CH2011 Extension Series No. 4

Bias-corrected transient scenarios at the local scale and at daily resolution

Sven Kotlarski [1,2], Martin Ivanov [2,3], Christoph Schär [2]

[1] Federal Office of Meteorology and Climatology MeteoSwiss, Switzerland, [2] Institute for Atmospheric and Climate Science, ETH Zurich
[3] Department of Geography, Climatology, Climate Dynamics and Climate Change, Justus-Liebig University of Giessen

---

C2SM Center for Climate Systems Modeling

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Home Affairs FDHA
Federal Office of Meteorology and Climatology MeteoSwiss

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

NCCR CLIMATE
Swiss Climate Research

OcCC
Organe consultatif sur les changements climatiques
Beratendes Organ für Fragen der Klimaänderung

Abstract

A new set of bias-corrected daily climate scenarios for 28 temperature and 27 precipitation stations in Switzerland is presented. The scenarios are based on the quantile mapping (QM) methodology, are available for 15 GCM-RCM model chains of the ENSEMBLES project and provide transient time series for the entire period 1980–2099. An extensive validation exercise provides strong confidence in the quality and robustness of the new QM-based scenarios. They are complementing the delta change-based local daily scenarios of CH2011 and provide an added value with respect to future changes in temporal climate variability and in extremes. They pave the way towards the next release of the Swiss national climate scenarios CH2018 to be released in 2018.

Reviewers

Ole Roessler (Institute of Geography, University of Berne), Jan Rajczak (Institute for Atmospheric and Climate Science, ETH Zurich)

Acknowledgements

This work was supported by the Swiss State Secretariat for Education, Research and Innovation SERI under project number C12.0089. We acknowledge the RCM data sets from the EU-FP6 project ENSEMBLES http://ensembles-eu.metoffice.com and the Swiss Federal Office of Meteorology and Climatology MeteoSwiss for providing observational data. Furthermore, we thank Dr. Jan Rajczak, ETH Zurich, for his valuable and extremely helpful input regarding the application of QM over the Swiss territory.

Recommended citation

Kotlarski, S., M. Ivanov, and C. Schär, 2017: Bias-corrected transient scenarios at the local scale and at daily resolution. *CH2011 Extension Series No. 4,* Zurich, 21 pp.

Addresses

Center for Climate Systems Modeling (C2SM)
ETH Zürich, CHN
Universitätsstrasse 16
CH-8092 Zürich

Institute for Atmospheric and Climate Science
ETH Zürich, CHN
Universitätsstrasse 16
CH-8092 Zürich

Federal Department of Home Affairs FDHA
Federal Office of Meteorology and Climatology MeteoSwiss
Operation Center 1
P.O. Box 257
CH-8058 Zurich-Airport

Download

Extension Article and Scenario Data
→ www.ch2011.ch

Contact

info@ch2011.ch

Bias-corrected transient climate scenarios are presented for 28 temperature and 27 precipitation sites in Switzerland and for 15 GCM-RCM chains of the ENSEMBLES project

The new scenarios are complementing the existing delta-change based daily local scenarios of CH2011

They provide an added value with respect to future changes in temporal climate variability and in extremes and pave the way towards the upcoming CH2018 scenarios

1| Background and Motivation

The CH2011 Swiss Climate Change Scenarios (CH2011 2011) provide climate scenario products that are widely used by the Swiss climate impacts community (e.g. CH2014-Impacts 2014) and that are ultimately relevant for policymaking in the frame of the Swiss national adaptation strategy (Bundesamt für Umwelt BAFU 2012). Despite their documented usefulness and applicability for a large range of climate impact assessments, the CH2011 scenarios at daily resolution are subject to certain limitations which partly relate to the fact that all CH2011 products are based on the delta change (DC) approach. DC provides climate change signals that can be used to scale observed mean conditions or observed time series in a historical reference period in order to obtain climate conditions valid for a future scenario period (e.g. Bosshard et al. 2011).

The change signals are obtained by comparing climate model output for the historical and the future period. Advantages of DC are a comparatively easy and straightforward implementation and the fact that the historical reference period is entirely based on observations and, hence, can be assumed to be realistic. Furthermore, the derived scenario conditions mimic the (observed) characteristics of the reference period to a large extent, for instance with respect to spatio-temporal climate variability and inter-variable relations (such as the inter-dependency between temperature and precipitation). The derived scenario products are therefore unlikely to suffer from strong distortions and inhomogeneities of these patterns.

The DC approach, however, is based on a number of assumptions that are not necessarily fulfilled. As a consequence, DC cannot account for the full spectrum of possible climatic changes which partly restricts its applicability. Limitations of DC and, hence, of the CH2011 scenarios relate to:

(1) **The neglect of changes in temporal variability:**
In order to obtain time series valid for a future period, DC scales observational series of the historical period. As a consequence, the temporal variability of the future series mimics the variability in the observations to a large extent. For instance, the future sequence of wet and dry days will almost be identical to the observed sequence. The same is true for variabilities on larger temporal scales (seasonal, annual). If future climate change will be associated with changes in temporal variability these changes will not be reflected by DC-based scenario time series.

(2) **The restricted ability to capture changes in extremes:**
The DC method applied in CH2011 scales the observed time series by the climate change signal of the 30-year mean value, which can considerably differ from the climate change signal of extremes. Potential changes in extreme conditions are therefore not (or only partly) reflected by the DC approach.

(3) **The assumption of stationary climate model biases:**
DC approaches are based on the comparison of climate model output for a historical reference and a future scenario period. The obtained climate change signals implicitly assume model biases that are similar in both periods. Non-stationary model biases which are likely to be present (e.g. Buser et al. 2009; Bellprat et al. 2013) and that would influence the derived climate change signal are not accounted for.

(4) **The non-transient setup with discrete scenario periods:**
Standard DC approaches provide climate change signals for discrete scenario time slices and with respect to a specific historical reference period. The CH2011 scenarios, for instance consider the three scenario periods 2020–2049, 2045–2074 and 2070–2099 with respect to the reference interval 1980–2009. Impact assessments that require transient meteorological input or that target scenario periods not considered by the specific DC implementation have to interpolate between the DC time slices provided, which involves assumptions about the evolution of climate change and climate variability in the non-covered periods (e.g. Farinotti et al. 2012).

The mentioned limitations of DC can be partly overcome by explicit bias correction (BC) approaches. Instead of scaling observations by a climate model-derived climate change signal, BC explicitly corrects for systematic biases in transient climate model output. This correction is based on a comparison of climate model output and observations within a common historical reference period and the establishment of a correction function to translate biased climate model output into de-biased transient time series. Similar to DC, BC can implicitly include a downscaling step to derive climate conditions representative for the site scale. Note, however, that BC can never correct for all kinds of climate model biases (see Chapter 5) which is the reason why BC is nowadays also referred to as bias adjustment.

The present CH2011 Extension provides bias-corrected temperature and precipitation scenarios at daily resolution for several sites in Switzerland based on the ENSEMBLES regional climate projections (van der Linden and Mitchell 2009). It thereby complements the CH2011 daily local scenarios (Bosshard et al. 2011). The BC scenarios are based on results obtained in the frame of the ELAPSE project *(Enhancing local and regional climate change projections for Switzerland)* which was funded by the Swiss State Secretariat for Education, Research and Innovation SERI, and on research that has been conducted in relation to precipitation scenarios (e.g. Rajczak et al. 2016b). Within ELAPSE several bias correction methods, namely a large number of variants of the quantile mapping (QM) methodology, were evaluated with respect to their reliability and robustness when applied over the Swiss territory. The ELAPSE project itself was closely related to the recent COST-Action VALUE *(Validating and Integrating Downscaling Methods for Climate Change Research; Maraun et al., 2015; www.value-cost.eu)* that aims at inter-comparing and evaluating statistical and dynamical climate downscaling methods.

The following Chapter provides an overview on the QM methodology and on the specific setups that were evaluated in the frame of ELAPSE. Chapter 3 then presents selected evaluation results. Based on one particular QM implementation, BC scenarios for several Swiss sites were produced which are presented and compared to the respective DC-based CH2011 product in Chapter 4. Chapter 5 provides details on how to use the new BC scenarios and on inherent limitations, before summarizing and concluding this report in Chapter 6. Note that Chapters 3 and 4 only provide a concise summary of the results obtained within the ELAPSE project. More details can be found in Ivanov et al. (2015) and Ivanov and Kotlarski (2017).

2| Quantile Mapping

2.1 Overview

In recent years, the quantile mapping (QM) method became increasingly popular to both correct for systematic biases in climate model simulations and to bridge the scale gap between coarse resolution climate model output and the site scale. Originating from the empirical transformation of Panofsky and Brier (1968), QM corrects for biases in the distribution of a simulated variable (the predictor) by comparing the model output against the observational distribution (the predictand; see the illustrative example in Figure 1). The matching of both distributions is achieved by establishing a correction function that transforms simulated quantiles into their observed counterparts. This correction function is then applied to each member of the simulated time series yielding a bias-corrected series with a distribution similar (or identical) to the observed one. The implicit assumption is that the model can predict ranked categories of the variable of interest, i.e. quantiles, but not its actual values (Déqué 2007). Simulated and observed quantiles can either be based on the full empirical distribution (non-parametric implementation; e.g. Déqué 2007; Themessl et al. 2012) or on a fitted theoretical distribution (parametric implementation, e.g. Piani et al. 2010).

