AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE
10.1002/2015JD024040

Key Points:
- Relating model biases to misrepresented processes is a key for model improvement and impact modeling
- We show using circulation types how the origin of some biases in GCM-RCM runs can be diagnosed
- We discuss the implications for bias adjustment and strategies to deal with biases

Correspondence to:
N. Addor,
naddor@ucar.edu

Citation:
Addor, N., M. Rohrer, R. Furrer, and J. Seibert (2016), Propagation of biases in climate models from the synoptic to the regional scale: Implications for bias adjustment, *J. Geophys. Res. Atmos.*, *121*, 2075–2089, doi:10.1002/2015JD024040.

Received 3 AUG 2015
Accepted 5 FEB 2016
Accepted article online 8 FEB 2016
Published online 4 MAR 2016

Propagation of biases in climate models from the synoptic to the regional scale: Implications for bias adjustment

Nans Addor[1,2], Marco Rohrer[3], Reinhard Furrer[4], and Jan Seibert[1,5]

[1]Department of Geography, University of Zurich, Zurich, Switzerland, [2]Research Applications Laboratory, National Center for Atmospheric Research, Boulder, Colorado, USA, [3]Oeschger Center for Climate Change Research and Institute of Geography, University of Bern, Bern, Switzerland, [4]Department of Mathematics, University of Zurich, Zurich, Switzerland, [5]Department of Earth Sciences, Uppsala University, Uppsala, Sweden

Abstract Bias adjustment methods usually do not account for the origins of biases in climate models and instead perform empirical adjustments. Biases in the synoptic circulation are for instance often overlooked when postprocessing regional climate model (RCM) simulations driven by general circulation models (GCMs). Yet considering atmospheric circulation helps to establish links between the synoptic and the regional scale, and thereby provides insights into the physical processes leading to RCM biases. Here we investigate how synoptic circulation biases impact regional climate simulations and influence our ability to mitigate biases in precipitation and temperature using quantile mapping. We considered 20 GCM-RCM combinations from the ENSEMBLES project and characterized the dominant atmospheric flow over the Alpine domain using circulation types. We report in particular a systematic overestimation of the frequency of westerly flow in winter. We show that it contributes to the generalized overestimation of winter precipitation over Switzerland, and this wet regional bias can be reduced by improving the simulation of synoptic circulation. We also demonstrate that statistical bias adjustment relying on quantile mapping is sensitive to circulation biases, which leads to residual errors in the postprocessed time series. Overall, decomposing GCM-RCM time series using circulation types reveals connections missed by analyses relying on monthly or seasonal values. Our results underscore the necessity to better diagnose process misrepresentation in climate models to progress with bias adjustment and impact modeling.

1. Introduction

Climate model simulations can show large departures from observations. These biases are often subtracted and then forgotten or postprocessed for impact modeling using pragmatic methods that typically do not account for the origins of the biases. Yet to improve climate projections, design robust bias adjustment methods, and derive uncertainty estimates, a better understanding of the physical processes leading to biases is necessary. Biases are the result of accumulated errors in the representation of various processes occurring at various spatial and temporal scales. A direct and critical consequence is that biases are difficult to reduce. Further, biases are not stationary over time [e.g., *Chen et al.*, 2015], and equally plausible assumptions on their evolution can lead to significantly different projected changes [*Buser et al.*, 2009]. In this article, we explore how a better diagnostic of the reasons behind biases can help with (1) the assessment of how much benefit could be gained by specific improvements (e.g., higher resolution in regional climate models (RCMs) or better synoptic circulation in general circulation models (GCMs)), (2) the better understanding of why some biases remain after bias adjustment, and (3) the design of more robust bias adjustment methods that are not purely empirical but account for known process misrepresentations.

Biases in RCM simulations emerge from processes at different spatial scales. At the regional scale, they were shown to stem from processes related to, for instance, soil moisture [e.g., *Bellprat et al.*, 2013], surface albedo [e.g., *Maraun*, 2012], or convective precipitation [e.g., *Ban et al.*, 2014]. This study focuses on biases inherited from the larger synoptic scale, i.e., those biases stemming from the synoptic circulation represented by the driving model, typically a GCM [e.g., *Noguer et al.*, 1998; *van Ulden and van Oldenborgh*, 2006]. Considering the synoptic situation links the air flowing over the region of interest to its origin and thereby can deliver major insights into the quality of process representation by climate models [*James*, 2006; *Maraun et al.*, 2011] and into the reasons behind the projected climate changes [e.g., *Cattiaux et al.*, 2013].

To describe the synoptic circulation, a common approach is to rely on circulation types (CTs), as they enable to reduce a multivariate atmospheric flow to a few classes [*Philipp et al*., 2010]. CTs are used routinely by weather forecasting services [e.g., *Schiemann and Frei*, 2010], and the identification of the synoptic situations causing extreme events is an area of active research [e.g., *Stahl and Demuth*, 1999; *Prudhomme and Genevier*, 2011; *Wilby and Quinn*, 2013]. Once such a relation between CTs and extreme events is established, it is tempting to use it to infer future changes in extremes based on the CT changes projected by climate models. Yet an important first step is to assess how well climate models represent CTs under present climate [*Huth et al*., 2008]. Several studies reported biases in the frequency of CTs simulated by climate models and found that they induce biases in regional surface variables, such as precipitation and temperature [*van Ulden and van Oldenborgh*, 2006; *van Ulden et al*., 2007; *Blenkinsop et al*., 2009; *Plavcová and Kyselý*, 2011].

Although the last years have seen the development of a profusion of bias adjustment methods, in general, these methods do not relate the biases to the (mis)representation of physical processes in the models, such as the representation of the synoptic circulation, but instead perform an empirical adjustment. An example is the widely used quantile mapping method, which relies on the transformation of the model output so that after the transformation, its cumulative distribution function matches that of the observations [*Piani et al*., 2010; *Themeßl et al*., 2011]. This method gained popularity in particular because of its ability to correct a wide range of statistical variables, such as the frequency of dry days or the variance of daily temperature [*Teutschbein and Seibert*, 2012]. Closer scrutiny, however, reveals that other aspects that are crucial for impact modeling are not corrected correctly: the diurnal temperature range [*Thrasher et al*., 2012], the precipitation under different circulation types [*Bárdossy and Pegram*, 2011], or multiday statistics [*Addor and Seibert*, 2014]. Further, although quantile mapping has been shown to retain the relationship between temperature and precipitation simulated by climate models, recent studies suggest that it does not correct this relationship when it is biased [*Wilcke et al*., 2013; *Li et al*., 2014]. Such issues may be fixed by further increasing the complexity of the statistical postprocessing. However, their recurrence reminds us that not identifying and accounting for the physical processes leading to the biases causes potentially significant errors to remain in statistically postprocessed simulations. It is hence fair to acknowledge that this kind of postprocessing is not a “correction” of the biases but rather an “adjustment,” and to refer to it as such.

To progress with the attribution of biases in RCM simulations and the understanding of their consequence for bias adjustment methods, this study explores how biases in synoptic circulation frequency propagate to the regional scale and influence our ability to bias adjust climate simulations. We address the following questions:

1. How well do GCM-RCMs of the ENSEMBLES project [*van der Linden and Mitchell*, 2009] capture the frequency and regime of CTs in the Alpine region under present climate?
2. How does this influence biases in GCM-RCM simulations of mean temperature and precipitation over Switzerland?
3. What are the implications for bias adjustment, in particular when using quantile mapping calibrated for different CTs?

2. Data and Methods

2.1. Climate Models and Evaluation Procedure

This study is based on 20 GCM-RCM simulations produced for the ENSEMBLES project (see Table 1 for the model list). These GCM-RCM combinations were selected by *Fischer et al*. [2012] to produce the probabilistic Swiss Climate Change Scenarios CH2011 [*CH2011*, 2011]. We evaluated each combination by extracting the grid points falling within the borders of Switzerland, an area of ~41,000 km^2. Model performance varies within this area, in particular with topography [e.g., *Addor and Seibert*, 2014], but here our focus is on the average performance over the whole country. We compared the simulated precipitation and temperature in winter (December–February, DJF) and summer (June–August, JJA) to gridded observations from the 2 km data sets RhiresD (precipitation) and TabsD (temperature) [*Frei*, 2013]. These two data sets were recently released by the Swiss Federal Office of Meteorology and Climatology and take advantage of the dense network of stations across the country (~420 for precipitation and ~90 for temperature). As a comparison, the E-OBS 8.0 data set [*Haylock et al*., 2008], which is popular for RCM evaluation, relies on precipitation and temperature data from only 37 Swiss stations. Observed and simulated fields were averaged over Switzerland and then compared.

