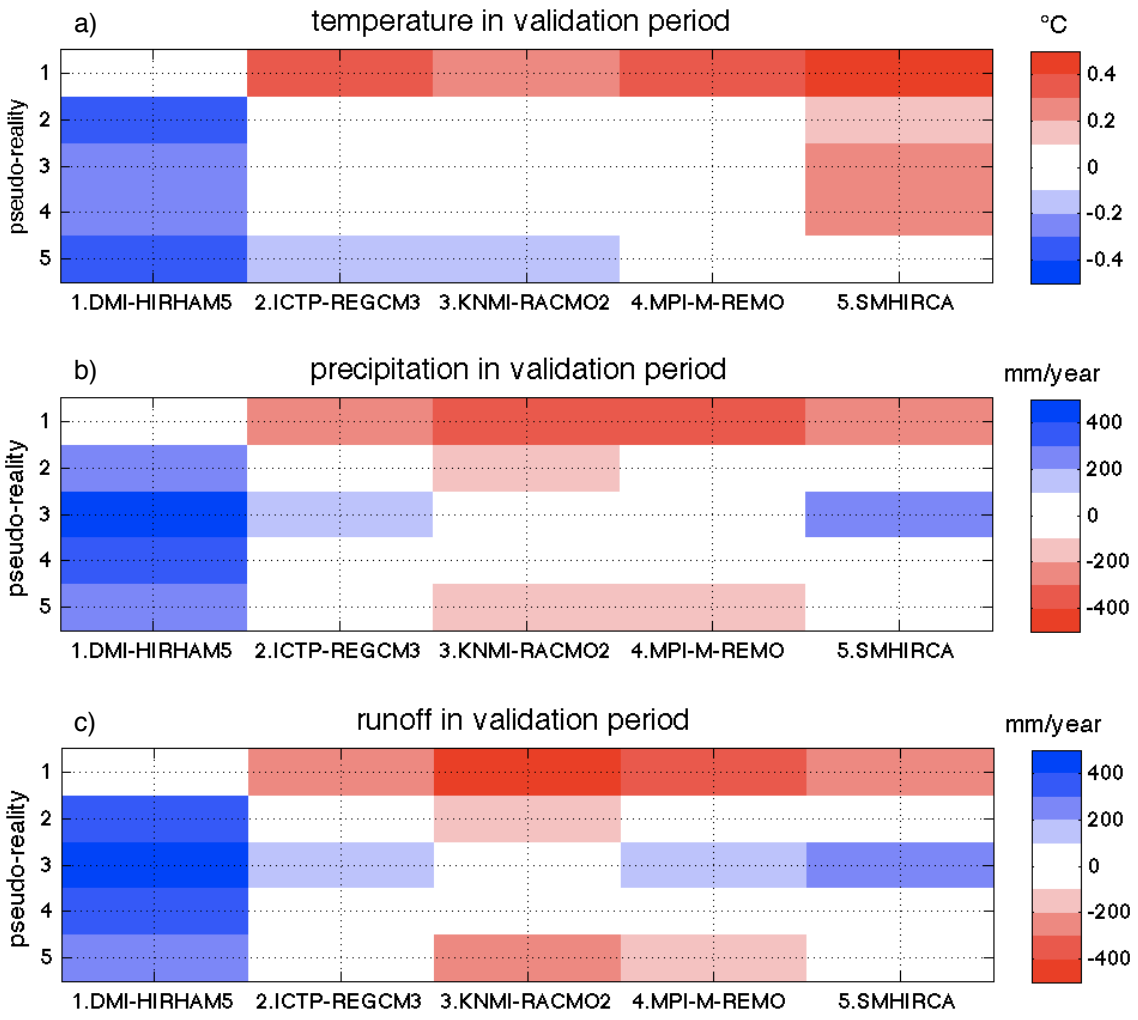


Figure 11: Differences in a) mean annual mean temperature in °C, b) mean annually accumulated precipitation in mm/year and c) mean annually accumulated runoff in mm/year for bias corrected time series in the validation period relative to the pseudo-reality they were corrected to. The numbers on the y-axis give the RCM time series that was used to bias correct the other time series to, since relative values are regarded, zero values on the diagonal are expected.

For the other RCMs the differences between the bias corrected time series and the pseudo-reality are smaller than 0.1 °C. These values are small compared to the absolute changes in the time series from the calibration period to the validation period, where a change of 0.1 °C accounts for 4%. Furthermore, for the RCMs the mean annual mean temperatures within the calibration period have a maximum range of 1.4 °C, compare Table 7, so a change of 0.1 °C accounts for 7%.

For mean annually accumulated precipitation, Figure 11b), the total range of differences in the mean annually accumulated precipitation spans from -400 mm/year for the KNMI-RACMO2 bias corrected time series, when bias corrected to the DMI-HIRHAM5, to 500 mm/year for the DMI-HIRHAM5 time series, when bias corrected to the KNMI-RACMO2. A difference in mean annually accumulated precipitation of 500 mm/year denotes a relative difference of about 16% compared to an average annual value of 3000 mm/year, and is in the same order of magnitude than the change between the calibration and validation period, compare Table 7.

In general all of the DMI-HIRHAM5 time series show higher mean annually accumulated precipitation values after the bias correction than the corresponding pseudo-realities. Furthermore, all time series bias corrected to the DMI-HIRHAM5 show a drier precipitation behaviour than the pseudo-reality.

An opposite behaviour is evident for the KNMI-RACMO2: All bias corrected time series of this RCM show a lower mean annual accumulated precipitation than their pseudo-realities. Furthermore, all the time series corrected to the KNMI-RACMO2 show higher precipitation values than the pseudo-reality within the validation period.

The remaining differences are lower than 100 mm/year, which correspond to a 3% decrease or increase of the mean annually accumulated precipitation compared to the pseudo-reality. However, these differences cannot be disregarded when compared to the trends present in the precipitation time series, with values ranging from 50 to 500 mm/year.

For the resulting mean annually accumulated runoff, shown in Figure 11c), the pattern is similar to the one of the precipitation values. The range lies between an increase in the runoff values up to 550 mm/year for the DMI-HIRHAM5 time series, when bias corrected to the KNMI-RACMO2 time series as the chosen pseudo-reality, and a decrease in runoff of 450 mm/year in the case where the KNMI-RCAMO2 time series is bias corrected to the DMI-HIRHAM5 time series. This indicates, that the amount of runoff in the validation period is strongly linked to the available precipitation.

In general, the differences of the bias corrected time series to the pseudo-realities in the validation period are high and ought not to be disregarded.

4.2.2 Dry Day Probability in the Validation Period

The mean annual dry day probabilities of the bias corrected precipitation time series are compared within the validation period. A dry days spectrum as presented in Section 4.1.4 is not suitable here, because the temporal variability within the RCMs differ, as do the corresponding histograms. Instead the total number of dry days in form of the annual dry day probabilities of the different time series is investigated. This feature is expected to have a good agreement in the bias corrected time series compared to the corresponding pseudo-reality.

Figure 12 shows the mean annual dry day probabilities of the uncorrected (on the diagonal) and bias corrected time series. The numbers on the y-axis indicate which RCM was used as pseudo-reality.

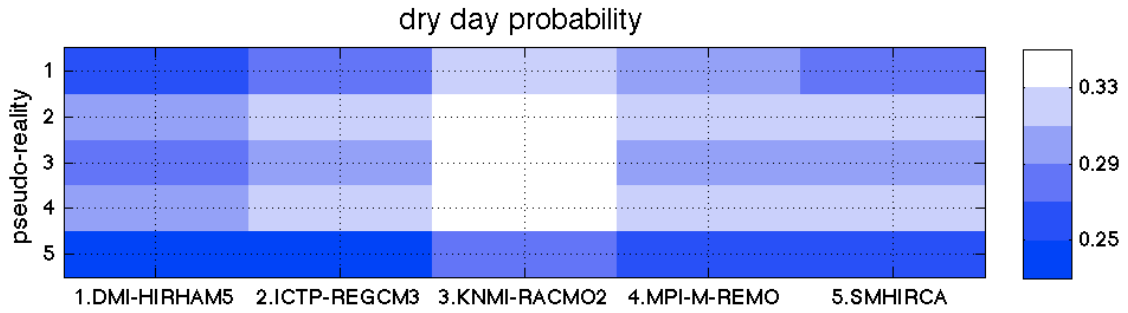


Figure 12: Dry day probabilities for the bias corrected time series and the pseudo-realities in the validation period. The values give the probability of a dry day occurrence in the time series, where a dry day is defined as a precipitation intensity smaller than 1 mm/day. The numbers on the y-axis give the RCM time series that was used to bias correct the other RCM time series to.