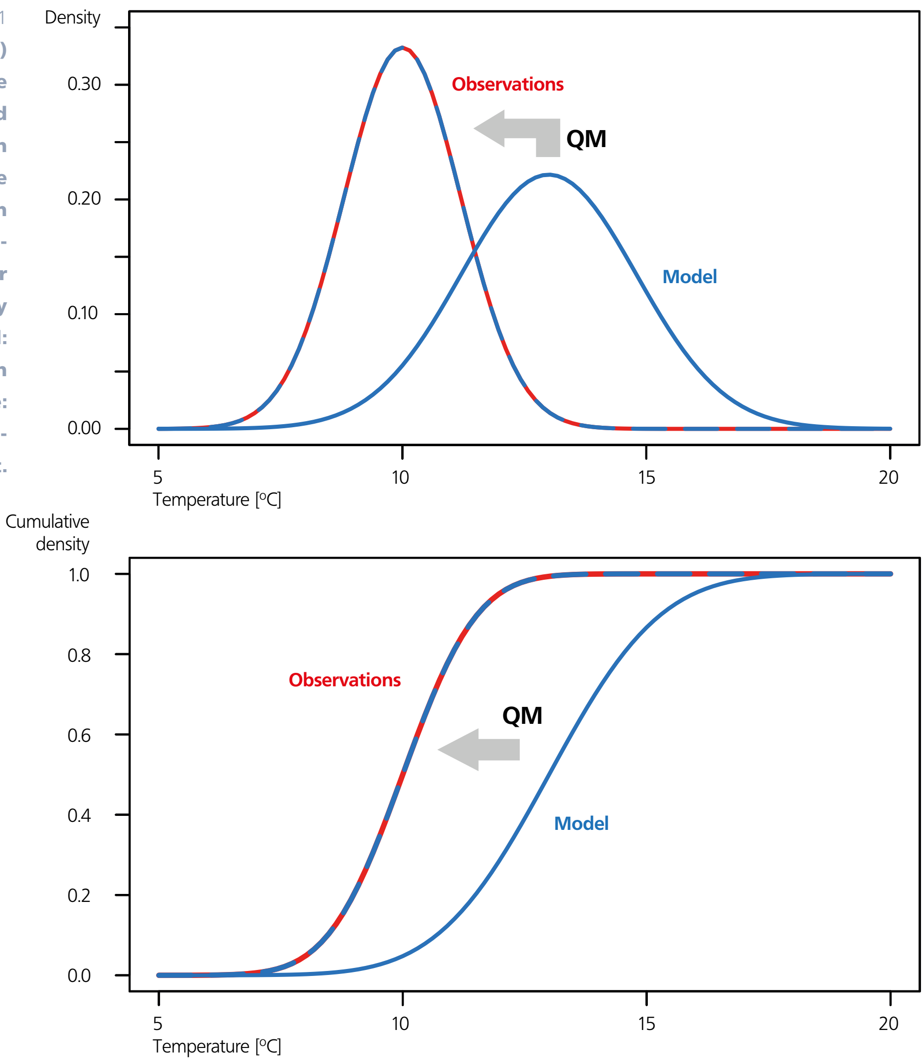


Figure 1
Illustration (idealized) of the QM method: The distribution of a simulated variable (here: daily mean temperature; solid blue line) is corrected to match the observed distribution (solid red line). Upper panel: Probability density function. Lower panel: Cumulative distribution function. Dashed blue line: Distribution of the bias-corrected model output.

If the observational reference against which QM is calibrated reflects the same spatial scale as the raw climate model output (which is the case if gridded observations of the same spatial resolution and reflecting grid cell area mean values are used), QM has to be considered as a mere bias correction. Most applications, however, implicitly include a downscaling component by calibrating QM against observations that reflect the site scale, i.e., against station measurements. This additional downscaling component is an attractive property of QM, but also afflicted with specific problems and limitations (see below). In this context QM can be classified as a MOS (Model Output Statistics) methodology that directly relates modelled predictors to their observed counterparts (Maraun et al. 2010) and that can be used to bias-correct and downscale coarse resolution global or regional climate model output in order to provide climate scenarios at the site scale. This application of QM in a climate scenario context is illustrated in Figure 2: In a first step, climate model output for a historical period is compared against observations for the same period and a quantile-dependent correction function is established. This calibrated function is in a second step applied to the entire climate model time series, yielding bias-corrected (and possibly downscaled) time series for both the historical and the future scenario period. By definition, the distributions of the observed and the bias-corrected series match each other in the historical calibration period.

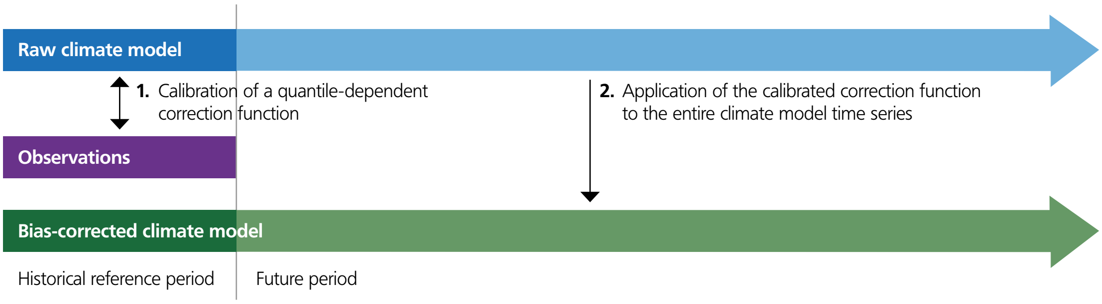


Figure 2
Application of QM in a climate scenario context. Step 1: Calibration of the correction function in a historical calibration period. Step 2: Application of the calibrated function in order to generate bias-corrected (and possibly downscaled) climate scenarios.

A large number of studies recently applied QM in a downscaling context and documented the method's general applicability and a performance that is often comparable or even superior to other empirical-statistical downscaling approaches (e.g. Boé et al. 2007; Themessl et al. 2011; Gudmundsson et al. 2012; Themessl et al. 2012). Several works explicitly applied QM as an interface between coarsely-resolved and potentially biased climate model output and subsequent climate impact models (e.g. Wood et al. 2004; Hagemann et al. 2011; Finger et al. 2012; Rajczak et al. 2016a). Although QM does not explicitly correct for a biased temporal variability in raw climate model output (Addor and Seibert 2014), Rajczak et al. (2016b) showed that for instance biases in wet day-dry day transition probabilities and in multi-day indices can be effectively removed. Furthermore, Wilcke et al. (2013) found that the separate application of QM to several meteorological variables retains inter-variable dependencies as represented by the climate model output (although it doesn't necessarily correct for biased inter-dependencies). Despite these promising results, issues concerning the applicability of QM in a downscaling context remain (see also Chapter 5). These include (1) non-stationarities of climate model biases (e.g. Buser et al. 2009; Maraun 2012; Bellprat et al. 2013) that can only partly be represented by QM, (2) a potentially distorted spatial climate variability (Maraun 2013), (3) a spurious influence on climatic trends due to variance inflation (Maraun 2013), (4) issues concerning the spatial representativeness of the underlying climate model output (Maraun and Widmann 2015), (5) a potentially dangerous application of QM in the presence of strong circulation biases in the underlying climate model output (Maraun and Widmann 2015), and (6) uncertainties in the calibrated correction function as a consequence of multi-decadal climate variability (e.g. Maurer et al. 2016).

2.2 Implementation for Switzerland

Within the ELAPSE project the applicability, robustness and the added value of QM with respect to DC-based approaches have been assessed for the Swiss territory. Both daily temperature and daily precipitation performance measures were considered. For this purpose 21 variants of QM were implemented, among them 20 non-parametric implementations. These variants differ with respect to (a) the derivation of distributional quantiles (linear interpolation between empirical percentiles versus direct use of all empirical data quantiles versus theoretical fit of the empirical distribution function) (b) the temporal resolution of the derived correction function (explicit consideration of each day of the year versus interpolation between the 12 centers-of-month) and (c) the treatment of new extremes (correction according to the 99th percentile versus correction according to the mean of all quantiles above the 99th percentile versus correction according to the maximum empirical quantile; for precipitation only: additive versus multiplicative correction). Additionally and for the case of precipitation only, a parametric implementation based on a mixture model that approximates the distribution of daily precipitation values by combining a Gamma and a generalized Pareto distribution (Frigessi et al. 2002) has been implemented. Furthermore, a correction of the mean temperature and precipitation bias, i.e. the simplest bias correction one can think of, has been considered and compared to the more complex QM methodology. We her refrain from listing details on the individual QM implementations but refer to Ivanov et al. (2015) and Ivanov and Kotlarski (2017) instead.

For each bias correction method the short-term stability of the correction function was tested in a cross-validation framework: The historical 40-year interval 1970–2009 was divided into the two sub-periods 1970–1989 and 1990–2009, each of which was corrected independently based on the calibration of the correction function over the other sub-period. Furthermore, the inter-variable consistency between the separately bias-corrected temperature and precipitation series has been assessed. The decadal-to-centennial scale stability of the bias correction methods was tested in a series of pseudo-reality experiments (e.g. Vrac et al. 2007). Here, the unknown and hence non-observed future "reality" (2070–2099) was provided by an individual climate model simulation which also served as calibration target in a historical calibration period (1980–2009) against which all other climate model simulations were calibrated.