Table 1. List of the Models Evaluated in This Study

ID	Institute	GCM	RCM
1	SMHI	BCM	RCA3
2	METNO	BCM	HIRHAM
3	ETHZ	HadCM3Q0[a]	CLM
4	HC	HadCM3Q0[a]	HadRM3Q0
5	UCLM	HadCM3Q0[a]	PROMES
6	VMGO	HadCM3Q0[a]	RRCM
7	METNO	HadCM3Q0[a]	HIRHAM
8	HC	HadCM3Q3[a]	HadRM3Q3
9	SMHI	HadCM3Q3[a]	RCA3
10	C41	HadCM3Q16[a]	RCA3
11	HC	HadCM3Q16[a]	HadRM3Q16
12	MPI	ECHAM5	REMO
13[b]	DMI[b]	ECHAM5[b]	HIRHAM[b]
14	KNMI	ECHAM5	RACMO
15	SMHI	ECHAM5	RCA3
16	ICTP	ECHAM5	REGCM
17	CNRM	ARPEGE	ALADIN
18	DMI	ARPEGE	HIRHAM
19	OURANOS	CGCM3	CRCM
20	GKSS	IPSL	CLM

[a]Note that in continuation we do not differentiate between HadCM3 different sensitivities (Q0, Q3, and Q16) and consider them as the same model.
[b]Model evaluated but excluded from the analyses and from the figures (see section 3.3).

An important aspect to take into account when comparing RCM simulations to gridded observations is observational uncertainties. Errors in observational data sets, which come in particular from the interpolation of station measurements to grid points and from measurement errors (e.g., precipitation undercatch), can be as large as the differences between RCM simulations and the gridded observations [*Addor and Fischer*, 2015]. Observational uncertainties are typically estimated by comparing several data sets or by cross validation, both methods presenting weaknesses. A source of error particularly relevant to this study is precipitation undercatch, which is particularly high in winter, when it is estimated to lead to an underestimation of precipitation by ~8% below 600 m and by ~40% above 1500 m above sea level [*Frei et al.*, 2015]. This is, however, not accounted for in the RhiresD data set used in this study, and to our knowledge, there is at present no data set covering Switzerland on a regional grid that accounts for precipitation undercatch. In this study we assume that up to 30% of the mean precipitation over Switzerland can be missed by gauges in winter.

Switzerland was chosen as a study area because it constitutes an interesting test bed in which to explore how well models capture the combined influence of the Alpine range and synoptic circulation on atmospheric conditions. Further, the size of the domain represents a compromise between an evaluation of specific grid cells, which may lead to an overinterpretation of model outputs because this resolution is finer than the model effective resolution, and an evaluation over a larger domain, in which opposite biases in different locations are more likely to partially mask each other.

The investigated period was 1980–2001. The end was determined by the availability of the ERA-40 reanalysis [*Uppala et al.*, 2005], which was used to derive the time series of "observed" CTs (section 2.2). A desired feature of any bias adjustment method to be applied to climate runs under future conditions is its ability to cope with potential bias nonstationarities over time [*Maraun*, 2012; *Teutschbein and Seibert*, 2013; *Chen et al.*, 2015]. We assessed bias nonstationarity during the study period by splitting the time series into two 11 years periods. Similar biases over the two periods would indicate bias stationarity under the current yet changing climatic conditions. Clearly, this would not yet imply stationarity in a climate scenario setting, but it constitutes a necessary test allowing us to start exploring the susceptibility of bias adjustment to natural climate variability and climate change.

2.2. Characterization of the Synoptic Circulation

For each GCM-RCM combination, we characterized the daily synoptic situations using an objective CT classification based on a combination of principal components analysis and hierarchical cluster analysis (PCACA) [*Philipp et al.*, 2010, 2014]. The classification relies on daily mean sea level pressure (SLP) fields within an area representative of the Alps (domain D06 in *Philipp et al.* [2010]; see Figure 1) and exists in three variants, using 9, 18, and 27 CTs. *Schiemann and Frei* [2010] compared 71 classification schemes and found that the nine-CT variant of the PCACA scheme captures winter daily precipitation over the Alps in a more reliable way than most other classification schemes with nine CTs and also better than a number of classification schemes using 18 or even 27 CTs, although its performance is weaker in summer than in winter. In this study we

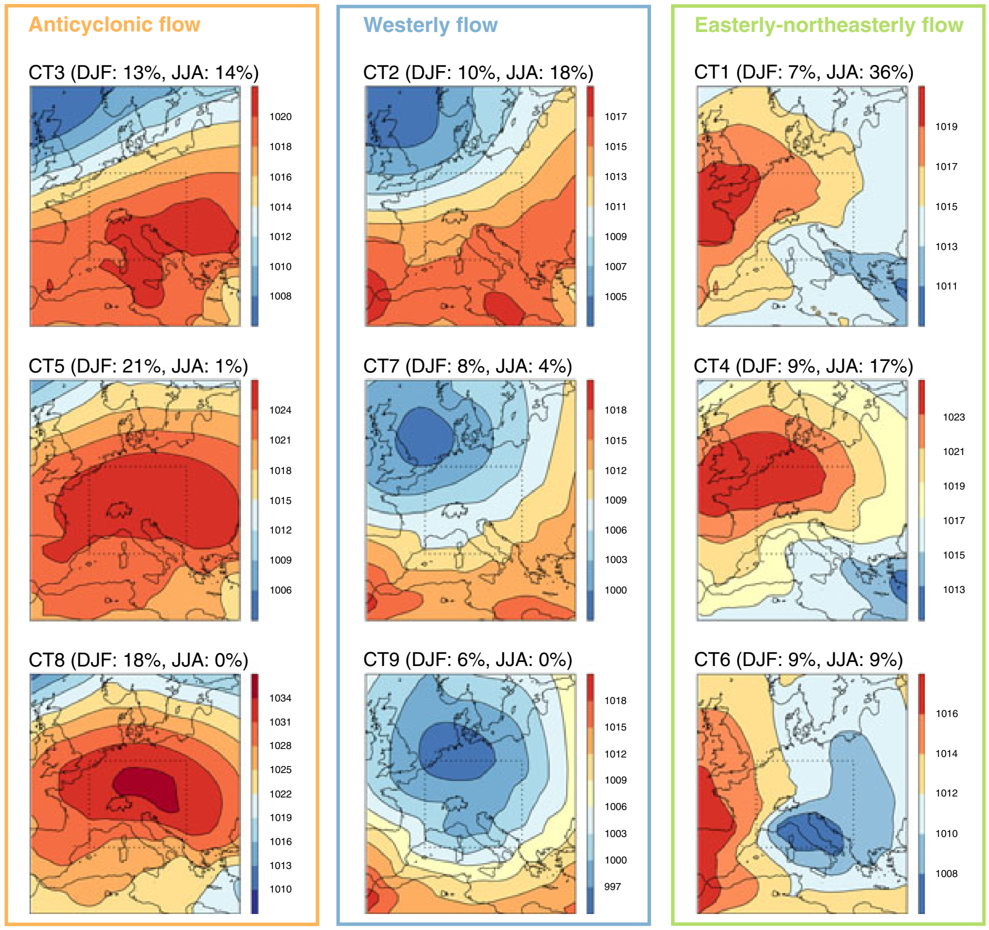

Figure 1. Mean daily sea level pressure [hPa] for the nine centroids of the PCACA09 classification determined over the Alps domain (dotted rectangle) including Switzerland (closed solid line within the rectangle). The frequency of occurrence of each CT in DJF and JJA is indicated in brackets. In the text and figures we refer to the CTs using their dominant flow and the following colors: anticyclonic (orange), westerly (blue), and easterly-northeasterly (green).

utilized the nine-CT variant of the PCACA scheme. CT classes can be determined at the seasonal level, but here we chose the classes determined at the annual level. The centroid of each CT and the time series of observed CTs were derived from ERA-40 reanalysis [*Uppala et al.*, 2005]. Using another reanalysis data set might have led to different CT time series, but we assume these differences to be overall smaller than the differences between the reanalysis and the climate simulations [*Riediger and Gratzki*, 2014]. We then assigned the simulated conditions of every day to the nearest CT in terms of Euclidian distance to its centroid [*Rohrer*, 2013]. This means that we used the same centroids (determined using ERA-40) to classify and then compare all the GCM-RCM simulations.