Since the dry days should be corrected to the value in the pseudo-reality, a horizontal pattern is expected. What is seen is a dependence on the pseudo-reality as well as on the RCM time series that was bias corrected. The values of the dry days probabilities range between lowest probabilities of 0.23, for the time series corrected to the SMHIRCA, to maximum values of 0.36, for the time series of the KNMI-RACMO2. The value becomes largest when the KNMI-RACMO2 time series is bias corrected to the ICTP-REGCM3. In general the KNMI-RACMO2 time series show higher dry day probabilities than the other time series indicating a strong dependence on the time series that is bias corrected. On the other hand the dependence on the pseudo-reality is seen if the time series are corrected to the SMHIRCA time series.

In general it can be said, that the probabilities are not corrected to the value of the pseudo-reality by the bias correction, but altered from the uncorrected RCM time series.

4.2.3 Annual Cycles of Precipitation, Temperature and Runoff

In this section it is tested if the seasonal cycles of runoff time series are altered by the QM bias correction. For a better understanding the seasonal cycles of temperature, precipitation and runoff time series of the calibration and the validation period are compared. Figure 13 shows the mean seasonal cycles in the two periods for all five pseudo-realities for temperature in the top row, precipitation in the middle row and runoff in the bottom row.

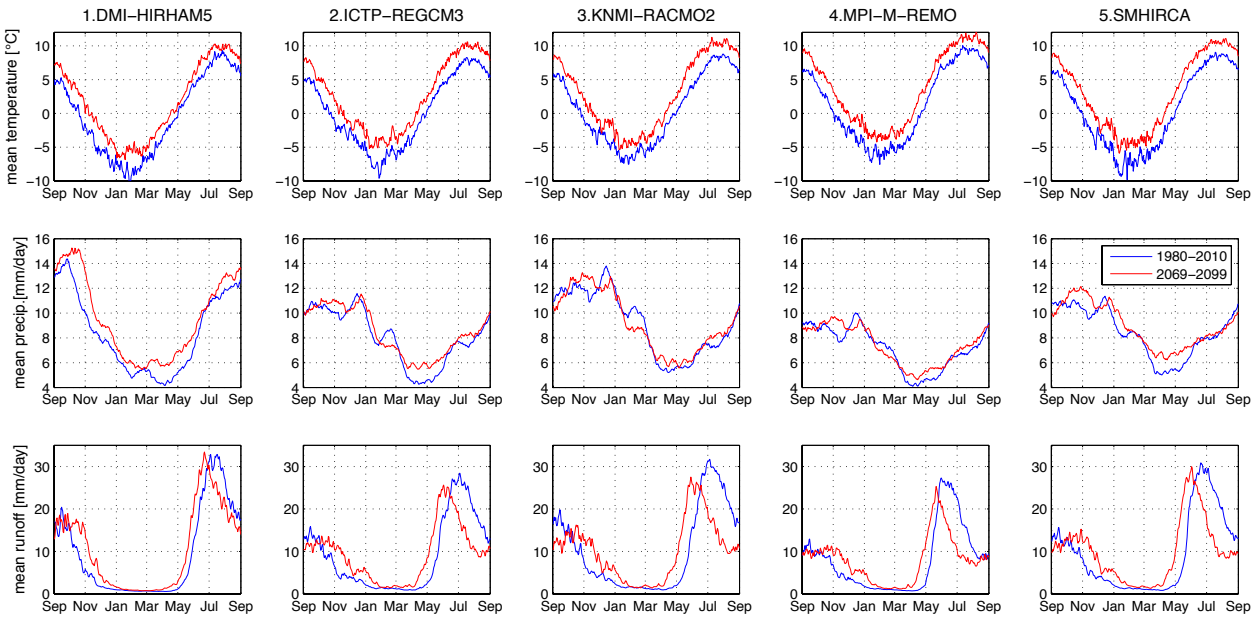


Figure 13: Mean seasonal cycles of temperature (top), precipitation (middle) and runoff (bottom) time series for the five RCMs used as pseudo-realities. The blue line is the average seasonal cycle in the calibration period from 1980 to 2010 and the red line gives the average seasonal cycle for the validation period from 2069 to 2099. The seasonal cycle of precipitation was smoothed by applying a 30-days running mean.

The seasonal cycles of temperature in the calibration period show maximum temperatures of 11 °C in August for the MPI-M-REMO and minimum temperatures in February of -8 °C for DMI-HIRHAM5. In the validation period the phase and amplitude of the seasonal cycles are not altered for any RCM time series compared to the calibration period, however there is an offset of about 2.5 °C in each of the RCM time series, compare with the values given in Table 7.

The smoothed seasonal cycle of precipitation shows larger differences for the different RCM time series: For DMI-HIRHAM5 the seasonal cycle of precipitation in the calibration period ranges from 4 mm/day in April to 14 mm/day in October. In the validation period there is a shift of the maximum precipitation values of about half a month, as well as an

increase of the precipitation intensities all over the year of about 1 mm/day. For the time series of the other four RCMs the amplitude in the calibration period is lower, ranging from 3 mm/day for MPI-M-REMO to about 4 mm/day for KNMI-RACMO2. The time series in the validation period for KNMI-RACMO2 and MPI-M-REMO show a similar amplitude and phase than the seasonal cycle in the calibration period. The same is true for ITCP-REGCM3, with the exception of April and May mean precipitation values, where there are higher average values in the validation period of about 2 mm/day compared to the calibration period. For the SMHIRCA the phase of the seasonal cycle did not change, but in the validation period the average maximum values in November and the average minimum values in April are higher by about 1 mm/day.

The seasonal cycles of runoff in the calibration period show a peak in end of June and July indicating the beginning of the spring flood, so the snow/glacier melting season with values between 25 mm/day to 32 mm/day depending on the RCM. This peak, which coincides with the maximum temperatures, is prolonged by the increase in precipitation and the melting of the remaining snow until autumn, with values around 15 mm/day. During spring the minimum average runoff values occur with an average runoff of 1 mm/day.

In the validation period a shift of the melting season towards an earlier time of the year of about one month for all the RCM time series is seen. In addition, the whole melting season is prolonged by about a month. The reason for the change in the melting season is the increase in temperatures. The hydrological model simulates freezing as soon as the temperature drops below 0.5 °C. The melting period coincides with the time in the year, when the seasonal cycle of the temperature time series is higher than 0.5 °C. For the ITCP-REGCM3, the KNIM-RACMO2 and the MPI-M-REMO time series the amplitude of the spring flood is reduced as well as the melting in autumn, because the amount of snow which fell in the winter season remained unchanged, due to no change in the precipitation amount. However, an increased evapotranspiration reduced the amount of water available. For the DMI-HIRHAM and the SMHIRCA the increase in evapotranspiration is balanced by the increase in precipitation.

Now the change in the seasonal cycle of runoff is considered for the bias corrected time series. The question is, whether the shift in the seasonal cycle of runoff between calibration and validation period is altered by the QM bias correction?

Figure 14 shows the runoff difference between the two periods for each of the bias corrected time series. The top row shows the difference between the seasonal cycle of runoff between the calibration period and the validation period for the pseudo-reality and the RCM time series that were bias corrected to it, given in the legend. The bottom row

shows all the time series that belong to the RCM, the uncorrected one and the time series bias corrected to the pseudo-reality, given in the legend.