Climate model data were provided by 15 transient regional climate scenarios of the ENSEMBLES project combining 11 different regional climate models (RCMs) with 6 different global climate models (GCMs) and assuming the SRES A1B emissions scenario (see Table 1). These 15 GCM-RCM chains provide output at a spatial resolution of about 25 km and correspond to the 14 chains used in CH2011 for the two later scenario periods (see Figure 2.4 in CH2011 (2011)) plus the additional chain DMI-BCM. As bias correction target, i.e. as local reference, homogenized daily temperature and precipitation time series for 28 stations of the Swiss National Basic Climatological Network (NBCN; Begert et al. 2007) were used (28 stations for temperature and 27 for precipitation; see Figure 3). These stations form a subset of those considered in the CH2011 local scenarios at daily resolution (Bosshard et al. 2011). In the cross-validation analysis the GCM-RCM data were spatially interpolated to station locations by means of an inverse distance weighting interpolation algorithm using the four nearest grid points, in accordance with CH2011 (2011) and Bosshard et al. (2011). In the pseudo-reality experiments the grid point that is closest to the respective station location has been extracted from each individual GCM-RCM chain which implies that pseudo-observations and model projections have the same spatial resolution (the one of the underlying GCM-RCMs). Hence QM in the pseudo-reality framework does not include an implicit downscaling step, but has to be considered as a pure bias-correction approach. Model chains having 360- or 365-day calendars were converted to a Gregorian calendar prior to the analysis by randomly adding missing values into the daily time series until the length of the respective Gregorian year is reached.

Based on the evaluation results and number of criteria (see below) one specific QM implementation has finally been selected to produce bias-corrected scenarios at daily resolution for all 28 NBCN stations and for all 15 GCM-RCM chains considered.

Figure 3

Location of the NBCN stations for which QM has been evaluated and for which bias-corrected climate scenarios were produced. Temperature: 28 stations (black and red). Precipitation: 27 stations (black and blue). The abbreviations refer to the unique station identifiers, see Ivanov et al. (2015) for details on the stations.

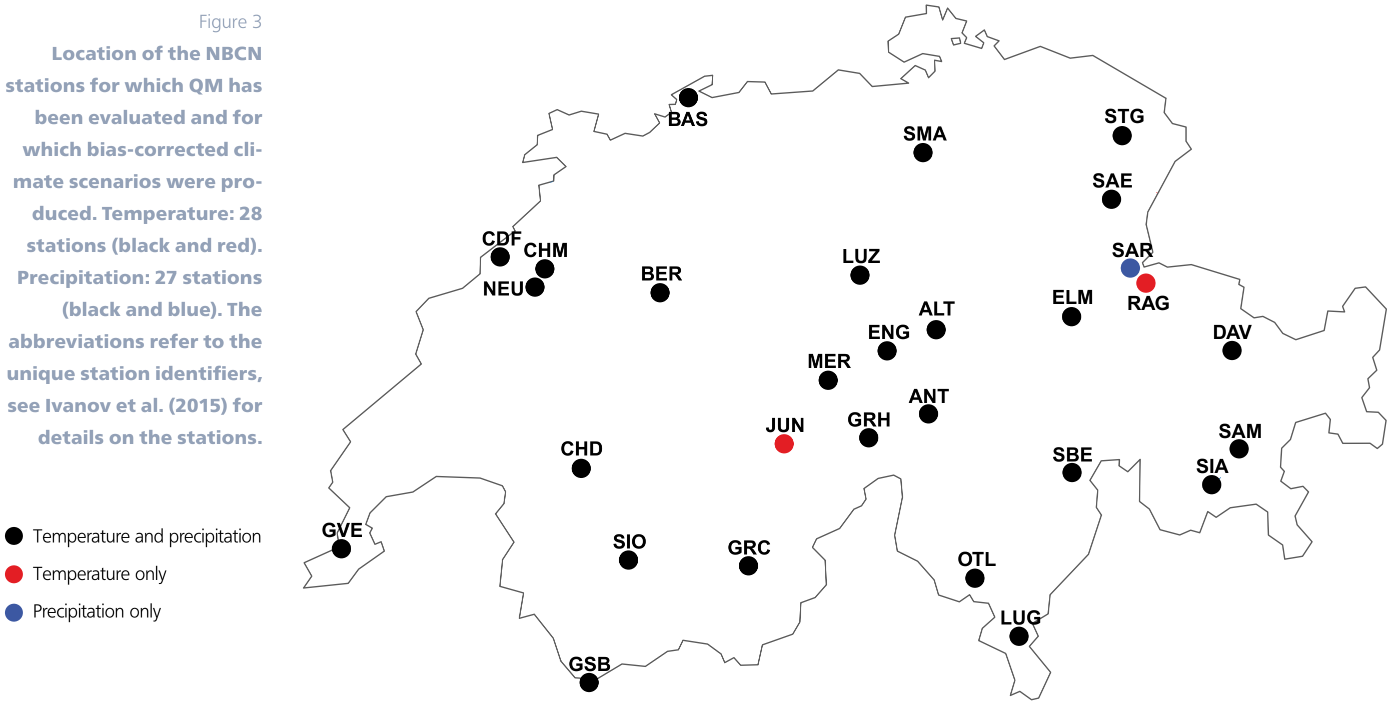


Table 1

List of the 15 employed GCM-RCM model chains from the ENSEMBLES project. The dark gray shading indicates the 10 model chains that were used for the CH2011 local scenarios at daily resolution (Bosshard et al. 2011).

Institution	GCM	RCM
CNRM	ARPEGE	Aladin
DMI	ECHAM5	HIRHAM5
ETHZ	HadCM3Q0	CLM
METO-HC	HadCM3Q0	HadRM3Q0
ICTP	ECHAM5	REGCM3
KNMI	ECHAM5	RACMO2
MPI-M	ECHAM5	REMO
SMHI	BCM	RCA
SMHI	ECHAM5	RCA
SMHI	HadCM3Q3	RCA
C4I	HadCM3Q16	RCA3
DMI	ARPEGE	HIRHAM5
DMI	BCM	HIRHAM5
METO-HC	HadCM3Q16	HadRM3Q16
METO-HC	HadCM3Q3	HadRM3Q3

3| Evaluation of the approach

We here present the results of the cross-validation exercise, the pseudo-reality evaluation and the assessment of inter-variable consistency. For reasons of brevity, no figures are shown for the two latter aspects and only a brief summary of the evaluation is provided. For details, the reader is referred to Ivanov and Kotlarski (2017). A range of evaluation metrics is considered, among them the seasonal mean bias, the bias of distributional percentiles and of return values as well as the bias of selected impact-relevant indices. Return values were estimated based on a generalized extreme value (GEV) distribution fitted to the seasonal block maxima (two well-separated daily maxima per season) by means of a modified maximum likelihood method using a Bayesian prior for the GEV shape parameter (cf. Frei et al. 2006). In all figures shown, the individual QM implementations are named according to the scheme *[CDF]_[EXTREMES]_[INT]* (see also Chapter 2). *[CDF]* describes the way in which the cumulative density function (CDF) of a given distribution is obtained, i.e., either by considering the individual quantiles *(step)* or by linearly interpolating between the percentiles *(linear)*. *[EXTREMES]* denotes the handling of new extremes that are corrected additively either according to the correction of the 99th percentile of the calibration period *(add)*, according to the mean correction for all empirical quantiles larger or equal to the 99th percentile *(addMean)* or according to the correction of the largest quantile of the calibration period *(addMax)*. Finally, the suffix *[INT]* denotes whether a temporal interpolation of the respective distributions between the 12 centers-of-month has been applied *(int)* or the correction function has been established separately for each DOY (empty). The method *mean* represents the simple correction of the mean bias (see above), *raw* denotes the raw RCM output. *Frigessi_int* is the parametric QM implementation for precipitation (mixture model; Frigessi et al. 2002). For further details the reader is referred to Ivanov et al. (2015) and Ivanov and Kotlarski (2017).

Figure 4 shows the cross-validation results for several temperature characteristics in the winter (DJF) and the summer (JJA) season. All QM implementations effectively decrease the partly substantial biases of the raw models (last row for each season) and, furthermore, are in most cases superior to a simple correction of the mean bias (second last row for each season). This is true for both distributional properties (percentiles, mean value, interquartile range) and estimated return values. Note that the raw models' biases include systematic deviations due to the scale mismatch between a climate model grid cell and a particular station, especially with respect to the reference elevation. Differences between the individual QM implementations are small. The methods that are only calibrated at the centres of the 12 calendar months and linearly interpolated in time *(int)*, are typically just as good as their daily-calibrated counterparts. Similar results are obtained for precipitation characteristics (Figure 5): The biases of the raw climate model output are effectively reduced by QM. Again, the individual QM implementations are close to each other and are in all cases superior to a simple correction of the mean bias. The parametric mixture model (uppermost row) typically shows a slightly worse performance compared to the non-parametric implementations. Besides distributional properties and estimated return values, QM is furthermore able to accurately represent further impact-relevant (multi-day) indices, see Figure 6 for selected examples. These results are in accordance with the recent study of Rajczak et al. (2016b) who cross-validated one specific QM implementation with respect to several precipitation indices.