The initial analysis showed that the GCM-RCM combination ECHAM5-HIRHAM (Table 1) presents unreasonably large biases over Switzerland in winter, and hence, this GCM-RCM was excluded from further analyses (more details in section 3.3).

2.3. Bias Propagation From the Synoptic to Regional Scale

We hypothesized that part of the biases in precipitation and temperature comes from circulation biases [*van Ulden et al.*, 2007], in particular from the underrepresentation and overrepresentation of specific CTs. For instance, a cold bias in the regional simulations may originate from an overly frequent occurrence of CTs associated with cold conditions. We considered both the origin of the atmospheric flow and average temperature and precipitation over Switzerland. This enabled us to determine which of the nine CTs lead to cold or warm, wet or dry conditions (Figures 1 and 2). As can be expected, in both winter and summer

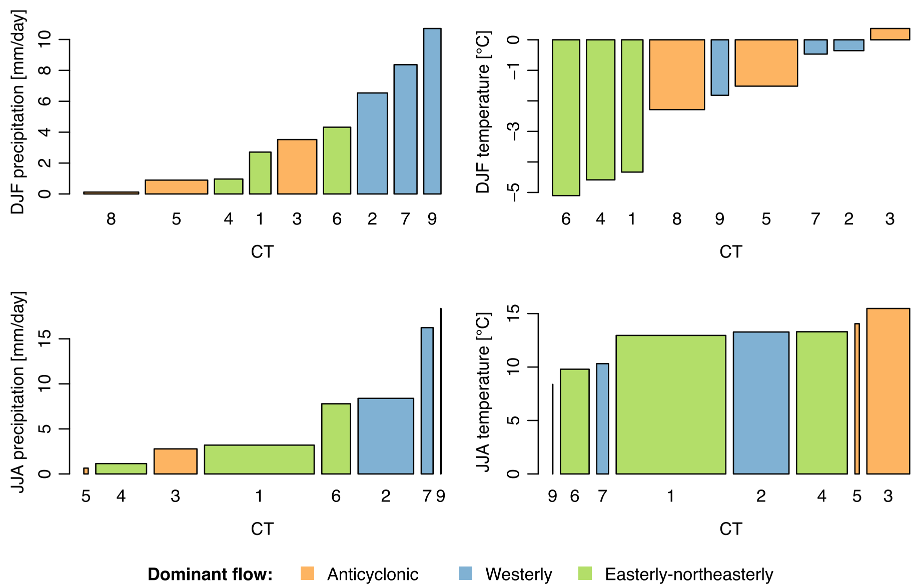

Figure 2. Mean daily (left column) precipitation amount and (right column) temperature observed over Switzerland in (top row) winter and (bottom row) summer for each CT (determined using ERA-40 SLP). The color of the bars corresponds to the dominant flow, their width is proportional to the frequency of occurrence of each CT, and they are ranked on the basis of their mean precipitation or temperature. Note that no days are attributed to CT 8 in JJA.

the CTs leading to the highest average daily precipitation correspond to westerly flows, advecting moist air from the Atlantic Ocean. In summer high-pressure systems over, or close to, Switzerland lead to the warmest conditions, while in winter the coldest conditions are caused by continental easterly-northeasterly flows. Given the good agreement between the large-scale flow and the mean precipitation and temperature, we decided to combine the nine CTs into three types of dominant flow: westerly, easterly-northeasterly, and anticyclonic flow (Figures 1 and 2). This choice has a double motivation. First, it was important for visualization and interpretation purposes to reduce the dimensionality of the problem while retaining the main CT differences, and second, since we perform CT-dependent bias adjustment, we needed enough data for each CT, which was not guaranteed for individual CTs with a low occurrence rate.

To assess how much benefit can be gained from an overall improvement of GCM simulations and how much bias in RCM simulations can be attributed to biases in CT frequency, we used two data sets: ERA-40-driven RCM simulations and time series constructed by resampling GCM-RCM simulations. Using a reanalysis to force an RCM improves fluxes at the domain boundaries, which has two main implications for our study. First, it improves the representation of the synoptic situation, since the CT classification used here is based on SLP, which is prescribed at the boundaries of the RCM domain by the reanalysis. Second, moisture and energy fluxes at the boundaries are also improved, which lead to more realistic precipitation and temperature simulations for individual CTs. To disentangle these two effects and assess the potential benefits of solely correcting biases in CT frequency, we resampled with replacement from the GCM-RCM runs. For each CT, we determined how many days ($N_{CT,OBS}$) were dominated by this CT in the observations, and randomly extracted with replacement $N_{CT,OBS}$ days from the $N_{CT,SIM}$ days dominated by the same CT in the GCM-RCM simulation (the difference between $N_{CT,OBS}$ and $N_{CT,SIM}$ is the bias in CT frequency). We completed the constructed time series by appending days from the other CTs repeating the same procedure. This leads to inconsistencies in the temporal sequence of the time series, but this is not an issue here since we restrict our attention to mean values, which are insensitive to the order of the time series values. With this resampling, precipitation and temperature biases for individual CTs remain essentially unchanged, but the frequency of each CT is corrected.

2.4. Bias Adjustment Using Quantile Mapping

We bias adjusted the RCM simulations by applying quantile mapping using two setups. For both setups quantile mapping was implemented using empirical quantiles [*Gudmundsson et al.*, 2012] and was applied on daily

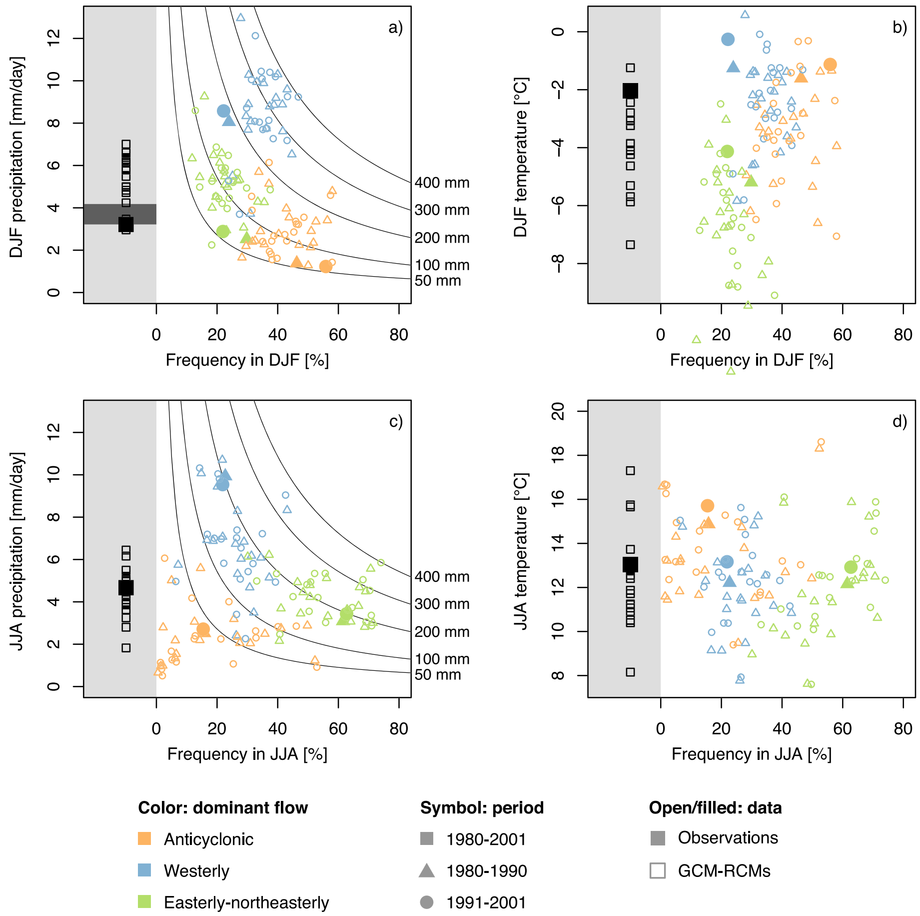

Figure 3. Model-observation differences in CT frequency (*x* axis) and in precipitation and temperature (*y* axis). Each open symbol designates one of the 19 GCM-RCM runs. The black symbols on the left of each panel correspond to seasonal averages over 1980–2001. The isolines in Figures 3a and 3c indicate the combinations of mean precipitation and frequency leading to the indicated seasonal (3 month) precipitation amounts. The dark gray area in Figure 3a represents the estimated undercatch of winter precipitation (30%).

precipitation and temperature averaged over Switzerland. In the first "standard" setup, the quantile mapping adjustment is conditional on the season, so two adjustments were used: one for summer and one for winter. In the second "CT-dependent" setup, the adjustment is conditional on both the season and the dominant flow, so six adjustments were used for temperature and six for precipitation. To calibrate the CT-dependent quantile mapping, we used all days with a particular flow during a particular season (e.g., anticyclonic flow in winter) from both the observations and a particular GCM-RCM run. The differences between the empirical cumulative distribution functions for these two data sets were then used to establish the transfer functions on which the quantile mapping method relies. This idea to apply a CT-dependent bias adjustment is inspired by *Bárdossy and Pegram* [2011], who observed that different CTs can lead to different biases in RCM simulations and, based on the fact that different CTs correspond to different atmospheric processes, proposed that the bias adjustment method should be CT dependent.