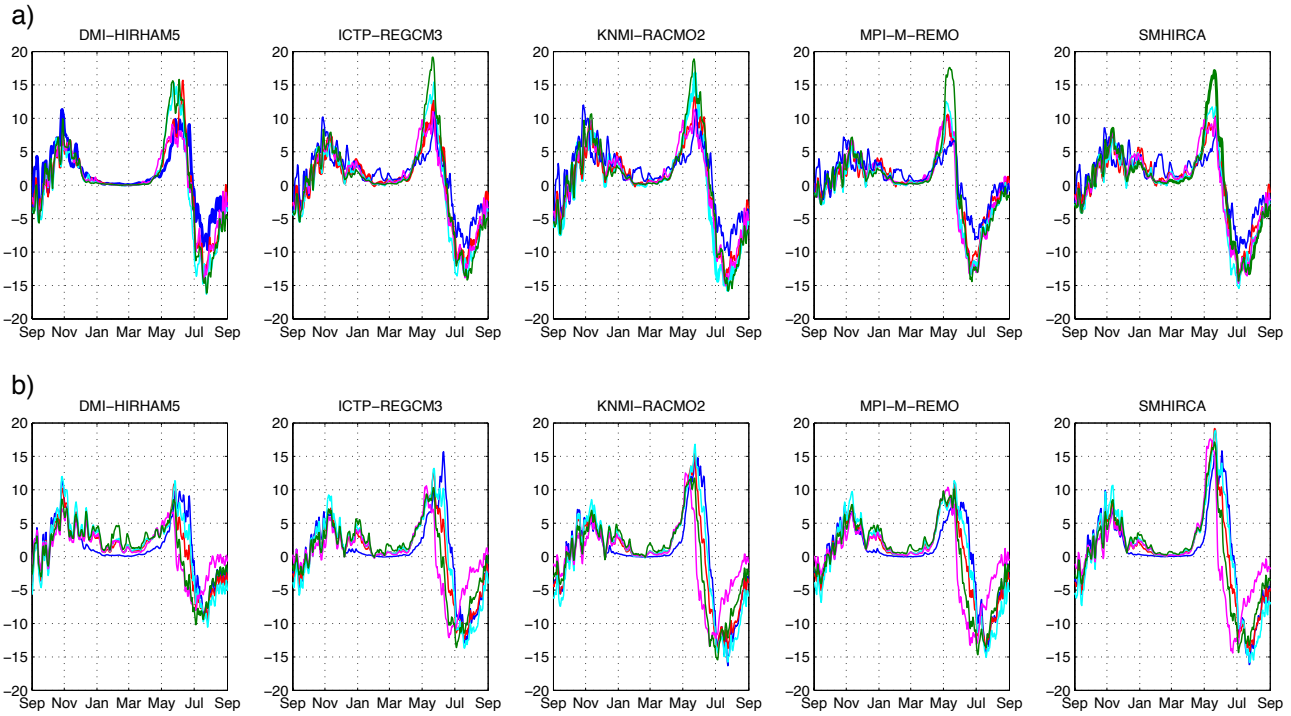


Figure 14: The differences in the seasonal cycle of runoff between the validation and the calibration period. a) The title gives the pseudo-reality and the legend shows the time series of which RCM was bias corrected to it. b) The title gives the RCM from which the time series are shown and the legend gives the corresponding pseudo-reality, that was used to bias correct the time series.

In general, the change in runoff between the validation and calibration period shows an increase in runoff for the validation period in May and a decrease of the runoff in July. This indicates the shift of the spring flood from July to May, increasing the runoff in May, but decreasing it in July, where the peak used to be in the calibration period. In addition, a delayed start of the winter season results in an increase in runoff in November.

When the pseudo-realities are compared to the time series corrected to them, the timing of the shift present in the pseudo-reality is well represented in the bias corrected time series, see Figure 14a). However the amplitude of the change between the runoff in the two periods is not altered, and depends on the RCM time series chosen to be bias corrected, see Figure 14b).

4.2.4 Future Scenarios - Changes in Temperature, Precipitation and Runoff

In order to assess the changes in temperature, precipitation and runoff the calibration time period from 1st September 1980 to 31st August 2010 was compared to the validation time period from 1st September 2069 to 31st August 2099. Table 7 gives information about the states of the three parameters in the pseudo-realities for the calibration and validation period.

Parameter	Pseudo-reality	Calibration period 1980 - 2010	Validation period 2069 - 2099
mean annual mean temperature [°C]	DMI-HIRHAM5	-0,52	1,72
	ICTP-REGCM3	0,12	2,63
	KNMI-RACMO2	0,44	2,89
	MPI-M-REMO	1,33	3,78
	SMHIRCA	0,33	3,04
mean annually accumulated precipitation [mm/year]	DMI-HIRHAM5	3061	3502
	ICTP-REGCM3	2938	3089
	KNMI-RACMO2	3344	3372
	MPI-M-REMO	2611	2693
	SMHIRCA	3081	3283
mean annually accumulated runoff [mm/year]	DMI-HIRHAM5	3432	3918
	ICTP-REGCM3	3289	3437
	KNMI-RACMO2	3720	3760
	MPI-M-REMO	2911	2950
	SMHIRCA	3469	3645

Table 7: Mean values of annual mean temperature, annually accumulated precipitation and annually accumulated runoff for the periods of 1st September 1980 to 31st October 2010 (calibration) and 1st September 2069 to 31st August 2099 (validation). The values are given for the five uncorrected RCM time series which are used as pseudo-realities.

For the mean annual mean temperature in the calibration period the values range between -0.5 °C for DMI-HIRHAM5 to 1.3 °C for MPI-M-REMO. For the validation period the values range from lowest future temperature of 1.7 °C for the DMI-HIRHAM5 to the largest future annual mean temperature values of 3.8 °C for MPI-M-REMO.

For the mean annually accumulated precipitation the values in the calibration period range from 2611 mm/year for the MPI-M-REMO to 3344 mm/year for the KNMI-RACMO2. For

the future period the minimum mean annually accumulated precipitation is found for MPI-M-REMO with 2693 mm/year and the highest values occur for DMI-HIRHAM5 with 3501 mm/year.

The connection of the temperature and especially the precipitation to runoff is seen in the range of the mean annually accumulated runoff values: For the calibration period the smallest runoff value of 2911 mm/year is found for the same RCM which had the smallest precipitation value and the highest temperature value, the MPI-M-REMO. The highest runoff value of 3720 mm/year is found for KNMI-RACMO2 which happened to have the largest precipitation amount for this period. This indicates that the influence temperature has on evapotranspiration, snow/glacier melt is minor to the amount of available water in the model provided by rain or snow. The same is true for the future runoff calculations in the validation period: Highest runoff values of 3918 mm/year occur in the same model which has highest precipitation values, the DMI-HIRHAM5 and lowest values of 2950 mm/year occur for MPI-M-REMO, which had the lowest amount of mean annual accumulated precipitation. These RCMs will be analysed with special attention when analysing the trends of the bias corrected time series.

The changes in mean temperature, mean annually accumulated precipitation and mean annually accumulated runoff between the validation and the calibration period are investigated and are referred to as trends. The trends of the time series should not be altered through the application of the bias correction compared to the uncorrected RCM time series. Figure 15 shows the trends of the three parameters: The diagonal always displays the value of the uncorrected RCM time series, the numbers on the y-axis give the number of the RCM used as the pseudo-reality to correct the other time series to in the calibration period.

For temperature, Figure 15a), all time series show a trend in the mean temperature in the range of 2.2 to 2.7 °C. These are within the range of global temperature increase of the A1B scenario, with 2.3 - 3.4 °C [IPCC]. The uncorrected RCM time series, lying on the diagonal, do show different trends for the mean temperature. The DMI-HIRHAM5 shows the smallest trend of 2.2 °C, whereas the SMHIRCA has the largest trend of 2.7 °C. The question is how these differences in the trends between the pseudo-realities affect the bias corrected time series?

It has to be kept in mind, that the influence of the chosen pseudo-reality is only taken from the calibration period. Whatever is present in the pseudo-reality later on, will not affect the bias corrected time series.

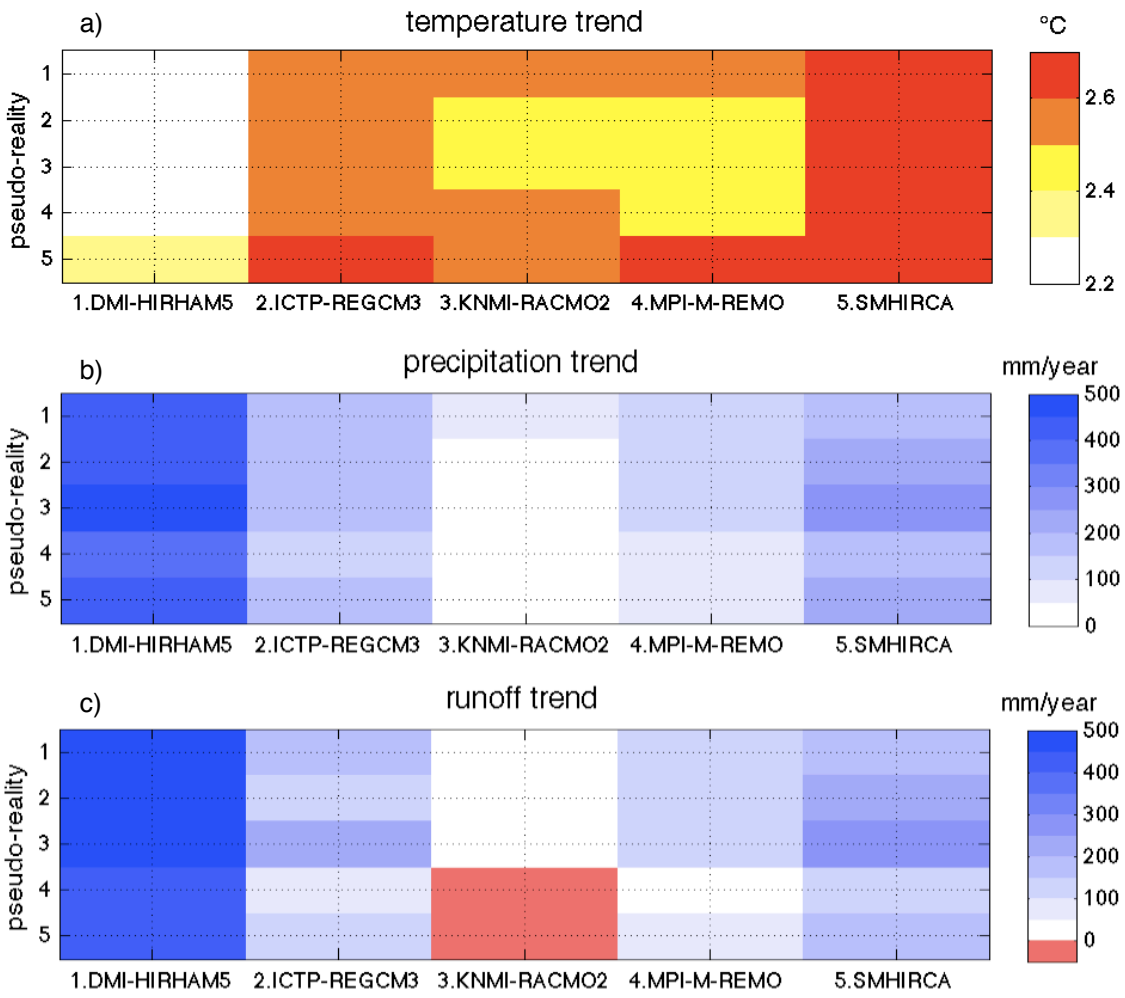


Figure 15: Changes in mean annual mean temperature in °C, mean annually accumulated precipitation in mm/year and mean annually accumulated runoff in mm/year for the uncorrected (diagonal) and bias corrected time series. The number on the y-axis gives the RCM time series that was used to bias correct the other time series to.