Similar evaluation results are obtained for the pseudo-reality exercise. For reasons of brevity the respective figures are not shown in this report, but are presented and discussed in Ivanov et al. (2015) and Ivanov and Kotlarski (2017). The satisfying model performance in the pseudo-reality framework indicates a long-term stability of the quantile-based correction functions also beyond the historical climate. Furthermore, the analysis of the inter-variable consistency reveals that QM substantially improves the joint distribution of daily temperature and precipitation that can be considerably biased in the raw model output (not shown here as well; see Ivanov and Kotlarski 2017).

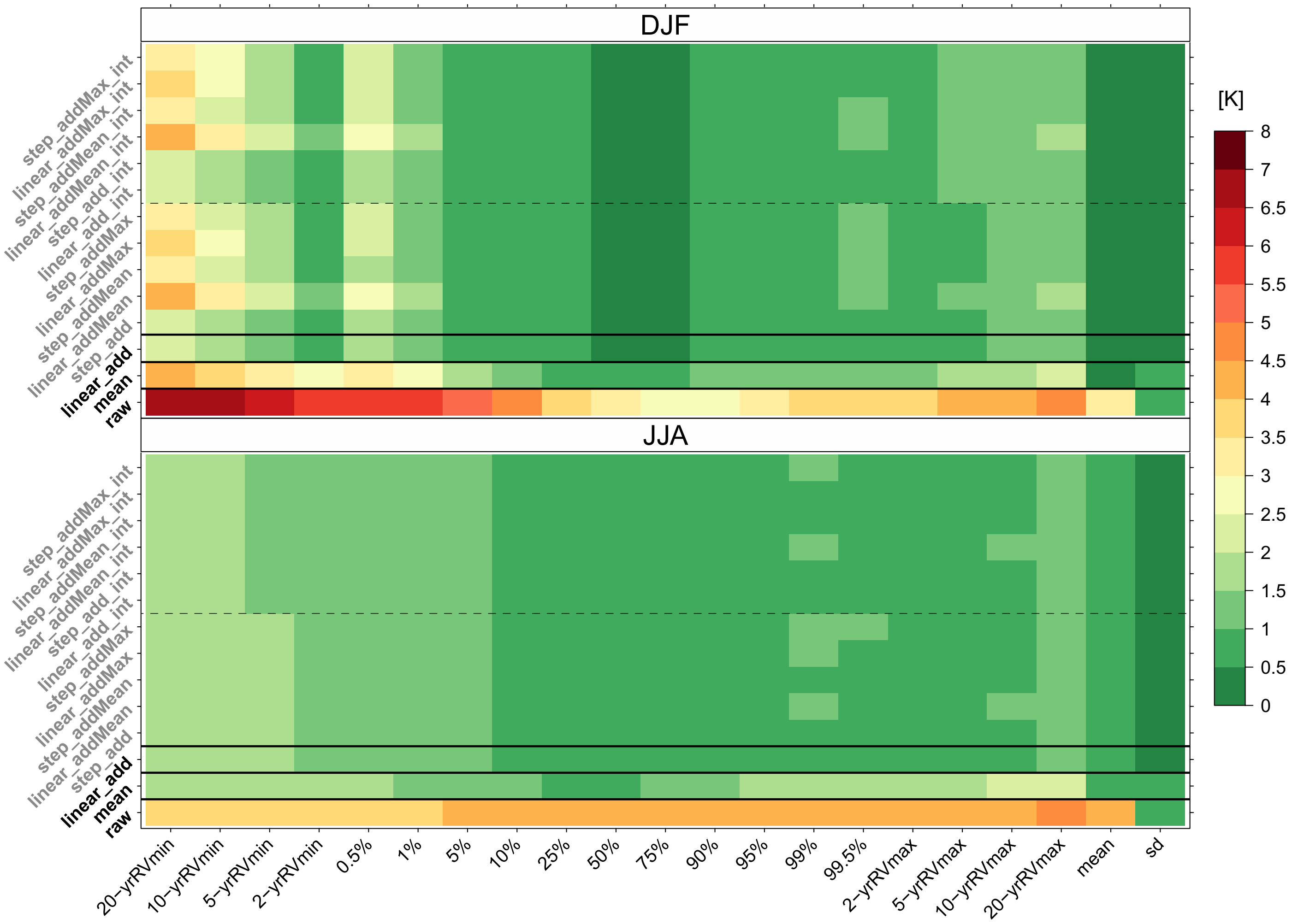


Figure 4

Mean absolute biases [K] of selected daily temperature characteristics in the cross-validation exercise, averaged over all GCM-RCM chains, stations, and the two cross-validation periods. X-axis: Temperature characteristics (RV: daily minimum and maximum temperature return values; X%: daily temperature percentiles; mean: mean temperature; sd: temperature standard deviation). Y-axis: Bias-correction methods (mean: simple correction of the mean bias; raw: uncorrected raw models). The methods above the dashed line (_int) are interpolated in time between the 12 centers-of-month. The linear_add method has finally been chosen for production of scenarios. Upper panel: winter (DJF), lower panel: summer (JJA).

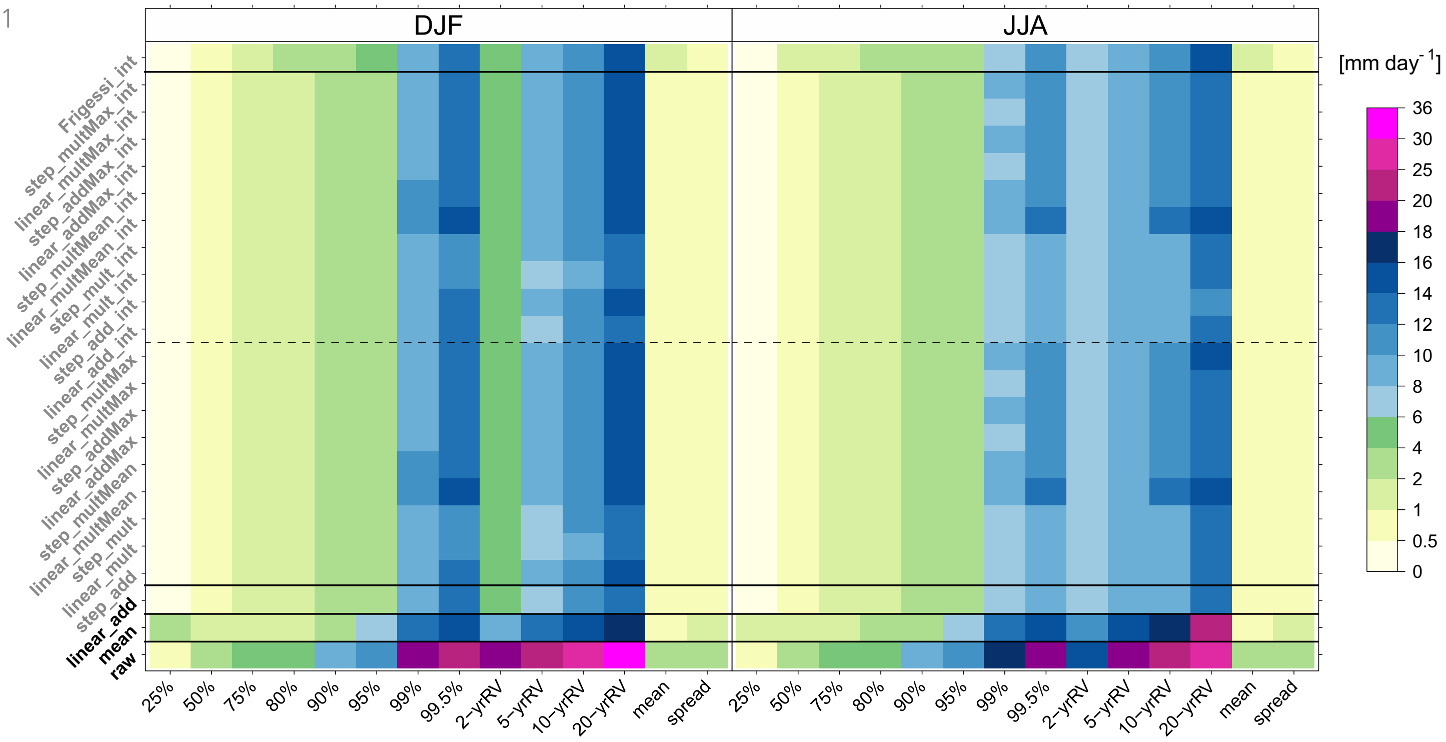


Figure 5

Mean absolute biases [mm day-1] of selected daily precipitation characteristics in the cross-validation exercise, averaged over all GCM-RCM chains, stations, and the two cross-validation periods. X-axis: Precipitation characteristics (X%: wet-day precipitation percentiles; RV: daily precipitation return values; mean: mean precipitation; spread: inter-quartile range). Y-axis: Bias-correction methods (Frigessi: parametric mixture model; mean: simple correction of the mean bias; raw: uncorrected raw models). The methods above the dashed line (_int) are interpolated in time between the 12 centers-of-month. The linear_add method has finally been chosen for production of scenarios. Left panel: winter (DJF), right panel: summer (JJA).