3. Results

3.1. Circulation Biases and Within Circulation Type Biases

Different CTs correspond to different regimes in terms of precipitation and temperature, and overall, the models capture these differences reasonably well (Figure 3). For instance, when in winter continental air is advected along the Alps by an easterly-northeasterly flow, the average temperature is lower than if the region was under the influence of a westerly or anticyclonic flow. This is captured by the models, and they also capture differences in CT frequency, for instance, that in winter, days with a dominant easterly-northeasterly flow are less frequent than those with a dominant anticyclonic regime (Figure 3b). Yet there is some considerable spread of the simulations among the observations, which reflects the existence of biases in the frequency (circulation biases) and the precipitation intensity/temperature of each CT (within CT biases). The contribution of these two kinds of biases to the mean seasonal bias is variable. For instance, in winter, all the models overestimate the frequency of westerly situations and several of them overestimate their mean precipitation intensity (Figure 3a). Since this circulation type is associated with high precipitation intensities (Figure 2), a small bias in the frequency leads to an important bias in the seasonal average (see the steep isolines in the upper part of Figure 3a). In contrast, an overestimation of the frequency of days with drizzle leads to much smaller changes in the seasonal average (see the flatter isolines in the lower part of Figure 3a). Further, as pointed out by *Bárdossy and Pegram* [2011], different CTs may show different biases. For instance in summer, models tend to overestimate precipitation amount on easterly-northeasterly flows but typically underestimate it on westerly situations (Figure 3c). Another important result is that there is a relative stability of CT biases over our study period; i.e., fluctuations in both circulation biases and within circulation type biases between the two subperiods are relatively minor (Figures 3a–3d). The open triangles and open circles of the same color (first and second subperiod, respectively) are quite well mixed, and importantly, the distance between the filled triangle and filled circle of the same color (variability in the observations) is smaller than the spread among the open symbols of this same color (biases), even when temperature increases (Figures 3b and 3d). It means that natural climate variability plays a small role in the model-observation differences considered here. As stated earlier, it does not imply bias stationary under changing climatic conditions, but it allows us to not discriminate between the first and second subperiod and to consider the whole 1980–2001 period in this study.

Winter precipitation in Switzerland is overestimated by 18 out of 19 models (black symbols in Figure 3a), which reflects a wide-spread issue among ENSEMBLES models already reported by *Fischer et al*. [2012], leading to the simulation of an unrealistically large snow accumulation and overall water resources. The overestimation of winter precipitation by the models is too large to be explained by precipitation undercatch alone or by interpolation errors and natural variability [*Addor and Fischer*, 2015]. Further, we suspect that one of the reasons explaining such a consistent positive bias is the lack of independence between the model runs, with two GCMs driving as many as 13 RCMs (Table 1), and we hypothesize that this bias stems partially from the synoptic fields inherited from the GCMs [*van Ulden and van Oldenborgh*, 2006].

3.2. RCM Response to Circulation Biases

To explore the influence of the synoptic field on RCM simulations, we represented the bias in temperature and precipitation as a function of the frequency bias in one CT (Figure 4). The clustering of points of the same colors indicates that the driving GCM has a considerable influence on the circulation bias, both in winter and summer. Although RCMs respond to the biases differently, we find that the CTs they simulate are to a large extent prescribed by the driving GCM. GCMs explain the majority of the variance of the bias in the CT frequency and as expected, the control of the GCM on CT frequency is stronger in winter than in summer (the variance explained is 0.73 versus 0.70 for westerlies, 0.57 versus 0.53 for easterlies-northeasterlies, and 0.88 versus 0.74 for anticyclonic situations, respectively).

The correlation between circulation biases and surface variables is significant for two combinations, which we show in Figures 4a and 4b. In winter, there is an almost systematic overestimation of the precipitation and of the frequency of westerlies: the models overestimate the precipitation by 71% and the frequency of westerlies by 57% on average (Figure 4a). The significant correlation between these two variables ($r = 0.57$) is consistent with the fact that an overestimation of westerlies leads to too much moist air being transported from the Atlantic and precipitating in the Alpine region. A similar conclusion was drawn by *van Ulden et al*. [2007] for the previous generation of RCM runs for Europe (PRUDENCE project [*Christensen and Christensen*, 2007]). Their study focused on Central Europe, for which they reported an overestimation of the west component of the geostrophic wind in winter, and established its contribution to wetter winters. Although here we did not consider the strength of the flow, their results are compatible with the overestimation of the frequency of westerlies we report, which suggests a persistent bias. Further, note that precipitation is overestimated for a large majority of the models and flow situations (Figure 3a), which indicates general issues with the representation of winter precipitation by the models in our study area.

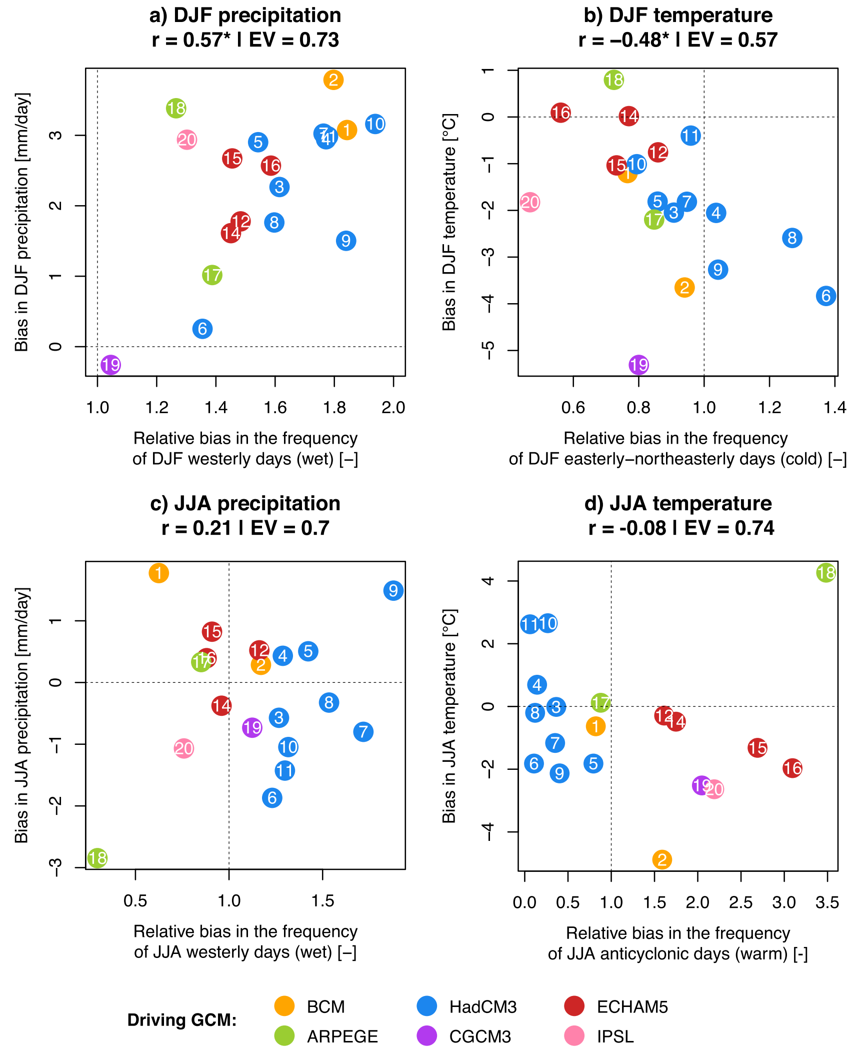

Figure 4. Bias in the frequency of a selected flow situation and in mean seasonal precipitation and temperature for the 19 GCM-RCMs. The Pearson's product moment correlation coefficient *r* (star indicates significance at the 5% level) and the explained variation *EV* of the frequency bias by the GCMs are indicated. Dots of the same color represent RCMs driven by the same GCM; the numbers refer to the different GCM-RCM combinations (see Table 1 for model IDs).