In general a vertical structure is seen in Figure 15a), indicating that the different trends for the temperature time series of the same RCM (all values in one column) show a higher resemblance than all the trends of time series, that were bias corrected to the same pseudo-reality (all values in one row). This is best seen for the two cases of the DMI-HIRHAM5 and the SMHIRCA: For the DMI-HIRHAM5 the trend was altered only for the time series corrected to the SMHIRCA. The time series of the SMHIRCA show different trends for the cases that ICTP-REGCM3 and KNMI-RACMO2 are the chosen pseudo-realities. This is an unexpected result because in all three cases, where the trend was altered the temperature values of the pseudo-reality in the calibration period were not extremely high or low. This indicates that the alteration of trends is independent of the values present in the calibration period. It can be noted, that only in the case of the

SMHIRCA temperature trends were reduced compared to the uncorrected RCM run, for all other cases, if the temperature trend of the time series was altered, it was enhanced compared to the uncorrected RCM time series. Additionally, the differences in the trends of the RCM time series are in the order of 0.1 °C, which is small compared to the trends in the different RCM time series, which are in the order of 2.5 °C.

For the trends in the mean annually accumulated precipitation, Figure 15b), the values range from 25 mm/year for KNMI-RACMO2 to 500 mm/year to DMI-HIRHAM5, which accounts for an increase of 16% to the total annually accumulated precipitation compared to the calibration period. Again a more vertical structure is seen in the pattern, indicating that the trends in the uncorrected RCM time series are not strongly altered compared to the bias corrected time series: For DMI-HIRHAM5 the trend was slightly decreased for the case it was bias corrected to the MPI-M-REMO pseudo-reality, which has the lowest mean annually accumulated precipitation values in the calibration period. For the KNMI-RACMO2 time series the bias correction slightly increased the trend in the case when the DMI-HIRHAM5 was the pseudo-reality, which is the RCM with the highest mean annually accumulated precipitation value in the calibration period. It seems that the precipitation trends are influenced by the amount of precipitation present in the pseudo-reality in the calibration period. Whenever a trend in mean annually accumulated precipitation is increased (decreased) the corresponding pseudo-reality had higher (lower) precipitation amounts than the uncorrected RCM in the calibration period. However, the changes between the time series of one RCM are small compared to the trend of the RCM time series in question.

For the trends in the mean annually accumulated runoff, Figure 15c), the pattern and also the total values are similar to the change in the mean annually accumulated precipitation. The values range between -50 to 500 mm/year, and are largest for the time series of the DMI-HIRHAM5 and smallest for the KNMI-RACMO2.

The largest difference between the two parameters is seen for the time series of the KNMI-RACMO2, where negative runoff trends occur for the case that the MPI-M-REMO and the SMHIRCA are the used pseudo-realities. Note, that the runoff trends of the bias corrected time series are negative although the uncorrected RCM time series does show a positive runoff trend.

The negative runoff trends arise from the combination of higher temperatures trends than present in the uncorrected RCM time series and only a small increase in precipitation. This means the amount of precipitation available in both periods is the same, but higher temperatures occur. Higher temperatures result in a higher rate of evapotranspiration and

snow/glacier melt. However, the amount of snow available to be melted did not change, hence the decrease of runoff in the future period arises from an enhanced evapotranspiration due to the higher temperatures. This effect is present in all the time series for the future period. However, it becomes only visible if the precipitation trend does not compensate the loss of water. In general, the trends in runoff are dominated by the trends in precipitation. However, in cases when precipitation change is small, the temperature trends become important.

5. Interpretation and Discussion

5.1 Interpretation of the Results within the Reference Period

The analysis carried out in the reference period gave a first impression on the performance of the two methods. From the histograms, the shape of the quantiles and the accumulated precipitation values it is seen, that those features are addressed and well corrected by the quantile mapping bias correction. So the quantile mapping bias correction is able to correct biases in the RCM time series concerning the statistical parameters and is able to reconstruct the shape of the observed precipitation intensities. This is consistent with the findings of Gudmundsson et al. [2012] and Piani et al. [2010] who also validated the quantile mapping bias correction within a reference period. However, a more meaningful validation is achieved if the quantile mapping bias correction is discussed in the pseudo-reality setting, which will be discussed in the next section.

For the VGLM mixture model bias correction the histograms and the shape of the observed precipitation intensities as well as the accumulated precipitation values are better represented for the summer time series compared to the winter. The bias is better reduced and the overall shape of the observed precipitation values is better represented. This can be attributed to the fact that in summer the variance in the observed time series is much lower than in the winter season. Figure 17 shows the daily variance in the observed time series, of the three different stations in a) and for the catchment time series in b).

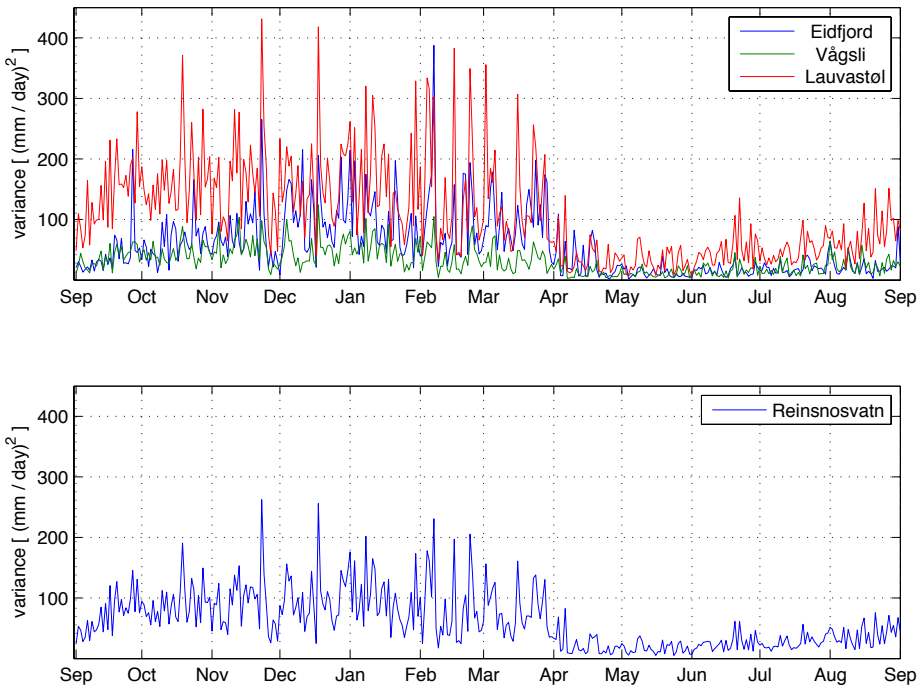


Figure 17: The daily seasonal cycle of precipitation variance over the period from 1961 to 1999 for the three station time series a) and the catchment time series b).

In winter the daily seasonal cycle of the precipitation variance is higher in all time series with values ranging between 50 (mm/day)² for Vågsli to 200 (mm/day)² for Lauvastøl, compared to summer with values ranging from 10 (mm/day)² for Vågsli to 100 (mm/day)² for Lauvastøl. This leads to the assumption that the VGLM mixture model bias correction works more accurate if the observed time series has a lower variance, because if the variance of the time series itself is small, the unexplained variance is also smaller. Therefore, the VGLM mixture model shows a better representation of the observed shape of the precipitation intensities for time series with a lower total variance. For Lauvastøl, where the variance is highest during summer, the fit of the VGLM mixture model is less accurate than for the other summer time series, see Figure 7.

Furthermore, for Lauvastøl winter precipitation intensities between 4 and 13 mm/day the values are overestimated by the fitted VGLM mixture model bias corrected time series, compare Figure 7. This can be explained, by the poor fit of the simple gamma distribution for this data set. Figure 18 shows the histogram of the observed time series for Lauvastøl and the fitted gamma distribution for winter and summer and as a comparison the histogram of the winter time series for Vågsli with a fitted gamma distribution.