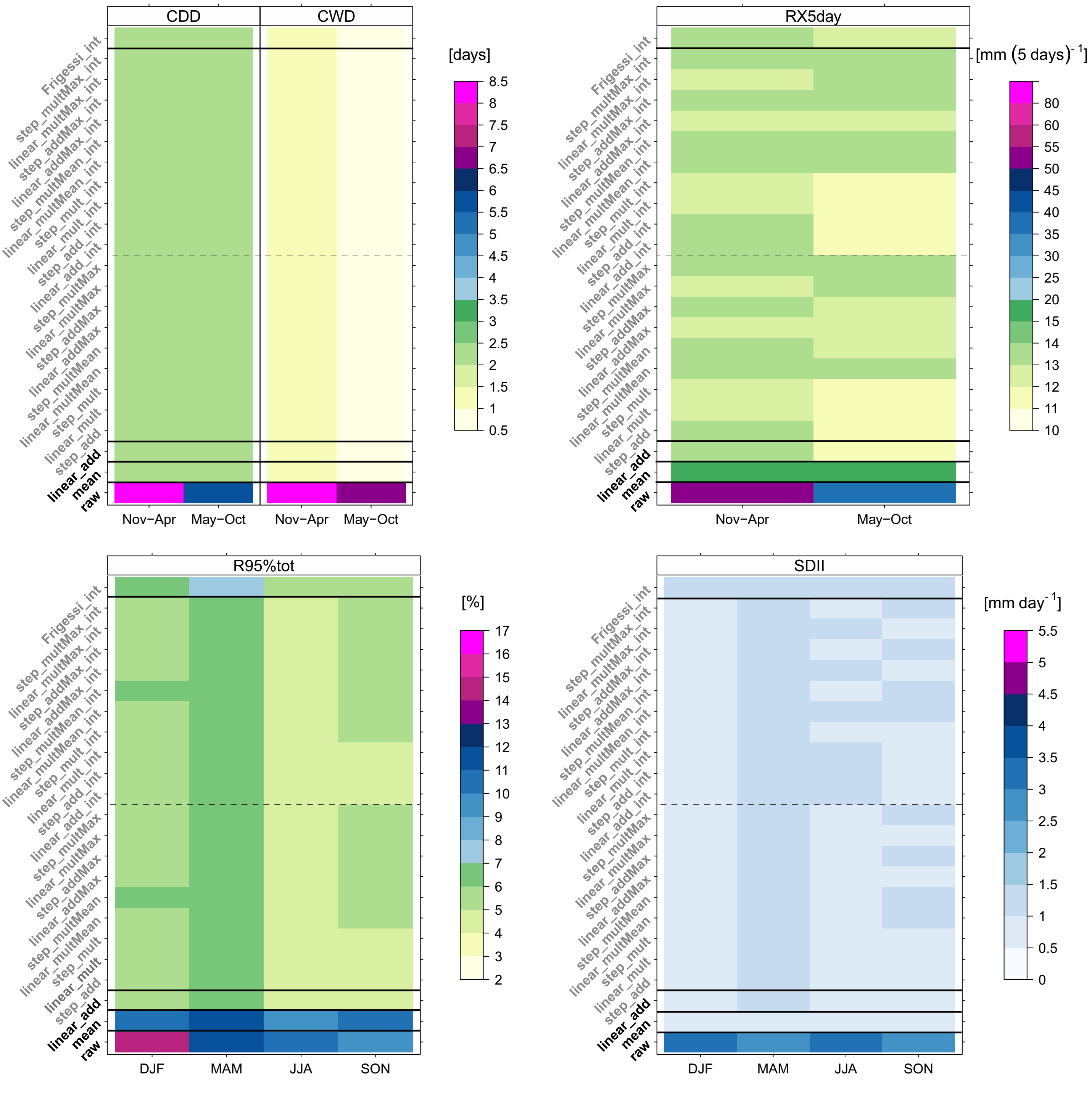


Figure 6

Absolute bias of precipitation impact indices in the cross-validation exercise, averaged over all GCM-RCM chains, stations, and the two cross-validation periods. X-axis: Season (winter: DJF, spring: MAM, summer: JJA, autumn: SON, cold season: Nov-Apr, warm season: May-Oct). Y-axis: Bias-correction methods. The linear_add method has finally been chosen for production of scenarios. Upper left panel: Maximum dry spell length (CDD) and maximum wet spell length (CWD). Upper right panel: Maximum 5-day accumulated precipitation (RX5day). Lower left panel: Precipitation fraction falling on very wet days above the 95th percentile (R95%tot). Lower right panel: Mean wet day intensity (Simple daily intensity index; SDII). Seasons: Winter (DJF), spring (MAM), summer (JJA), autumn (SON).

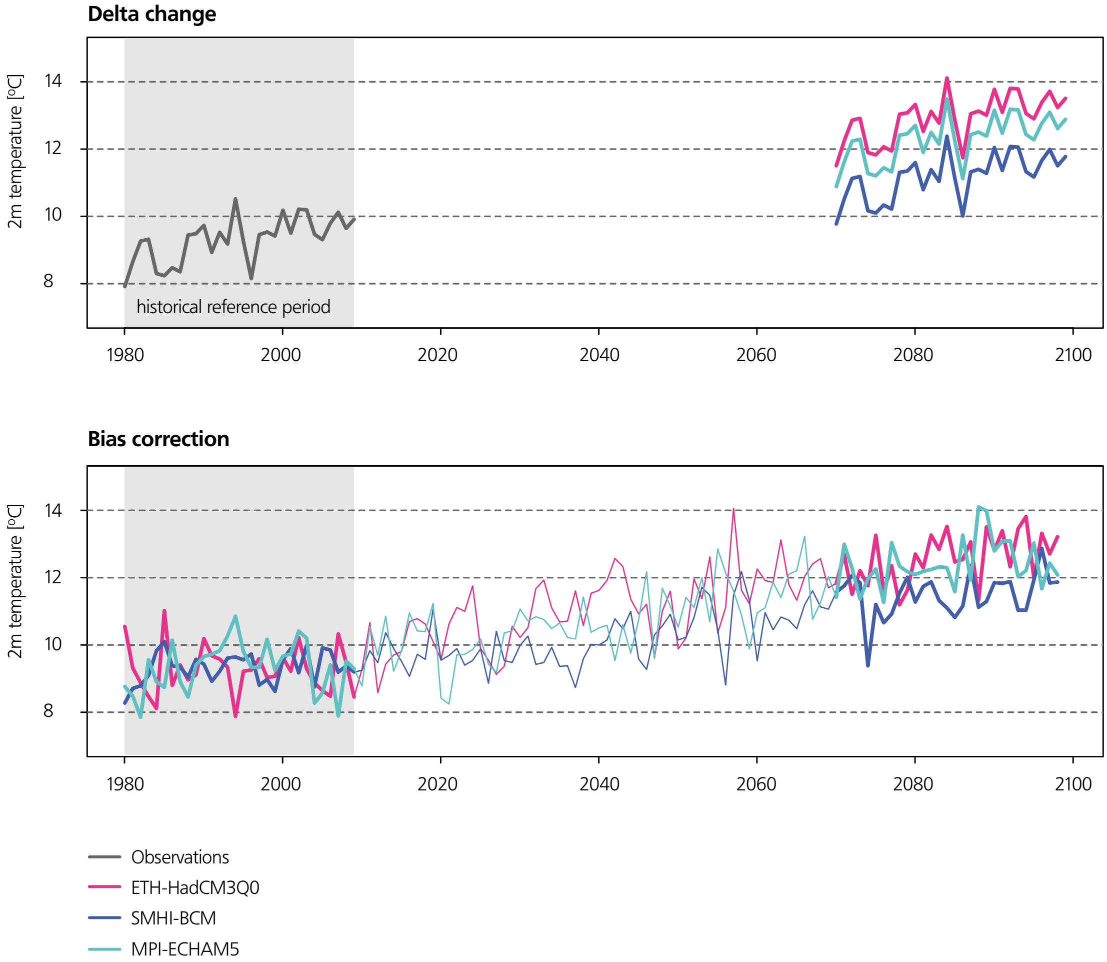
4| Bias-corrected versus Delta Change Scenarios

Based on the evaluation results summarized in the previous Chapter one specific QM implementation has finally been selected to produce bias-corrected local climate scenarios: **The non-parametric linear QM method, calibrated for each day of the year, with additive correction for new extremes according to the correction of the 99th percentile of the calibration period** (method *linear_add* in Figures 4, 5 and 6). The same methodology has also been applied in previous works (e.g. Boé et al. 2007). This choice was based on the performance in the cross-validation and pseudo-reality exercises as well as on the ease of implementation, including the required computational resources. Note that for many evaluation metrics, other methods reveal a similar performance, and could have been chosen as well. Using the *linear_add* method, transient bias-corrected local climate scenarios for the period 1980–2099 at daily resolution for temperature and precipitation were finally produced for the 28 Swiss NBCN stations considered and for all 15 ENSEMBLES model chains of Table 1. As calibration target for QM observations in the period 1980–2009, i.e. the same reference observations as used for the DC-based CH2011 scenarios, were chosen. We here present selected aspects of a comprehensive comparison against the DC-based CH2011 local scenarios and against climate change signals as represented by the raw GCM-RCM output. For this purpose, future time series for the period 2070–2099 were constructed/extracted for QM, DC and raw climate model output for each site, and climate change signals of selected indicators were computed with respect to the reference data in the historical period 1980–2009 (which, for the DC-based scenarios, are the observations themselves). In all three cases only the 10 GCM-RCM chains covered by the CH2011 daily local scenarios were considered (dark gray background in Table 1).