As mentioned before, we excluded the model ECHAM5-HIRHAM from this study. We found that the winter precipitation it simulates over Switzerland is about 328% of the observed amount and that 71% of its winter days are westerlies, which should be compared to the observed frequency of 23%. This agrees with the connection of biases in westerly frequency and precipitation amounts, but because of these particularly large biases, we decided to exclude this model from all the analyses and figures presented in this study. The tendency of ECHAM5-HIRHAM to overestimate winter precipitation to a larger extent than other ENSEMBLES model chains was also reported for Spain by *Turco et al.* [2013], but they did not investigate the reasons behind this bias.

The bias in winter temperature is significantly negatively correlated ($r = -0.48$) with the bias in the frequency of easterly-northeasterly flow, which in the models and in the observations, leads to below average temperature in Switzerland (Figure 3b). In other words, these two biases are consistent. The main axis of the cloud of points in Figure 4b does not include the origin, which is consistent with the underestimation of winter temperature for almost all the model-CT combinations (Figure 3b).

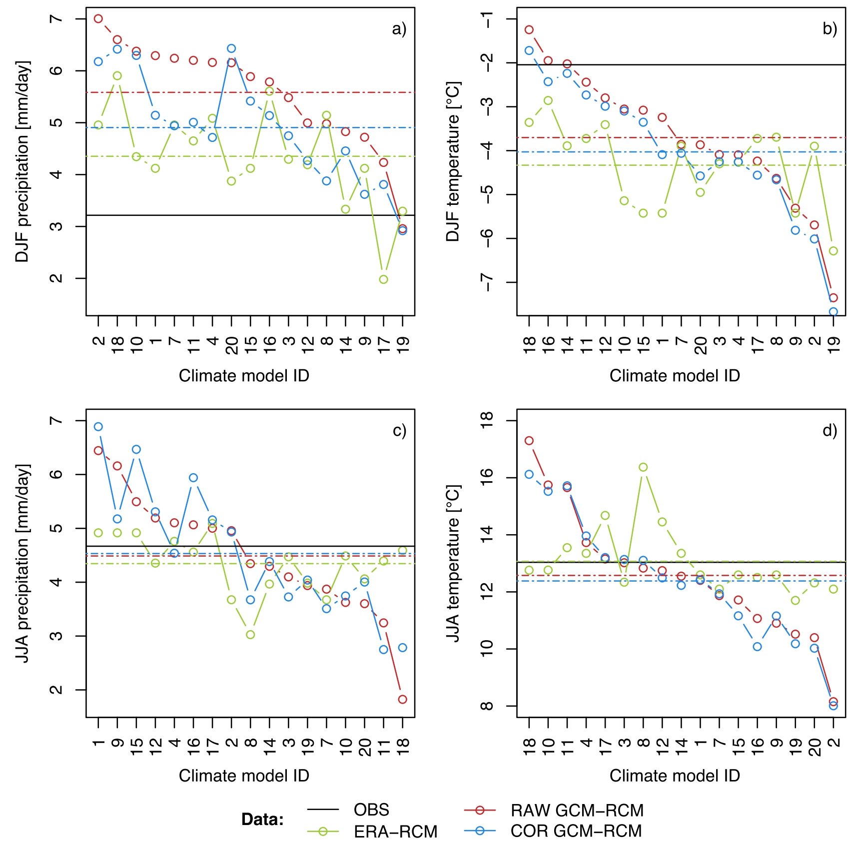

Figure 5. Assessment of the potential benefit of correct CT frequencies (see Table 1 for model IDs). The blue line indicates RCM simulations in which only the CT frequency is corrected (by resampling), while the green line indicates RCM model simulations in which boundary conditions (in particular SLP, on which the CT classification is based) are prescribed by ERA-40. The dashed colored lines indicate the seasonal mean across the models. Models are sorted based on the mean seasonal (a and c) precipitation or (b and d) temperature of the raw GCM-RCM simulations (red line). The RCMs RRCM and PROMES do not appear here: we experienced technical problems with the output files of ERA-40-PROMES, and no ERA-40-RRCM runs were available on the ENSEMBLES data portal.

The influence of synoptic situation is much weaker in summer: the correlations between the circulation biases and the surface variables (mean precipitation and temperature) are not significant. RCMs driven by the same GCM show similar circulation biases, but these can lead to quite different temperature biases. This likely reflects that processes at smaller spatial scales, such as soil moisture feedbacks and convective precipitation, are more influential during the summer months. Figure 4d also highlights that some models provide reasonable estimates of mean summer temperature, despite boundary conditions with clear deficiencies. For instance, the frequency of situations dominated by an anticyclonic circulation is clearly underestimated by HadCM3 and overestimated by ECHAM5 (blue versus red dots in Figure 4d). Yet some RCMs driven by these GCMs manage to capture the mean observed temperature surprisingly well. This raises the question whether these models get the right answer for the right reasons and suggests that projected changes in CTs should be interpreted with caution, given the important biases under present climate conditions.

The results in Figure 4 depict the correspondence between biases in mean precipitation and temperature, and a single CT. To account for the three CTs at once and to estimate the potential benefits of reducing CT frequency biases, we used constructed time series based on the resampling of GCM-RCM simulations. Figure 5a shows that the correction of the CT frequency leads to a decrease of winter precipitation bias by 32% (mean across the models). This provides further support to the idea that biases in CT frequency directly contribute to the wet winter bias in Switzerland. The overestimation of winter precipitation is reduced even further when ERA-40 is used to force the RCMs, which is expected because ERA-40 not only reduces CT frequency biases but also provides a more realistic boundary forcing. In contrast to winter precipitation,