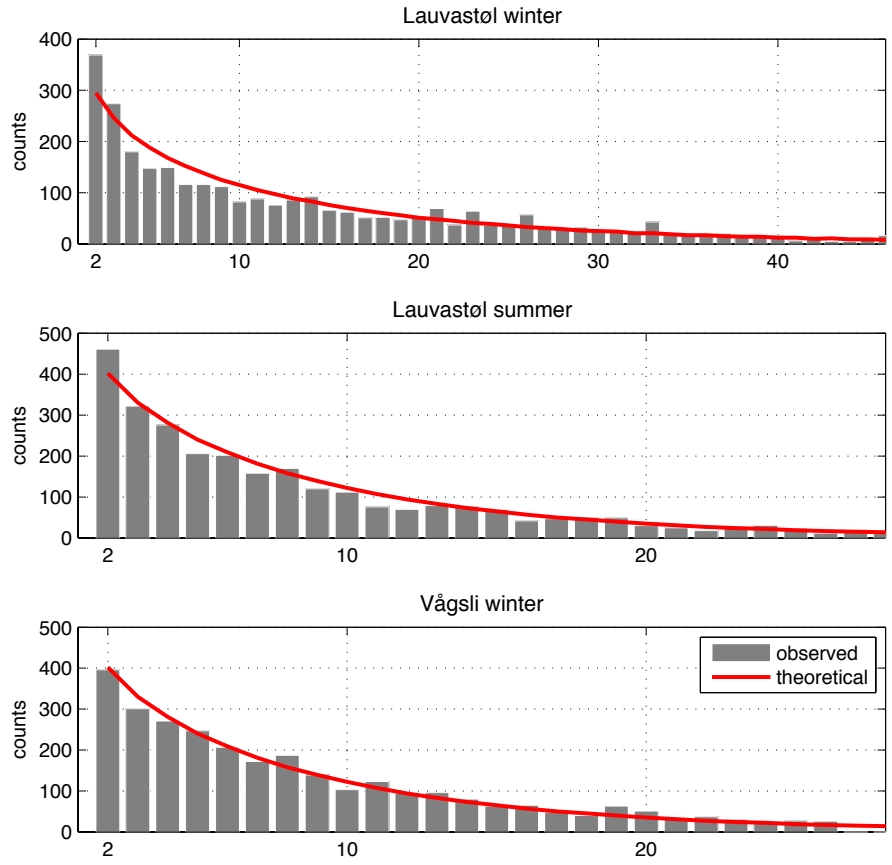


Figure 18: Histograms of the observed time series (grey bars) and the fitted gamma distributions (red line) for Lauvastøl in winter a) and summer b) in the reference period. As a reference the histogram of Vågsli in winter is shown in c).

It is seen, that for Vågsli the data is much closer to the theoretical gamma distribution, whereas for Lauvastøl in both seasons the observed time series seems not to follow the shape of the theoretical gamma distribution fitted to the data. This indicates, that for certain precipitation data sets the gamma distribution does not represent the shape of the observations well and there may be a need for another distribution function.
For the highest quantile of the simulations from the VGLM mixture model, the strongest overestimation is found for Lauvastøl, especially for the summer season, compare Figure 8. This might arise from the fact, that for Lauvastøl the observed time series experiences the highest number of high precipitation intensities. For the winter season a threshold of 80 mm/day is passed 8 times for the observed time series of Lauvastøl, only once for Eidfjord and Reinsnosvatn and in the observed time series of Vågsli that high a value is not reached. For summer a similar behaviour is seen, a threshold of 55 mm/day is passed in the observed time series of Lauvastøl 6 times, for Eidfjord the threshold was passed once and for the other two time series do not reach this value at all.
The sensitivity of the extreme value distribution to fit such very high values seems to be high within the model structure of the VGLM, and might give better results for a lower sensitivity.

For the dry day probability, the dry spells spectrum and the HSS, the two correction methods show very different results: For the total number of dry days in the time series, both correction methods show a good correction.
From the dry spells spectrum however it can be seen, that the representation of the observed temporal structure of the quantile mapping bias corrected time series is more sporadic and not consistently well. As an exception, the temporal structure of the time series of Reinsnosvatn is very well represented in the quantile mapping bias corrected time series for both, winter and summer. A similar result is found for the HSS, where an improvement of the HSS compared to the uncorrected time series is found for the quantile mapping bias corrected time series of the three station time series. However, an almost perfect reproduction of the temporal structure is found for Reinsnosvatn within both seasons.
This indicates, that quantile mapping bias correction does not attempt to correct the temporal structure in the RCM time series to fit the observed, and is only able to reconstruct the observed dry spells spectrum if the uncorrected time series already shows a good correspondence. In the case that the RCM is driven by a GCM, as necessary to

asses future climate changes, the temporal structure will not be corrected by the quantile mapping bias correction.

For the VGLM mixture model bias correction, the dry spells spectra show a far too high representation of dry spells of one day duration. The total amount of dry days is however well represented. This indicated that the logistic model lacks the ability to model consecutive dry days. The HSS shows a similar result indicating, that compared to the uncorrected time series, an improvement is achieved, but a general good agreement is not reached.

The fact, that the information for the logistic model is taken from the uncorrected RCM time series might lead to a worse representation of the observed temporal structure, if the RCM is not driven by reanalysis data. An improvement could be reached, if a dependency of the dry day occurrence is linked to the dry day occurrence of the day before, so including an autoregressive model. This could be derived from the observational time series.

5.2 Interpretation of the Results from the Pseudo-reality

The pseudo-reality setting gave interesting results for the quantile mapping bias correction, allowing a more thorough validation of this method.

For the comparison of the mean values in temperature, precipitation and runoff, a strong deviation from the chosen pseudo-reality is found, for some of the models. Is that an inability of the quantile mapping bias correction to cope with future time series, or can this be explained otherwise?

For mean annual mean temperature, see Figure 11, the strongest deviation from the pseudo-reality is found for the time series of the DMI-HIRHAM5 with strong negative values and the time series corrected to the DMI-HIRHAM5 with positive temperatures compared to the pseudo-reality. This can be explained when Table 7 is considered, where it is seen, that the DMI-HIRHAM5 shows by far the lowest temperature values in the calibration period as well as in the validation period and in addition the lowest trend of all the RCM time series.

When the mean annual mean temperature values of the bias corrected RCM runs in the validation period are compared to their pseudo-reality, the DMI-HIRHAM5 time series, the positive values indicate, that within the calibration period the mean values were corrected to fit the low DMI-HIRHAM5 temperature value, however, since all of the four RCM time series, that were bias corrected show higher trends, see Figure 15, the resulting temperature in the validation period is higher than the one of the DMI-HIRHAM5.

For the time series of the DMI-HIRHAM5 bias corrected to the other four RCM time series, the negative values are caused, due to the small trend present in the time series which is not altered by the bias correction, see Figure 15. Although the mean values within the calibration period were increased to fit the pseudo-reality value, the small trend in the DMI-HIRHAM5 time series inhibited the time series to reach as high values as the pseudo-realities have within the validation period. Hence, the negative values for all of the bias corrected DMI-HIRHAM5 time series relative to the corresponding pseudo-reality.

A similar explanation can be found for the time series of the SMHIRCA. All the bias corrected time series of the SMHIRCA show higher temperature values than their corresponding pseudo-reality, this might be explained through the strong temperature trend in the time series, see Figure 15. This strong positive trend combined with the correction of the mean temperature values in the calibration period, where the mean annual mean temperature of the SMHIRCA lies e.g. 1 °C below the one of the MPI-M-REMO causes the overall higher temperature values. The lower temperature values of the time series bias corrected to the SMHIRCA can be explained in a similar way. The bias

correction did not change the trend of the temperature time series, and in the calibration period the mean values were adjusted to fit the one of the SMHIRCA. However, since the pseudo-reality shows the largest trend, no other bias corrected RCM time series reaches as high temperature values in the validation period.

For the mean annually accumulated precipitation values, the DMI-HIRHAM5 time series show the strongest deviations relative to their pseudo-reality, with a higher amount of precipitation. The time series, that were bias corrected to the DMI-HIRHAM5, show less precipitation relative to the pseudo-reality.

For the time series of the DMI-HIRHAM5 higher mean annually accumulated precipitation amounts are found than for the corresponding pseudo-reality because of the strong trend of mean annually accumulated precipitation within DMI-HIRHAM5 time series, see Figure 15. After the mean values in the calibration period were adjusted to fit the pseudo-realities, the strong trend in the DMI-HIRHAM5 time series, which was not altered by the bias correction, caused these high precipitation values compared to the pseudo-realities, which all show much smaller trends. Therefore, the largest difference is found for the case in which the KNMI-RACMO2 time series is the chosen pseudo-reality: It has the highest mean annually accumulated precipitation in the calibration period, Table 7, and shows the lowest trend in the time series, see Figure 15.