To illustrate the added value of the new bias-corrected scenarios, Figure 7 shows exemplary time series for both delta-change based and bias-corrected scenarios at the station of Zurich-Fluntern (SMA) for three selected GCM-RCM chains and for the case of annual mean temperature (i.e., the daily temperature scenarios were aggregated to annual mean values). The upper panel presents the delta change-based CH2011 scenarios. These are available for selected time slices only (here: reference period 1980-2009 and scenario period 2070-2099) and with gaps in-between. The historical observations serve as reference for all three model chains and, by definition, all three scenarios exactly mimic the temporal variability of these reference observations. The scenarios only differ with respect to the general increase of the temperature level which, in this case, is largest for ETH-HadCM3Q0 and smallest for SMHI-BCM. Possible changes in temporal variability between the reference and the scenario period are not accounted for. In contrast, the bias-corrected scenarios (lower panel; simple correction of the bias in annual mean temperature) can pick up such changes while maintaining the general level of temperature increase. The scenario time series are transient, i.e. without gaps, and each individual scenario reflects the temporal variability of the underlying GCM-RCM chain, providing a better account on internal climate variability in the scenario ensemble. In the historical reference period, all scenarios agree on the mean temperature level (as this period served as calibration period for the bias correction) but exhibit their individual temporal variability.

Figure 7

Illustration of the difference between delta change-based and bias-corrected climate scenarios for the case of annual mean temperature [°C] at the station Zurich-Fluntern (SMA) and for three GCM-RCM chains. Upper panel: Delta change-based scenarios for the period 2070–2099; note that the observations serve as historical reference for all three model chains. Lower panel: Transient bias-corrected scenarios for the period 1980–2099 (here: simple correction for the bias in annual mean temperature). In both panels the gray background denotes the historical reference period for which observations are available and which served as calibration period for QM.



For a quantitative comparison of the previous DC-based and the new QM-based scenarios and raw climate model output, Figure 8 shows the range of climate change signals for seasonal mean temperature, for the 1st (winter) and the 99th (summer) percentile of daily mean temperature and for the interannual variability of seasonal mean values. Both DC and QM approximately preserve the mean seasonal temperature changes as represented by the raw climate models. For DC this result is obtained by definition as the DC-based scenarios have been developed based on the raw models' change signals (Bosshard et al. 2011). QM slightly distorts the raw change signals, which can be explained by the quantile-dependent, i.e. temperature-dependent, correction function. Regarding changes in extreme temperature conditions, DC is not able to pick up the increase of the 1st percentile of winter mean daily temperature (i.e. the loss of moderate cold extremes) as simulated by the climate models. The climate change signal of this index is obviously associated with modifications of the lower tail of the daily temperature distribution which cannot be represented by the simple DC approach that simply relies on changes of the mean. QM, however, is able to partly pick this increase of the 1st percentile. A similar reasoning applies to changes of interannual temperature variability. The changes simulated by the underlying climate model chains (mostly decreases in winter and increases in summer; see also Fischer et al. 2012) are approximately represented, though partly modified, by the QM-based scenarios but not by DC. In the latter case, interannual variabilities remain unchanged as DC simply scales the observed time series of the historical reference period and cannot account for changes in temporal variability.

In the case of precipitation (Figure 9), both DC and QM approximately conserve the raw models' change signals (left panel). Note, however, that in a few cases QM shows stronger precipitation increases in winter than represented by the raw climate model output. This can partly be explained by a strong dependence of the correction function on the respective quantile considered. It is not possible to assess whether, in these cases, the raw and DC-based change signals or the QM-based scenarios are more realistic. Concerning changes in 99th percentile of daily precipitation both DC and QM represent the simulated increase in wintertime, but DC fails to reproduce the stable level in summer. In this season, the climate change signal for extreme precipitation qualitatively differs from the respective signal for mean precipitation (see also Rajczak et al. 2013). The latter decreases and as DC is based on changes in mean precipitation amounts and scales the precipitation amounts on every single day with the respective climate change factor, also extreme precipitation amounts decrease in DC. This feature is also illustrated by Figure 10 in a spatially explicit context: For most NBCN stations both the winter and the summer 99th percentiles either don't change or even increase in the raw model output (averaged over the 10 GCM-RCM chains considered in this comparison exercise; left column). In winter this increase – including its basic spatial pattern – is represented by both DC and QM (middle and rightmost columns). In summer, however, only QM is able to pick up the simulated change of extreme precipitation amounts while DC shows a decrease for the entire country corresponding to the decrease of mean summer precipitation.

DC is furthermore not able to reproduce the simulated changes in precipitation transition probabilities (Figure 9, right panel). The underlying GCM-RCM chains, for instance, simulate a decrease of the probability for a dry day to be followed by another dry day (p00) in winter and an increase in summer (Figure 8, right panel), mainly as a consequence of changes in the wet day frequency. These changes are represented by QM (see also Rajczak et al. 2016b), but DC shows no or only small changes in p00 as (1) the temporal sequence of the scenario time series is identical to the sequence in the scaled observational reference time series, and (2) dry days (wet days) in the historical period mostly remain dry days (wet days) in the scenario unless daily precipitation amounts fall below or above the wet-day threshold by application of the DC scaling.

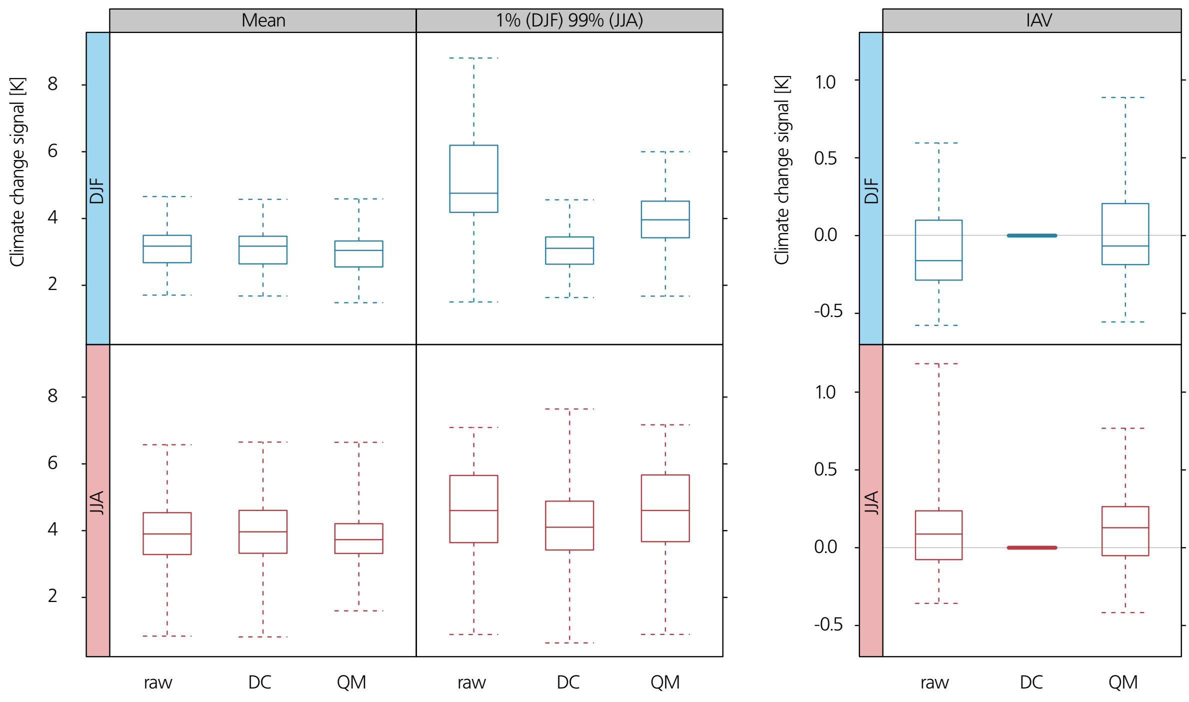


Figure 8

Additive temperature climate change signal between 1980–2009 and 2070–2099 for raw model output (raw), CH2011 DC-based scenarios (DC) and QM-based scenarios (QM) for winter (DJF, upper rows) and summer (JJA, lower rows). Left panel (Mean, 1% (DJF) 99% (JJA)): Seasonal mean temperature and 1st (DJF) and 99th (JJA) daily temperature percentile. Right panel (IAV): Interannual variability of seasonal mean temperature. The boxplots represent the signals for the 10 CH2011 GCM-RCM chains (dark gray background in Table 1) and the 28 temperature stations considered. Each box is defined by the lower and upper quartiles, the intermediate line segment is the median, and the whiskers extend to the data extremes.