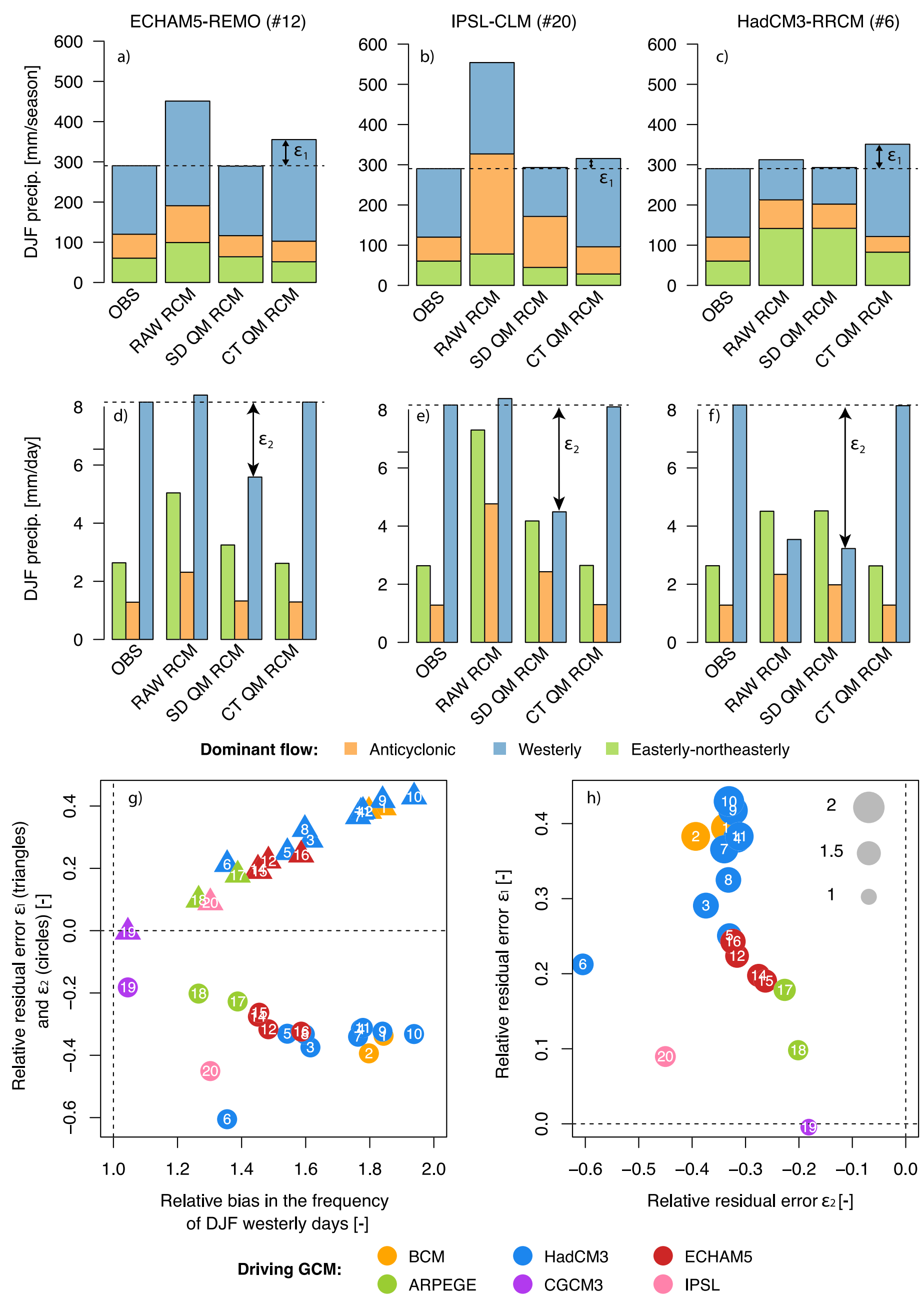

Figure 6. An illustration of residual errors in bias adjusted time series and of their relation with large-scale circulation biases. (a–f) Winter precipitation as observed (OBS), as simulated by three GCM-RCMs without bias adjustment (RAW RCM) and after bias adjustment using standard and CT-dependent quantile mapping (SD QM RCM and CT QM RCM, respectively); the colors correspond to the dominant flow. Arrows indicate the residual error in mean winter precipitation after CT-dependent bias adjustment (ε_1) and in mean winter precipitation on westerly days after standard bias adjustment (ε_2). (g) Residual error (expressed relatively to the observed precipitation amount) as a function of the bias in the frequency of westerly flows in winter; the colors correspond to the driving GCM. (h) Same information as in Figure 6g, but with more emphasis on the driving GCM; the size of the dots is proportional to the bias in the frequency of westerly flows in winter.

winter temperature does not benefit from a correct CT frequency (Figure 5b). The reason is that even if precipitation is overestimated on westerly days and anticyclonic days, removing westerly days (mean daily precipitation ~8 mm) and replacing them by anticyclonic days (mean daily precipitation ~2 mm) will lead to a decrease of the seasonal precipitation amount. This is also illustrated in Figure 3a): the isolines are steeper

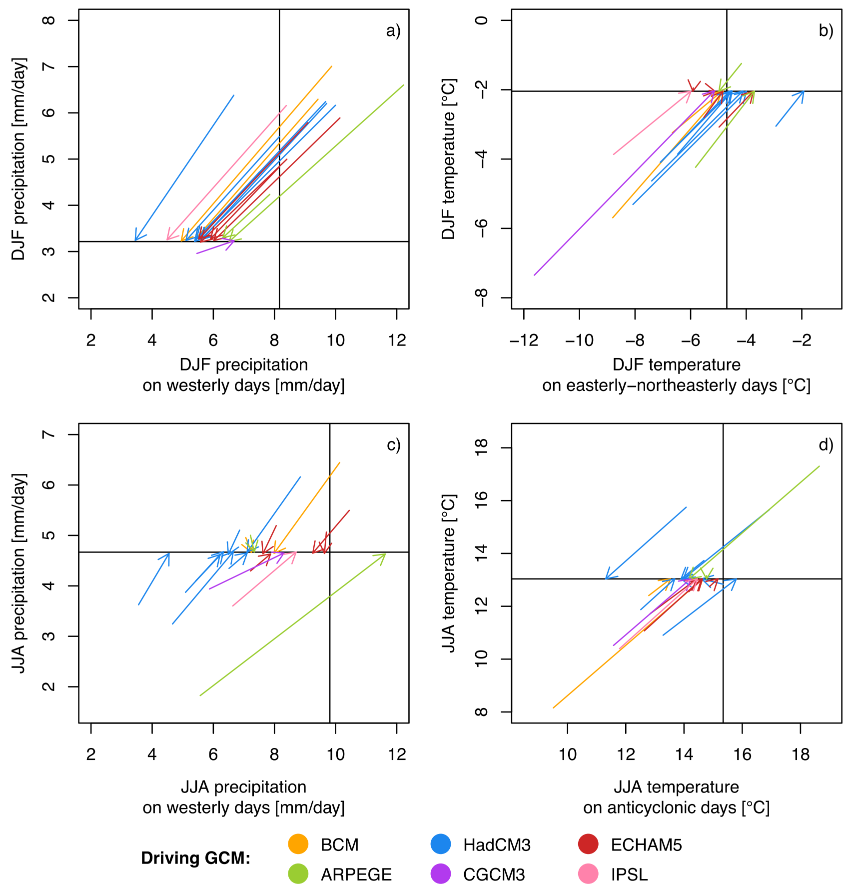

Figure 7. Changes in model-observation differences as a result of standard quantile mapping. Each arrow corresponds to a GCM-RCM run, the tail and head represent the simulated variable before and after postprocessing, respectively. The *y* axis indicates the variable quantile mapping has been trained to correct at the seasonal scale, while the *x* axis represents the same variable for a specific CT, which has not been considered in the calibration of the quantile mapping method. The solid black lines represent the observed value of the two variables.

around the blue points than around the orange points, meaning that changing the frequency of westerly days has a larger impact on the seasonal precipitation amount than changing the frequency of anticyclonic days by the same amount. This however does not apply to winter temperature: resampling from the three dominant flows, which are affected by a similar cold bias (Figure 3b), does not significantly change the seasonal temperature. In summer there is almost no benefit from correcting CT frequency biases (Figures 5c and 5d).

3.3. Residual Errors in Bias-Adjusted Time Series

We tested the sensitivity of bias-adjusted time series by applying quantile mapping using two setups: one conditional on the season and one on both the season and the CT. Both variants perform well at the tasks they were trained for, but their lack of consideration for the physical processes leading to biases prevents them from correcting for other aspects. As expected, when quantile mapping is calibrated for each season, the mean seasonal precipitation is correct (Figures 6a–6c), and when it is calibrated for each CT, the precipitation on each of the CT is correct (Figures 6d–6f). However, significant biases remain in the corrected time series of winter precipitation, as quantile mapping introduces compensation errors. For instance, for the bias-adjusted simulations to match the observed seasonal precipitation despite the frequency overestimation of westerly flow, standard quantile mapping leads to an underestimation of precipitation on westerly flow situations (ε_2 in Figures 6d–6f). Similarly, the seasonal precipitation is overestimated (ε_1 in Figures 6–6c) when a CT-dependent adjustment is applied because the overestimation of westerly days is not accounted for.

These residuals errors are present for all models. Figure 6g shows that ε_1 and ε_2 tend to increase in absolute terms with circulation biases, represented here by the bias in the frequency of winter westerly situations. For instance, the model HadCM3-RCA3 (#10) presents both the largest residual error in the seasonal winter amount (43%) and the largest overestimation of the westerly frequency (93%) of all models considered. Note the particular situation of the models IPSL-CLM (#20) and HadCM3-RRCM (#6). Their circulation bias in winter is lower than the average, but their absolute residual error ε_2 is larger than that of the other models. This likely results from the misdistribution of precipitation among three circulation types in the raw (uncorrected) model runs. In the observations, the ratio of mean precipitation on a westerly day to an easterly-northeasterly day is 3.1:1, but it is largely underestimated in both IPSL-CLM (1.1:1, Figure 6e) and in HadCM3-RRCM (0.8:1, Figure 6f). The correct ratio is reached when CT-based adjustment is applied, but the standard quantile mapping fails to reestablish the balance between the different regimes. Even after the adjustment, it still rains more in average on easterly-northeasterly days than on westerly days in HadCM3-RRCM (Figure 6f), which leads to a significant ε_2. The importance of the driving GCM discussed earlier can also be seen from the residual errors, which increase with the deficiencies in the large-scale circulation (Figure 6g). The combination of residual errors (ε_1, ε_2) largely depends on the driving GCM as can be seen from clusters of dots related to one GCM (Figure 6h).