For the time series corrected to the DMI-HIRHAM5 the mean values in the calibration period were adjusted, but none of the bias corrected time series exhibit as strong a trend and hence do not reach the highest values within the validation period, which is seen for the DMI-HIRHAM5, see Table 7.

For the bias corrected KNMI-RACMO2 time series lower values than those of the corresponding pseudo-realities are evident. This can be explained by adjustment of the time series by the bias correction in the calibration period, where the mean annually accumulated precipitation values are strongly decreased to fit the value of the corresponding pseudo-reality, see Table 7. However the KNMI-RACMO2 time series show nearly no trend from the calibration period to the validation period, hence the bias corrected time series do not reach as high values as the pseudo-realities within the validation period. The strongest difference is seen for the DMI-HIRHAM5, where the trend in the pseudo-reality is strongest.

For the time series bias corrected to the KNMI-RACMO2 higher mean annually accumulated precipitation values are seen. This is explained by the increase of the precipitation values in the calibration period to fit the ones of the pseudo-reality, see

Table 7. In addition the bias corrected time series show a higher trend than the KNMI-RACMO2 and therefore reach higher mean annually accumulated precipitation values within the validation period. The same argument can be followed for the differences in the other time series.

The mean values of temperature and precipitation within the validation period can be explained by the trends of the RCM that is bias corrected and the mean values in the calibration period of the pseudo-reality. This shows that the differences of the mean values in the validation period are not biases introduced by the quantile mapping bias correction method, but rather are artefacts from the original time series. The mean values in the validation period for a given time series is hence the product of the differences in the mean value in the calibration period relative to its pseudo-reality and the trend of the given time series:

$$\begin{aligned} T_{mean,val,bc} &= diff\,(T_{mean,cal},\ T_{mean,val,ps}) + trend \\ &= diff\,(T_{mean,cal},\ T_{mean,cal,ps}) + diff(\ T_{mean,val}, T_{mean,cal}\ ) \\ &= T_{mean,val} - T_{mean,cal,ps} \end{aligned} \tag{12}$$

The pattern of the runoff closely follows the precipitation pattern, it makes therefore sense to discuss only the differences in the two parameters, because the main part can be explained by the additional precipitation available to the hydrological model. Figure 19a) shows the difference between the runoff and precipitation relative to the pseudo-reality. Again it becomes evident, that the main differences occur for the DMI-HIRHAM5 and the KNMI-RACMO2 time series, or if these are the chosen pseudo-reality.
In order to explain this behaviour the other parameters of the HBV model were regarded. Figure 19b) shows the mean annually accumulated evapotranspiration in the validation period relative to the corresponding pseudo-reality. This parameter has a very similar patter than the difference between the runoff and the precipitation. With a reduced evapotranspiration of 40 mm/year it accounts for 66% of the water that is available for the runoff in addition to the enhanced precipitation. For time series that were bias corrected to the DMI-HIRHAM5 the evapotranspiration is overestimated by the bias corrected time series by about 40 mm/day, which causes a lower runoff in these model runs, because there is less water as input available. For the time series of the DMI-HIRHAM5, which were bias corrected to other pseudo-realities the evapotranspiration is underestimated,

which indicates a surplus of the water amount that is available in the hydrological model. This parameter is strongly connected to the temperature, which shows the same pattern for the DMI-HIRHAM5 time series, where higher temperatures causes higher evapotranspiration rates.

For the KNMI-RACMO2 time series there is a small increase of the evapotranspiration, which does not explain the difference in runoff and precipitation, some other mechanism has to explain the difference of 30 mm/year.

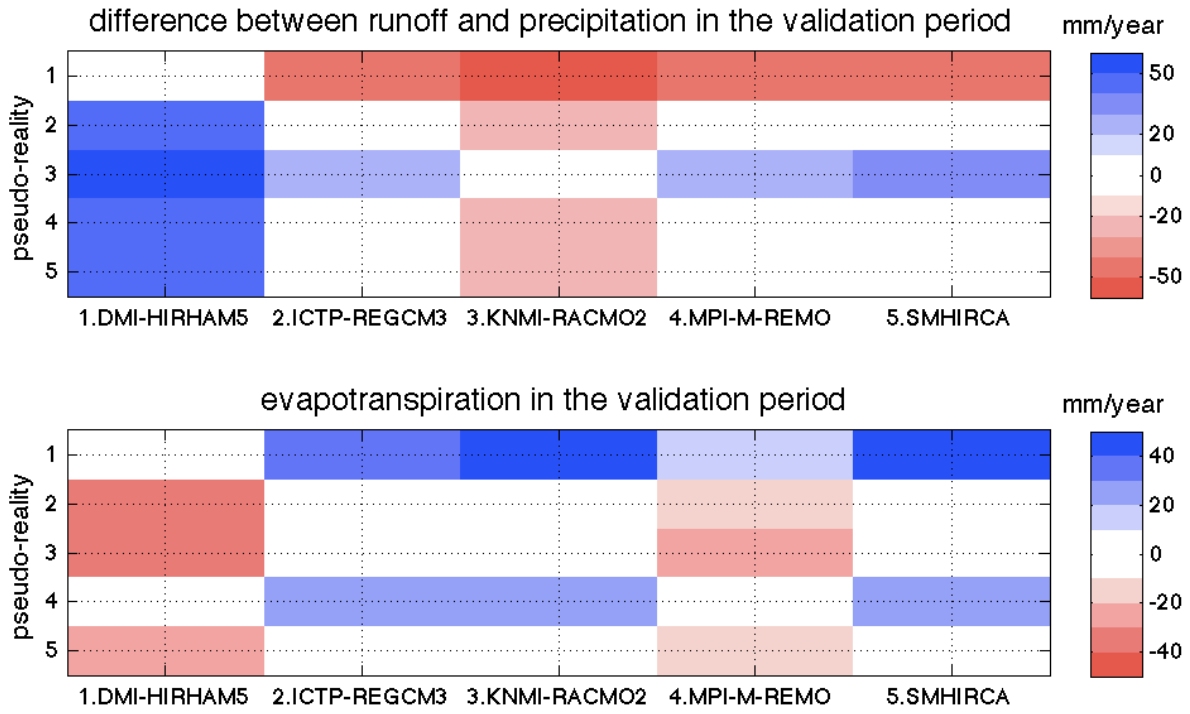


Figure 19: a) Difference between the relative runoff and precipitation of the bias corrected time series to their pseudo-reality in the validation period. b) Differences in mean annually accumulated evapotranspiration in mm/year for bias corrected time series in the validation period relative to the pseudo-reality they were corrected to. The number on the y-axis gives the RCM time series that was used to bias correct the other time series to, since it relative values are regarded, zero values on the diagonal are expected.

The results for the analysis of the annual cycle of runoff give a good insight in the changes of the temporal structure of the bias corrected time series. As a reminder, the bias correction in the pseudo-reality setting was carried out with monthly data. This becomes evident when Figure 14a) is considered: All of the bias corrected time series show the shift of the spring flood at the same time of the year as the corresponding pseudo-reality. There is however a spread in the amplitude of the change within the bias corrected time series, which arises from the fact that for all time series of one RCM the amplitude stays the same, see Figure 14b). Hence, the amplitudes of a bias corrected RCM time series remains unchanged, but the time of the change is altered by the bias correction.

6. Conclusion and Outlook

From the analysis in the reference period a understanding of the VGLM mixture model bias correction method is gained. This method has the huge potential to bias correct RCM time series. For the simulations a range for future scenarios is gained. The method deals with unexplained variance and is able to correct dry day occurrences. At this stage of the model development, the method is especially successful, if the observed time series follows a gamma distribution and has a low variance. In the further development, the logistic model should be improved by implementing an autoregressive process. In addition, the gamma distribution could be replaced by a more flexible probability distribution, which would be able to fit the observed data even if it did not follow a gamma distribution.

As soon as the method is fully developed, it could also be tested within a pseudo-reality setting, to test its ability to reconstruct future trends.

It has to be kept in mind, that running the HBV model with the 100 simulations from the VGLM mixture model will be at a large expense, therefore the analyses should be carried out for precipitation first. If these results are satisfactory, the results of the HBV model will give already a range of future scenarios, which gives an impression of the uncertainty within the RCM time series.

Within the reference period, the quantile mapping bias correction approach shows a good representation of the observed histograms and the accumulated precipitation amounts. All these features are forced to fit the observations by construction. However, this bias correction approach does neither implement a specific dry day correction nor does it deal with the unexplained variance present in the observed time series, hence it is unable to downscale RCM grid box average time series to a local scale. This becomes evident, when the catchment time series is calculated from quantile mapping bias corrected stations, here the catchment time series does no longer show the observed features. This indicates the inability of the quantile mapping bias correction method to distinguish between features observed at a local scale and the areal average of the RCM grid box. Therefore, the observed time series variance will be inflated. If the catchment time series, as an areal mean value is directly bias corrected, much better results are seen especially for the representation of the dry days. This is in agreement with Maraun [2013], who also recommends to avoid the downscaling step to local scale, but to rather directly bias correct the catchment time series from the RCM grid box time series.