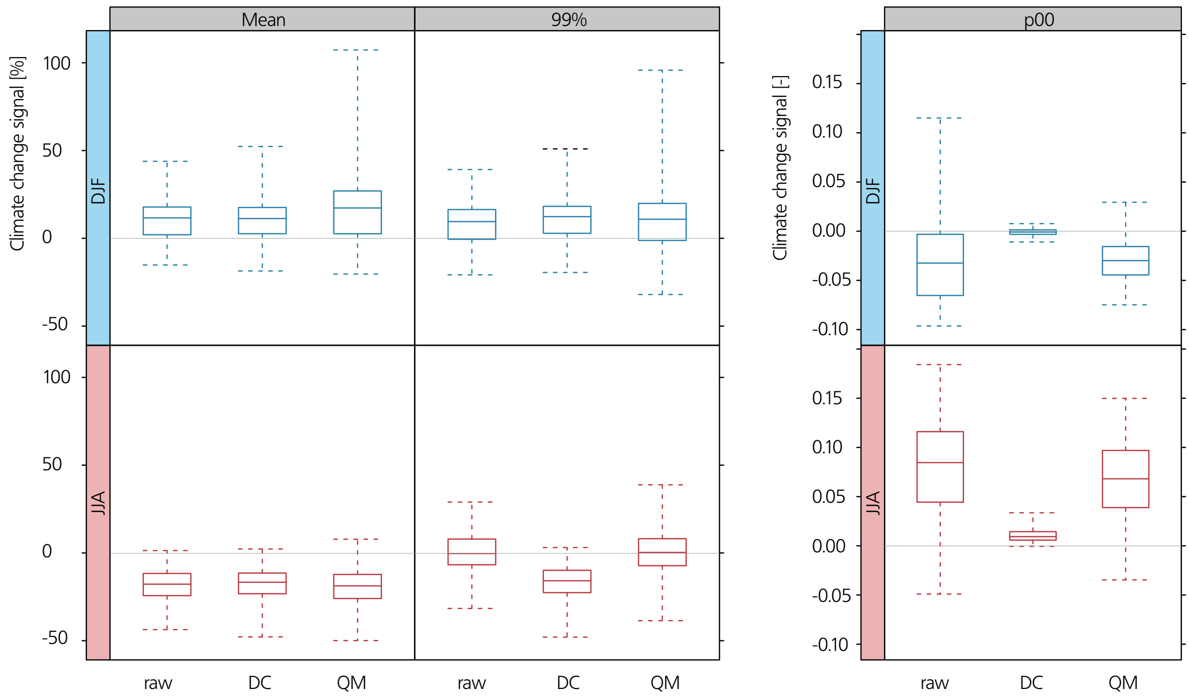


Figure 9

Precipitation climate change signals between 1980–2009 and 2070–2099 for raw model output (raw), CH2011 DC-based scenarios (DC) and QM-based scenarios (QM) for winter (DJF, upper rows) and summer (JJA, lower rows). Left panel (Mean, 99%): Multiplicative climate change signal for seasonal mean precipitation and for the 99th percentile of daily precipitation amounts. Right panel (p00): Additive climate change signal for the dry day-dry day transition probability (p00). The boxplots represent the signals for the 10 CH2011 GCM-RCM chains (gray background in Table 1) and the 27 precipitation stations considered. Each box is defined by the lower and upper quartiles, the intermediate line segment is the median, and the whiskers extend to the data extremes.

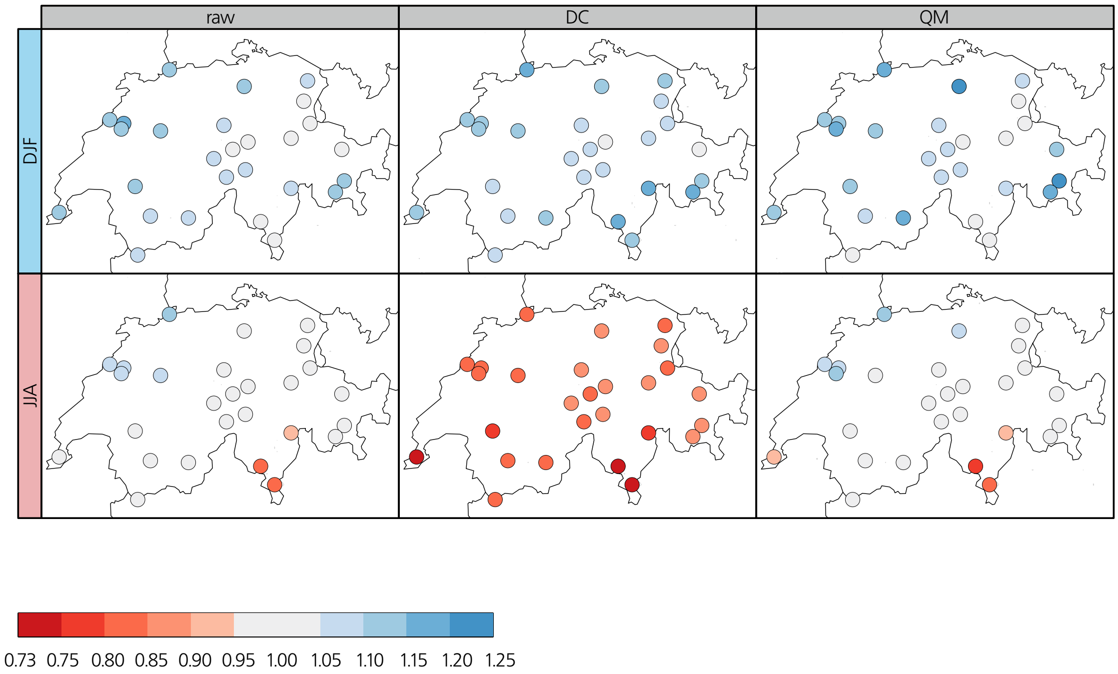


Figure 10

Mean multiplicative climate change signal at NBCN sites for the seasonal 99th percentile of daily precipitation between 1980–2009 and 2070–2099 for raw model output (raw, left column), CH2011 DC-based scenarios (DC, middle column) and QM-based scenarios (QM, rightmost panels) for winter (DJF, upper row) and summer (JJA, lower row). The climate change signals are averaged over the 10 CH2011 GCM-RCM chains (dark gray background in Table 1).

5| Limitations and Instructions for Use

The new transient bias-corrected scenarios for 15 GCM-RCM model chains presented in this CH2011 extension are available from www.ch2011.ch. They are complementing the DC-based CH2011 daily local scenarios (Bosshard et al. 2011) and are based on one emission scenario only (SRES A1B). For each of the 28 (temperature) and 27 (precipitation) NBCN stations considered, they provide transient daily temperature and precipitation time series for the period 1980–2099. Observations in the period 1980–2009 served as calibration target for QM. Hence, by definition the distribution of the bias-corrected data in this period corresponds to the distribution in the observations, but the bias-corrected time series exhibit their own temporal variability which is reflecting the variability in the underlying GCM-RCM chains and has no temporal correspondence with the observations.

As shown above the new scenarios provide an added value with respect to the DC-based CH2011 local daily scenarios. However, limitations concerning their usage in climate impact studies apply which mainly concern

(1) **Remaining biases:**
We here speak of "bias-corrected" scenarios, but the applied bias-correction procedure only targets the climatological distribution of temperature and precipitation and does not necessarily correct for biases in the temporal variability. However, there are strong indications that also the temporal variability on multiple temporal scales is accurately represented by QM-based scenarios (Ivanov and Kotlarski 2017; Rajczak et al. 2016b).

(2) **Spatial climate variability:**
The new QM-based scenarios were produced independently for the individual 28 (27) NBCN stations and have to be interpreted as single site scenarios. Any combination of the scenarios for two or more of these sites (for instance in hydrological applications) has to first verify whether spatial climate variability, i.e. the correspondence between scenarios for two or more sites, at the required temporal scale is accurately represented. Especially for close-by sites that are located in the same grid cell of the underlying GCM-RCM chain, spatial variability might be misrepresented as RCM subgrid variability is not accounted for (e.g., Maraun 2013).

(3) **Inter-variable relations:**
The new scenarios have been produced independently for the two variables temperature and precipitation. Hence, a correspondence between daily series of these two variables for a specific site is not necessarily given. However, recent works suggest a satisfying conservation of inter-variable dependencies and even some improvement of biased interdependencies by QM (Ivanov et al. 2015; Ivanov and Kotlarski 2017; Wilcke et al. 2013).

(4) **Trends and variance inflation:**
The applied deterministic QM methodology implicitly includes a downscaling step from coarse-resolution climate model output to the site scale. The latter typically exhibits a stronger temporal climate variability. As a consequence, the application of QM implies a statistical inflation of temporal variability of the climate model time series in many cases, which can have spurious influences on long-term climatic trends (Maraun 2013).

(5) **Stationarity of the bias correction function:**
In contrast to DC, QM can partly account for non-stationarities of climate model biases. If, for instance, model temperature biases are larger for high than for low temperature quantiles during the historical reference period, future warming is likely to increase the mean temperature bias. Such non-stationarities of the mean bias can be represented by QM. However, QM still assumes a stationarity of the quantile-based bias correction function itself. In case that GCM-RCM scenario simulations are subject to non-stationary biases of individual temperature or precipitation quantiles, these non-stationarities are not represented by QM-based climate scenarios. However, the pseudo-reality evaluation (Section 4; Ivanov et al. 2015; Ivanov and Kotlarski 2017) provides basic confidence in the applicability of QM also in a long-term context.