As already stressed, bias-adjusting DJF precipitation using standard quantile mapping leads to a degradation of the simulations of precipitation on westerly days (Figure 7a). But in other cases, the situation on specific CTs is improved by the bias adjustment of the overall seasonal conditions. This applies for instance to DJF temperature on easterly-northeasterly days (Figure 7b) and to JJA temperature on anticyclonic days (Figure 7d). JJA precipitation on westerly days is improved by the correction of JJA precipitation for some models, but it is degraded for other models (Figure 7c). Overall, Figures 7a–7d illustrate that while quantile mapping leads by construction to a good match between the distribution of the observations and that of the postprocessed time series, it also alters other aspects of the model simulations. Whether these alterations are beneficial or detrimental depends on the variable of interest and on the GCM-RCM.

4. Discussion

4.1. Implications for Bias Adjustment and Impact Modeling

Overall, quantile mapping enables substantial bias reductions. Yet we show that when circulation biases are large, they lead to significant residual errors in bias-adjusted time series. In such situations the limitations of ad hoc adjustments not accounting for misrepresented processes clearly appear. RCMs inherit circulation biases from their driving GCM, which leads to biases in surface variables crucial for impact modeling. These circulation biases illustrate the difficulty to capture the present and future climate dynamics using models [*Shepherd*, 2014]. They also represent a challenge for the interpretation of climate projections and for the bias adjustment of RCM simulations. Although we were able to identify their contribution to biases in winter precipitation and temperature, it is unclear how to correct circulation biases by postprocessing, since they are deeply rooted in the simulations. Indeed, as atmospheric circulation is at the core of climate projections, it needs to be handled with care, and pragmatic approaches often used when preparing climate models outputs for impact simulations might not be appropriate. It remains to determine how best account for circulation biases, in particular on how we could correct them a posteriori without compromising the reliability of the simulations. Concretely, should we select and remove days of overrepresented CTs and replace them by days of underrepresented CTs?

One way to progress would be to embrace the fact that different models have different deficiencies. We argue that model errors should be better diagnosed (for instance using CTs) and this should guide the choice of the bias adjustment method. That is, instead of using the same bias adjustment for all the models of the ensemble in a “one fits all” approach, model-specific bias adjustment methods could be used. For instance, choosing between a standard or a CT-based bias adjustment method may be a choice that is model dependent. In the case of HadRCM3-RRCM, the biases within CTs are so important (Figure 6f) that a CT-based approach may be considered to give better results than the standard approach. Another essential aspect that will influence the adequacy of a CT-dependent bias adjustment is a future change in CT frequencies. In cases where RCM biases are found to be CT-dependent and climate scenarios indicate a change of CT frequencies, then model biases cannot be assumed to be stationary in time. The contribution of CT-dependent errors to bias nonstationarities requires further research, but if these errors are significant, this might call for CT-dependent bias adjustment methods.

When it comes to impact modeling, a key question is how much residual errors in bias-adjusted time series matter. This depends on the sensitivity of the impact model to the bias adjustment [*Muerth et al*., 2013; *Teng et al*., 2015] and on the other sources of uncertainty involved [*Chen et al*., 2011; *Addor et al*., 2014]. In some cases, the time series might be deemed "good enough," for instance, if their residual errors are smaller than those in the observations, resulting from interpolation and measurement processes. In others, for instance when the impact model is highly distributed or if the focus is on a few extreme events, then these residual errors may become a concern.

For our area of interest, applying quantile mapping in its standard form leads for several GCM-RCMs to an overestimation of the mean precipitation falling on days dominated by an easterly-northeasterly flow, which are in average more than 2°C colder than days dominated by the two other flows. A potential consequence is an overestimation of the snow pack, which would degrade hydrological simulations in snow-dominated catchments. Only few studies have so far evaluated intervariable relationships in the context of bias adjustment and have done so using statistical indicators (e.g., correlation). We envisage that hydrological models could be used to evaluate RCM simulations and bias adjustment methods in an integrated way. In this approach observed catchment runoff is compared to simulated runoff using RCM simulations as input. The advantage is that the combined effect of temperature and precipitation and the significance of biases in these two variables for runoff are directly taken into account.

4.2. Implications for Climate Model Development and Evaluation

RCMs are designed to disaggregate GCM simulations in space and time. They, however, do not have the purpose nor have the ability as illustrated by this and other studies to correct biases in the frequency of synoptic situations. It implies that even if the resolution of the RCMs used in this study was further increased, the frequency of westerlies simulated by the GCM (and hence the water vapor introduced at the domain boundaries) would remain overestimated, and consequently, winters would remain too wet. This is a case in which the added value of dynamical downscaling is limited by the boundary conditions provided by the GCM [*Racherla et al*., 2012]. The fact that GCMs can introduce significant biases in RCM simulations underscores the importance to evaluate RCM simulations not only by using a reanalysis-RCM setting (i.e., with close to correct synoptic fields) but also by using a GCM-RCM setting (i.e., with potentially significantly biased synoptic fields). The former allows for the separation between downscaling deficiencies and deficiencies of the boundary forcing and is a crucial step in model development and tuning. Yet since RCMs are overwhelmingly used to gain insight into future conditions, for which no reanalysis is available, it is essential to also evaluate them in a GCM-RCM setting, to test how well they cope with potentially significant circulation biases. We argue that this is particularly true in the framework of bias adjustment studies and impact assessments.

Finally, the importance of the circulation biases under present conditions for the models investigated here implies that caution should be used when be interpreting projected changes in the frequency of circulation types. Now that the simulations from the Coordinated Regional Downscaling Experiment (CORDEX) [*Giorgi and Gutowski*, 2015] start being used for regional assessments of climate change and impact studies over Europe, we recommend to assessing the reliability with which CT are simulated in these new simulations, before inferring impacts from projected changes in CT frequency. The ENSEMBLES runs used in this study rely on GCMs from Coupled Model Intercomparison Project Phase 3 (CMIP3) generation. The comparison performed by *Perez et al*. [2014] suggests that GCMs from the following generation (CMIP5, which provide the forcing for CORDEX simulations) capture better the synoptic situation over the northeast Atlantic region. However, it remains to be evaluated whether this improvement is significant enough to lead to a decrease of biases in regional precipitation and temperature.

5. Conclusions and Outlook

Accounting for the processes leading to biases is essential to progress with bias adjustment, but these processes are difficult to pinpoint. One way to identify them is to perform a CT-dependent evaluation of climate models. This enables to relate the synoptic scale to the regional scale and can reveal model deficiencies overlooked when performing model evaluation at the seasonal time scale. We used a CT-dependent evaluation to explore precipitation and temperature over Switzerland, which revealed that the overestimation of the frequency of westerly situations was a significant contributor to the overly wet conditions simulated by the large majority of the ENSEMBLES GCM-RCMs in winter. However, we could not explain biases in other variables (e.g., summer temperature), which require a focus on other scales and variables (e.g., soil moisture) [*Bellprat et al*., 2013]. Further, our study investigated mean precipitation and temperature over Switzerland, but future studies could concentrate on smaller, well-chosen areas to address questions such as how well atmospheric conditions on different sides of the Alps are reproduced under different synoptic situations or how well wet or dry extremes in relation to different CTs are captured by the GCM-RCMs.

As for bias adjustment, we propose that a better understanding of what we correct is necessary. It is well established that bias adjustment only improves selected aspects of climate simulations and leaves inconsistencies in the postprocessed time series, which we refer to as residual errors. However, we go beyond this simple fact and illustrate how the existence and the amplitude of these residual errors can be related to misrepresentation of specific atmospheric processes (i.e., the frequency of specific CTs). We show that compromises are necessary, since the improvement of one aspect of an RCM simulation can lead to a degradation of other aspects.

We propose that the choice of the bias adjustment method should better account for two key elements: (i) the origin of the biases, which we argue must be better diagnosed, and (ii) the purpose of the bias-adjusted time series. In the context of precipitation over Switzerland, it is important to determine what is most valuable: the right seasonal precipitation, or how it is distributed among different synoptic circulations? The answer to this question is application dependent and will rely on the experience of impact modelers.