Within the pseudo-reality setting, the inability of the quantile mapping to correct the dry day structure becomes clearly evident, since the total number of dry days within the bias corrected time series does not show a clear dependency on the pseudo-reality or the uncorrected RCM time series, see Figure 12. If a future RCM time series was corrected using an observed time series within the calibration period, the dry day occurrences for the validation period would not be well corrected by the quantile mapping bias correction.
For the seasonal cycles of the runoff, the change from the calibration to the validation period was well reconstructed regarding the point of time, however, the amplitude remained the one from the uncorrected RCM time series. If the method would be applied using observations, the temporal structure of the bias corrected time series would resemble the one from the observations and the amplitude would remain the amplitude of the uncorrected RCM, as long as the bias correction approach is applied on monthly data subsets.
A positive outcome of the analyses is the fact that the general trend present in the RCM time series is not crucially changed by the application of the quantile mapping bias correction. This refutes the concerns Maraun [2013] has expressed concerning the effect the quantile mapping bias correction might have on trends within the RCM time series. In this analysis longer time series of almost 90 years are considered compared to Maraun [2013], who looked at a time span of 30 years, and it was found that for the considered time period, the trends of the bias corrected time series were similar to the uncorrected RCM trends.

To further investigate these methods, at first a general analysis of the representation of the extremes would be suggested for the reference time period. In this context the Bries Score, the HSS or the extreme dependency score could be applied to the bias corrected time series. For the extremes a good performance is expected from the VGLM mixture model, since it includes an extreme value distribution. However, in order to validate the performance for the extremes on a profound level, the extreme value theory should be studied, before applying any of the scores with an insufficient understanding of the material.
Additionally, within the pseudo-reality setting other bias correction methods or the delta change method could be implemented and compared to the performance of the quantile mapping bias correction. It is expected, that this setting will point out other difficulties of these methods and will give a more thorough understanding of how the methods work.

Acknowledgments

At first, I would like to thank my family for their support in all areas not only during my master thesis but also throughout my whole studies.
I want to thank my two supervisors, Douglas Maraun and Katja Matthes, for their help and advise in both professional and personal questions, to my colleagues who answered all of my in-between questions. Especially I would like to thank Geraldine Wong, who helped me with the implementation of the probabilistic method and gave me a fresh heart to continue my thesis.
A special thanks goes to Eli Alfnes, from Statkraft, who tutored me during my time in Norway, awaked my interest in the topic and helped me with various questions during my writing time.
I want to thank Statkraft for providing the observed data, without them I could not have carried out my analysis.
Finally I would like to say thank you to my friends who had to bear me and supported me throughout the whole time, especially to Kerstin Kretschmer, Elisabeth Frenken and Florian Schütte.

References

Akaike, H. (1974), A new look at the statistical model identification, *IEEE Transactions on Autommatic Control*, 19, 716–723.

Bronstert, A., V. Kolokotronis, D. Schwandt, and H. Straub (2007), Comparison and evaluation of regional climate scenarios for hydrological impact analysis: General scheme and application example, *International Journal of Climatology*, 27 (12), 1579–1594.

Christensen, J.H., B. Hewitson, A. Busuioc, A. Chen, X. Gao, I. Held, R. Jones, R.K. Kolli, W.-T. Kwon, R. Laprise, V. Magaña Rueda, L. Mearns, C.G. Menéndez, J. Räisänen, A. Rinke, A. Sarr and P. Whetton (2007), Regional Climate Projections. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. *Cambridge University Press*, Cambridge, United Kingdom and New York, NY, USA.

Christensen, J. H., F. Boberg, O. B. Christensen, and P. Lucas- Picher (2008), On the need for bias correction of regional climate change projections of temperature and precipitation, *Geophysical Research Letter*, 35, L20709, doi:10.1029/2008GL035694.

Doolittle, M. H. (1888), Association ratios, *Bulletin of Philosophical Society of Washington*, 7, 122-127.

Engen-Skaugen, T. (2007), Refinement of dynamically downscaled precipitation and temperature scenarios, *Climate Change*, 84:365-382, doi: 10.1007/s10584-007-9251-6

Frigessi, A., O. Haug, and H. Rue (2003), A dynamic mixture model for unsupervised tail estimation without threshold selection, *Extremes*, 5, 219 – 235.

Gudmundsson, L., J. B. Bremnes, J. E Haugen and T. Engen-Skaugen (2012), Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations–a comparison of methods, *Hydrology and Earth System Sciences*, *16*, 3383-3390, doi: 10.5194/hess-16-3383-2012.

Heidke, P. (1926), Berechnung des Erfolges und der Güte der Windstärkenvorhersage im Sturmwarnungsdienst, Geogr. Ann., 8, 310–349

IPCC, 2007: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. *Cambridge University Press*, Cambridge, United Kingdom and New York, NY, USA, 996 pp.

Landgren, O.A., T. Engen Skaugen, J. E. Haugen and E. J. Førland (2012), An evaluation of temperature and precipitation from global and regional climate models over Scandinavia, *Norwegian Meteorological Report*, no. 21a/2012, 1503-8025

Lindström G., B. Johansson, M. Persson, M. Gardelin and S. Bergström (1997), Development and test of the distributed HBV-96 hydrological model, *Journal of Hydrology*, 201, pp 272-288, doi:10.1016/S0022-1694(97)00041-3

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

Maraun, D. (2012), Nonstationarities of regional climate model biases in European seasonal mean temperature and precipitation sums, *Geophysical Research Letter*, 39, L06706, doi: 10.1029/2012GL051210.

Maraun D. (2013), Bias Correction, Quantile Mapping, and Downscaling: Revisiting the Inflation Issue, *Journal of Climatology,* 26, 2137–2143, doi:10.1175/JCLI-D-12-00821.1

McCullagh, P., and J. A. Nelder (1989), Generalized linear model, *Chapman & Hall/CRC*, Vol. 37.

Norwegian Meteorological Institute, http://met.no/English/Climate_in_Norway/

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

Räisänen J. and O. Räty (2012), Projections of daily mean temperature variability in the future: cross-validation tests with ENSEMBLES regional climate simulations, *Climate Dynamics,* 0930-7575, doi:10.1007/s00382-012-1515-9

Sælthun, N. R. (1995), „Nordic“ HBV model, *Norwegian Water Resources and Energy Administration*, ftp://terra.uio.no/pub/dagrunv/HBVMOD.PDF

Teutschbein, C. and J. Seibert (2012), Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods, *Journal of Hydrology*, 456–457, 12–29, doi:10.1016/j.jhydrol.2012.05.052.

Themeßl, M. J., A. Gobiet, and A. Leuprecht (2010), Empirical-statistical downscaling and error correction of daily precipitation from regional climate models, *International Journal of Climatology,* 31, 1530-1544, doi:10.1002/joc.2168.

Vrac, M., and P. Naveau (2007), Stochastic downscaling of precipitation: From dry events to heavy rainfalls, *Water Resources Research*, 43, W07402, doi:10.1029/2006WR005308.

Wilks, D. S. (2005), Statistical Methods in the Atmospheric Sciences, *Academic Press, Elsevier LTD, Oxford*, 3rd revised edition

Wong, G., D. Maraun, M. Widmann, J. Eden and T. Kent (2013), A stochastic model output statistics approach for correcting and downscaling precipitation including extremes, Draft

Yang, W., J. Andréasson, L. P. Graham, J. Olsson, J. Rosberg and F. Wetterhall (2010), Distribution-based scaling to improve usability of regional climate model projections for hydrological climate change impacts studies, *Hydrology Research*, 41.3-4, doi:10.2166/nh.2010.004

Yee, T. W. and A.G. Stephenson (2007), Vector generalized linear and additive extreme value models, *Extremes*, 10:1-19, doi: 10.1007/s10687-007-0032-4

Appendix

A. Dry Day Correction

Table A.1 and A.2 give the effect of the applied dry day threshold on observations and the RCM time series for the reference period for winter and summer, respectively. The observed time series used to fit the RCM data to is then only the dry day corrected.

Station		Original TS [mm/year]	Dry day corrected TS [mm/year]	dry day threshold [mm/day]
Eidfjord	RCM	571,3	545,6	2,4
	OBS	516,7	514,2	1
Vågsli	RCM	770,7	723,2	4,1
	OBS	371,9	367,0	1
Lauvastøl	RCM	765,9	756,5	1,2
	OBS	716,3	716,3	1
Reinsnosvatn	RCM	605,9	597,5	1,0
	OBS	573,1	568,7	1

A.1: Effect of the dry day correction on observed and RCM winter mean accumulated precipitation values.

Station		Original TS [mm/year]	Dry day corrected TS [mm/year]	dry day threshold [mm/day]
Eidfjord	RCM	408,2	355,9	3,5
	OBS	224,6	220,1	1
Vågsli	RCM	617,8	511,9	6,5
	OBS	213,5	207,6	1
Lauvastøl	RCM	499,8	473,7	2,1
	OBS	401,0	401,0	1
Reinsnosvatn	RCM	309,5	304,5	2,1
	OBS	393,0	365,1	1

A.2: Effect of the dry day correction on observed and RCM summer mean accumulated precipitation values.