6| Conclusions

A new set of bias-corrected climate scenarios for Switzerland is presented. These scenarios are based on the QM methodology, cover daily temperature and precipitation and are available for 15 GCM-RCM model chains and 28 (27) stations in Switzerland. A detailed evaluation of the approach carried out within the ELAPSE research project indicates a satisfying and robust performance for a range of climate indicators at different temporal scales. A comparison of the new scenarios to the previous DC-based CH2011 local daily scenarios furthermore reveals an added value of QM with respect to changes in temporal climate variability, changes in moderate extremes and transient applications. However, also the new scenarios are subject to potential shortcomings and do not necessarily correct for all bias characteristics in the underlying GCM-RCM chains. Also, simulated changes in very rare events beyond the moderate extremes considered here have to be considered as highly uncertain. Therefore, a prudent and well-thought-through application in any climate impact assessment is recommended.

In a contextual sense, the new bias-corrected scenarios are meant to complement existing scenario products, in particular the DC-based local daily scenarios of CH2011. They hence enable a more comprehensive description and quantification of the influence of climate model downscaling and postprocessing strategies on the outcomes of climate impact assessments. Any strategy has its specific and inherent advantages but is also afflicted with its specific shortcomings and might have its specific fields of applications. In the case presented here, QM and DC can be associated with certain modifications of the raw models' climate change signals which will ultimately impact the results of subsequent model applications (such as the application of hydrological models). The availability and consideration of several scenario products obtained by independent postprocessing strategies is hence essential for the robustness of climate impact assessments.

Given the promising evaluation results presented above and in the available literature, the QM approach will represent a central approach in the frame of the upcoming CH2018 Swiss climate change scenarios for bias-correcting and downscaling the newest generation of GCM-RCM experiments provided by the CORDEX initiative www.cordex.org Hence, a further purpose of the scenario product presented in this report is to bridge the gap between the previous DC-based CH2011 scenarios and the upcoming CH2018 release (www.ch2018.ch).

References

Addor, N., and J. Seibert, 2014: Bias correction for hydrological impact studies – beyond the daily perspective, *Hydrological Processes,* **28,** 17, 4823-4828, doi: 10.1002/hyp.10238.

Begert, M., G. Seiz, N. Foppa, T. Schlegel, C. Appenzeller, and G. Müller, 2007: Die Überführung der klimatologischen Referenzstationen der Schweiz in das Swiss National Basic Climatological Network (Swiss NBCN), *Arbeitsberichte der MeteoSchweiz,* **215,** Bundesamt für Meteorologie und Klimatologie, MeteoSchweiz, 43 pp., available from http://www.meteoschweiz.admin.ch/content/dam/meteoswiss/de/Ungebundene-Seiten/Publikationen/Fachberichte/doc/arbeitsbericht215.pdf

Bellprat, O., S. Kotlarski, D. Lüthi, and C. Schär, 2013: Physical constraints for temperature biases in climate models, *Geophysical Research Letters,* **40,** 4042-4047, doi: 10.1002/grl.50737.

Boé, J., L. Terray, F. Habets, and E. Martin, 2007: Statistical and dynamical downscaling of the Seine basin climate for hydrometeorological studies, *International Journal of Climatology,* **27,** 1643-1655, doi: 10.1002/joc.1602.

Bosshard, T., S. Kotlarski, T. Ewen, C. Schär, 2011: Spectral representation of the annual cycle in the climate change signal, *Hydrology and Earth System Sciences,* **15,** 2777-2788, doi: 20.5194/hess-15-2777-2011.

Bundesamt für Umwelt BAFU, 2012: Anpassung an den Klimawandel in der Schweiz. Ziele, Herausforderungen und Handlungsfelder, Erster Teil der Strategie des Bundesrates vom 2. März 2012, www.bafu.admin.ch/ud-1055-d

Buser, C. M., H. R. Künsch, D. Lüthi, M. Wild, and C. Schär, 2009: Bayesian multi-model projection of climate: bias assumptions and interannual variability, *Climate Dynamics,* **33,** 849-868, doi: 10.1007/s00382-009-0588-6.

CH2011, 2011: Swiss Climate Change Scenarios CH2011, published by C2SM, MeteoSwiss, ETH, NCCR Climate, and OcCC, Zurich, Switzerland, 88 pp., ISBN: 978-3-033-03065-7.

CH2014-Impacts, 2014: Toward Quantitative Scenarios of Climate Change Impacts in Switzerland. Published by OCCR, FOEN, MeteoSwiss, C2SM, Agroscope, and ProClim. Bern, Switzerland, 136 pp.

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

Farinotti, D., S. Usselmann, M. Huss, A. Bauder, and M. Funk, 2012: Runoff evolution in the Swiss Alps: projections for selected high-alpine catchments based on ENSEMBLES scenarios, *Hydrological Processes,* **26,** 1909-1924, doi: 10.1002/hyp.8276.

Finger, D., G. Heinrich, A. Gobiet, and A. Bauder, 2012: Projections of future water resources and their uncertainty in a glacierized catchment in the Swiss Alps and the subsequent effects on hydropower production during the 21st century, *Water Resources Research,* **48,** W02521, doi: 10.1029/2011WR010733.

Fischer, E.M., J. Rajczak, and C. Schär, 2012: Changes in European summer temperature variability revisited, *Geophysical Research Letters,* **39,** L19702, doi: 10.1029/2012GL052730.

Frei, C., R. Schöll, S. Fukutome, J. Schmidli, and P.L. Vidale, 2006: Future change of precipitation extremes in Europe: Intercomparison of scenarios from regional climate models, *Journal of Geophysical Research,* **111:** D06105, doi: 10.1029/2005JD005965.

Frigessi, A., O. Haug, and H. Rue, 2002: A dynamic mixture model for unsupervised tail estimation without threshold selection, Extremes, 5(3), 219-235, doi: 10.1023/A:1024072610684.

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.

Hagemann, S., C. Chen, J. O. Haerter, J. Heinke, D. Gerten, and C. Piani, 2011: Impact of a Statistical Bias Correction on the Projected Hydrological Changes Obtained from Three GCMs and Two Hydrology Models, *Journal of Hydrometeorology,* **12,** 556-578, doi: 10.1175/2011JHM1336.1.

Ivanov, M.A., S. Kotlarski, and C. Schär, 2015: Assessment of performance of available quantile mapping methods for the current climate for stations in Switzerland, including inter-parameter consistency. Assessment of the long-term stability of quantile mapping in a pseudo-reality framework, *ELAPSE Deliverable B1,* Internal Report. Can be obtained from the authors on request.

Ivanov, M.A., and S. Kotlarski, 2017: Assessing distribution-based climate model bias-correction methods over an Alpine domain: Added value and limitations, *International Journal of Climatology,* **37,** 2633–2653, doi: 10.1002/joc.4870.

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

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

Maraun, D., and Coauthors, 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., M. Widmann, J. M. Gutiérrez, S. Kotlarski, R. E. Chandler, E. Hertig, J. Wibig, R. Huth, and R.A.I. Wilcke, 2015: VALUE: A framework to validate downscaling approaches for climate change studies, *Earth's Future,* **3,** doi: 10.1002/2014EF000259.

Maraun, D., and M. Widmann, 2015: The representation of location by a regional climate model in complex terrain, *Hydrology and Earth System Sciences,* **19**(8), 3449-3456, doi: 10.5194/hess-19-3449-2015.

Maurer, E.P., D. L. Ficklin, and W. Wang, 2016: Technical Note: The impact of spatial scale in bias correction of climate model output for hydrologic impact studies, *Hydrology and Earth System Sciences,* **20,** 685-696, doi: 10.5194/hess-20-685-2016.

Panofsky, H.W. and G.W. Brier, 1968. Some Applications of Statistics to Meteorology, The Pennsylvania State University Press, Philadelphia.

Piani, C., J. O. Haerter, and E. Coppola, 2010: Statistical bias correction for daily precipitation in regional climate models over Europe. *Theoretical and Applied Climatology,* **99,** 187-192, doi: 10.1007/s00704-009-0134-9.

Rajczak, J., S. Kotlarski, N. Salzmann, and C. Schär, 2016a: Robust climate scenarios for sites with sparse observations: a two-step bias correction approach, *International Journal of Climatology,* **36,** 1226-1243, doi: 10.1002/joc.4417.

Rajczak, J., S. Kotlarski, and C. Schär, 2016b: Does quantile mapping of simulated precipitation correct for biases in transition probabilities and spell-lengths? *Journal of Climate,* **29,** 1605-1615, doi: 10.1175/JCLI-D-15-0162.1.

Themessl, M.J., A. Gobiet, and A. Leuprecht, 2011: 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.

Themessl, M.J., A. Gobiet, and G. Heinrich, 2012: Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal, *Climatic Change,* **112,** 449-468, doi: 10.1007/s10584-011-0224-4.

van der Linden, P., and J. F. B. Mitchell, 2009: ENSEMBLES: Climate Change and its Impacts: Summary of research and results from the ENSEMBLES project, Met Office Hadley Centre, Exeter, UK, 160 pp.

Vrac, M., M. L. Stein, K. Hayhoe, and X.-Z. Liang, 2007: A general method for validating statistical downscaling methods under future climate change, *Geophysical Research Letters,* **34,** L18701, doi: 10.1029/2007GL030295.

Wilcke, R.A.I., T. Mendlik, and A. Gobiet, 2013: Multi-variable error correction of regional climate models, *Climatic Change,* **120,** 871-887, doi: 10.1007/s10584-013-0845-x.

Wood, E., L. R. Leung, V. Sridhar, and D.P. Lettenmaier, 2004: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs, *Climatic Change,* **62,** 189-216.