B. Akaike Information Criterion

The Akaike information criterion (AIC) is a measure of the relative quality of a statistical model, for a given set of data. The AIC deals with the trade-off between the complexity of the model and the goodness of fit of the model. As such, AIC provides means for model selection.

In the general case, the AIC is calculated as followed:

$$AIC = 2k - 2\ln(L)$$

where k is the number of parameters in the statistical model, and L is the maximised value of from the likelihood function for the estimated model.

Given a set of candidate models for the data, in this case the six parameters that can be fixed, the preferred model is the one with the minimum AIC value.

The AIC does not only reward goodness of fit, within the $2\ln(L)$ part of the equation, but also includes a penalty, the $2k$ part of the equation, that is increasing the AIC depending on the number of estimated parameters. This penalty discourages over-fitting.

For the model selection of the VGLM mixture model, different parameters of the model structure were fixed and the AIC was calculated to see which model structure is favourable. Tables B.1 and B.2 show the winter and summer values of the AIC calculations, respectively, the coloured values were the minimum values and this model structure was used for the simulations.

No	Fixed parameter	Eidfjord	Vågsli	Lauvastøl	Reinsnosvatn
1	xi	11401,17	9734,34	14175,99	12989,14
2	xi,tau	11394,96	9728,55	14143,39	12987,14
3	xi, m, tau	11469,46	9745,00	14153,77	12999,64
4	xi, tau, gamma	11426,75	9727,83	14215,77	13005,64
5	xi, tau, lambda	11425,27	9740,69	14162,5	13011,12
6	xi, tau, sigma	11392,12	9727,51	14218,49	12990,16
7	xi, tau, m, lambda	11449,61	9749,97	14146,28	13015,28
8	xi, tau, m, sigma	11373,10	9735,64	14156,39	12960,86
9	xi, tau, m, gamma	11421,97	9735,46	14124,19	13014,43
10	xi, tau, lambda, gamma	11452,70	9761,47	14179,50	13103,94
11	xi, tau, lambda, sigma	11452,41	9729,79	14219,87	13031,24
12	xi, tau, gamma, sigma	11382,84	9730,30	14172,03	12989,42

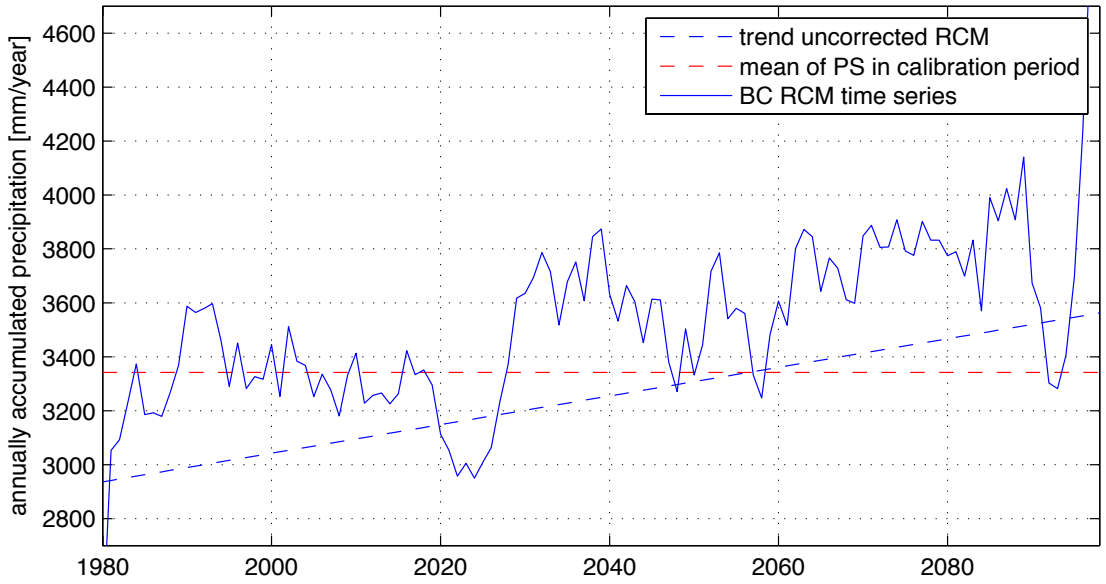
B.1: AIC values for the winter time series.

No	Fixed parameter	Eidfjord	Vågsli	Lauvastøl	Reinsnosvatn
1	xi	7442,17	6654,80	10471,35	9265,79
2	xi,tau	7440,17	6652,80	10496,64	9337,90
3	xi, m, tau	7370,57	6716,90	10469,37	9257,15
4	xi, tau, gamma	7438,19	6661,46	10469,16	9310,18
5	xi, tau, lambda	7409,35	6652,69	10454,89	9309,57
6	xi, tau, sigma	7438,26	6660,38	10533,71	9305,55
7	xi, tau, m, lambda	7365,09	6714,90	10532,65	9286,35
8	xi, tau, m, sigma	7436,26	6665,23	10524,77	9281,50
9	xi, tau, m, gamma	7404,72	6657,94	10467,49	9263,79
10	xi, tau, lambda, gamma	7438,25	6701,81	10525,78	9368,91
11	xi, tau, lambda, sigma	7436,23	6670,09	10497,39	9364,44
12	xi, tau, gamma, sigma	7364,14	6667,44	10500,88	9292,42

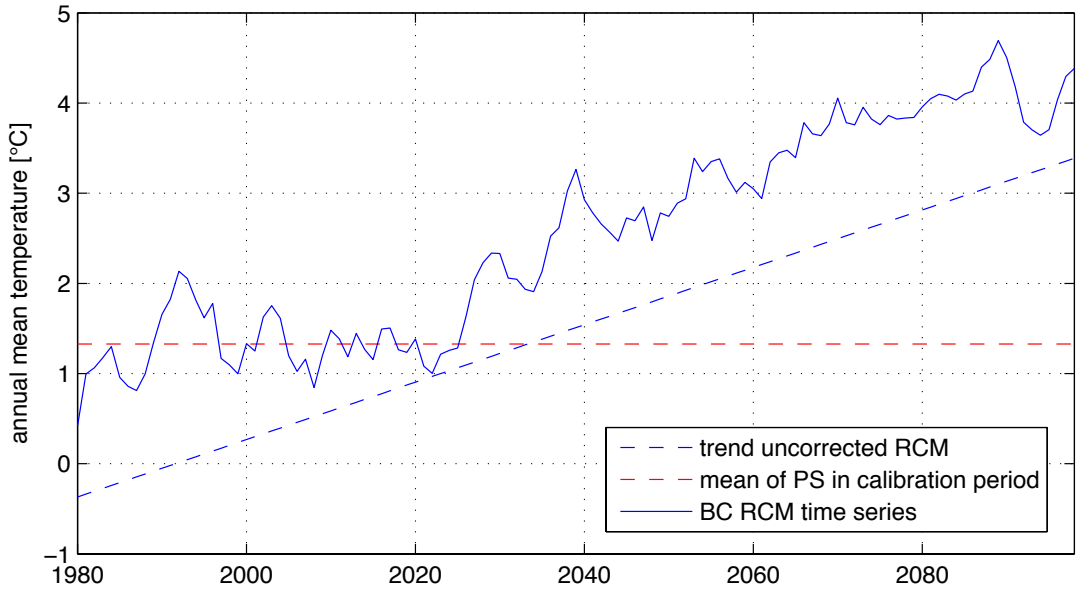
B.2: AIC values for the summer time series.

C. Schematic of Trend Conclusion

In order to illustrate the equation (12), Figure C.1 and C.2 show an exemplary bias corrected time series for precipitation and temperature, respectively.



C.1: Bias corrected annually accumulated precipitation time series of DMI-HIRHAM5 in the case, when corrected to the KNMI-RACMO2 (blue line). The dashed red line shows the mean annually accumulated precipitation of the KNMI-RACMO2 in the calibration period and the dashed blue line gives the trend in the uncorrected DMI-HIRHAM5 time series.



C.2: Bias corrected annual mean temperature time series of SMHIRCA in the case, when corrected to the MPI-M-REMO (blue line). The dashed red line shows the mean annual mean temperature of the MPI-M-REMO in the calibration period and the dashed blue line gives the trend in the uncorrected SMHIRCA time series.

Erklärung

Hiermit erkläre ich, Nadine Mengis, dass ich die vorliegende Arbeit selbständig und ohne fremde Hilfe angefertigt und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Die eingereichte schriftliche Fassung der Arbeit entspricht der auf dem elektronischen Speichermedium. (Name der Datei: 908102)
Weiterhin versichere ich, dass diese Arbeit noch nicht als Abschlussarbeit an anderer Stelle vorgelegen hat.

_____

Datum, Unterschrift