Overall, circulation biases represent a real challenge for both the dynamical downscaling and the bias adjustment of RCM simulations, because they are deeply rooted in the atmospheric circulation simulated by GCMs. As more recent GCMs improve the representation of atmospheric circulation, this should contribute to the reduction of RCM biases. Further research is nevertheless necessary to better understand the processes behind biases in GCM-RCM simulations and to clarify how downscaling and bias adjustment methods should account for them.

Acknowledgments
We are grateful to the contributors to the COST733 Action "Harmonisation and Applications of Weather Types Classifications for European Regions" for their systematic analysis of CT classifications and for making the classification software *cost733class* available [*Philipp et al*., 2014]. We thank the EU FP6 Integrated Project ENSEMBLES projects for the GCM-RCM simulations (http://ensemblesrt3.dmi.dk/) and the Swiss Federal Office of Meteorology and Climatology, MeteoSwiss for the RhiresD and TabsD data sets (these data sets can be requested from MeteoSwiss). The CT classification of the GCM-RCM simulations is available from Marco Rohrer (marco.rohrer@giub.unibe.ch). This article benefited from constructive comments from Sven Kotlarski and two anonymous reviewers and from stimulating and inspiring discussions with Renate Wilcke and Andreas Prein. Nans Addor thanks the University of Oregon, USA, for hosting him while working on this study. This research was supported by the Swiss National Science Foundation (grant 200021_131995) and by the University of Zurich (University Research Priority Programs on Global Change and Biodiversity).

References

Addor, N., and E. M. Fischer (2015), The influence of natural variability and interpolation errors on bias characterization in RCM simulations, *J. Geophys. Res. Atmos.*, *120*, 180–10,195, doi:10.1002/2014JD022824.

Addor, N., and J. Seibert (2014), Bias-correction for hydrological impact studies—Beyond the daily perspective, *Hydrol. Process.*, *28*, 4823–4828, doi:10.1002/hyp.10238.

Addor, N., O. Rössler, N. Köplin, M. Huss, R. Weingartner, and J. Seibert (2014), Robust changes and sources of uncertainty in the projected hydrological regimes of Swiss catchments, *Water Resour. Res.*, *50*, 7541–7562, doi:10.1002/2014WR015549.

Ban, N., J. Schmidli, and C. Schär (2014), Evaluation of the convection-resolving regional climate modeling approach in decade-long simulations, *J. Geophys. Res. Atmos.*, *119*, 7889–7907, doi:10.1002/2014JD021478.

Bárdossy, A., and G. Pegram (2011), Downscaling precipitation using regional climate models and circulation patterns toward hydrology, *Water Resour. Res.*, *47*, 1–18, doi:10.1029/2010WR009689.

Bellprat, O., S. Kotlarski, D. Lüthi, and C. Schär (2013), Physical constraints for temperature biases in climate models, *Geophys. Res. Lett.*, *40*, 4042–4047, doi:10.1002/grl.50737.

Blenkinsop, S., P. D. Jones, S. R. Dorling, and T. J. Osborn (2009), Observed and modelled influence of atmospheric circulation on central England temperature extremes, *Int. J. Climatol.*, *1660*, 1642–1660, doi:10.1002/joc.

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, *Clim. Dyn.*, *33*(6), 849–868, doi:10.1007/s00382-009-0588-6.

Cattiaux, J., H. Douville, and Y. Peings (2013), European temperatures in CMIP5: Origins of present-day biases and future uncertainties, *Clim. Dyn.*, *41*(11–12), 2889–2907, doi:10.1007/s00382-013-1731-y.

CH2011 (2011), Swiss climate change scenarios CH2011, edited by C2SM, MeteoSwiss, ETH, NCCR Climate, and OcCC, Zurich, Switzerland.

Chen, C., J. O. Haerter, S. Hagemann, and C. Piani (2011), On the contribution of statistical bias correction to the uncertainty in the projected hydrological cycle, *Geophys. Res. Lett.*, *38*, L20403, doi:10.1029/2011GL049318.

Chen, J., F. P. Brissette, and P. Lucas-Picher (2015), Assessing the limits of bias-correcting climate model outputs for climate change impact studies, *J. Geophys. Res. Atmos.*, *120*, 1123–1136, doi:10.1002/2014JD022635.

Christensen, J. H., and O. B. Christensen (2007), A summary of the PRUDENCE model projections of changes in European climate by the end of this century, *Clim. Change*, *81*(S1), 7–30, doi:10.1007/s10584-006-9210-7.

Fischer, A. M., A. P. Weigel, C. M. Buser, R. Knutti, H. R. Künsch, M. A. Liniger, C. Schär, and C. Appenzeller (2012), Climate change projections for Switzerland based on a Bayesian multi-model approach, *Int. J. Clim.*, *32*(15), 2348–2371, doi:10.1002/joc.3396.

Frei, C. (2013), Interpolation of temperature in a mountainous region using nonlinear profiles and non-Euclidean distances, *Int. J. Clim.*, *34*(5), 1585–1605, doi:10.1002/joc.3786.

Frei, C., J. Christensen, M. Déqué, D. Jacob, R. Jones, and P. L. Vidale (2003), Daily precipitation statistics in regional climate models: Evaluation and intercomparison for the European Alps, *J. Geophys. Res.*, *108*(D3), 4124, doi:10.1029/2002JD002287.

Giorgi, F., and W. J. Gutowski Jr. (2015), Regional Dynamical Downscaling and the CORDEX Initiative, *Annu. Rev. Environ. Resour.*, *40*(1), 467–490, doi:10.1146/annurev-environ-102014-021217.

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, *Hydrol. Earth Syst. Sci.*, *16*(9), 3383–3390, doi:10.5194/hess-16-3383-2012.

Haylock, M. R., N. Hofstra, A. M. G. Klein Tank, E. J. Klok, P. D. Jones, and M. New (2008), A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006, *J. Geophys. Res.*, *113*, D20119, doi:10.1029/2008JD010201.

Huth, R., C. Beck, A. Philipp, M. Demuzere, Z. Ustrnul, M. Cahynová, J. Kyselý, and O. E. Tveito (2008), Classifications of atmospheric circulation patterns: Recent advances and applications, *Ann. N. Y. Acad. Sci.*, *1146*, 105–152, doi:10.1196/annals.1446.019.

James, P. M. (2006), An assessment of European synoptic variability in Hadley Centre Global Environmental models based on an objective classification of weather regimes, *Clim. Dyn.*, *27*, 215–231, doi:10.1007/s00382-006-0133-9.

Li, C., E. Sinha, D. E. Horton, N. S. Diffenbaugh, and A. M. Michalak (2014), Joint bias correction of temperature and precipitation in climate model simulations, *J. Geophys. Res. Atmos.*, *119*, 153–162, doi:10.1002/2014JD022514.

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

Maraun, D., T. J. Osborn, and H. W. Rust (2011), The influence of synoptic airflow on UK daily precipitation extremes. Part II: Regional climate model and E-OBS data validation, *Clim. Dyn.*, *39*, 287–301, doi:10.1007/s00382-011-1176-0.

Muerth, M. J., B. Gauvin St-Denis, S. Ricard, J. A. Velázquez, J. Schmid, M. Minville, D. Caya, D. Chaumont, R. Ludwig, and R. Turcotte (2013), On the need for bias correction in regional climate scenarios to assess climate change impacts on river runoff, *Hydrol. Earth Syst. Sci.*, *17*(3), 1189–1204, doi:10.5194/hess-17-1189-2013.

Noguer, M., R. Jones, and J. Murphy (1998), Sources of systematic errors in the climatology of a regional climate model over Europe, *Clim. Dyn.*, *14*(10), 691–712, doi:10.1007/s003820050249.

Perez, J., M. Menendez, F. J. Mendez, and I. J. Losada (2014), Evaluating the performance of CMIP3 and CMIP5 global climate models over the north-east Atlantic region, *Clim. Dyn.*, 1–18, doi:10.1007/s00382-014-2078-8.

Philipp, A., J. Bartholy, C. Beck, M. Erpicum, P. Esteban, X. Fettweis, R. Huth, P. James, S. Jourdain, and F. Kreienkamp (2010), Cost733cat—A database of weather and circulation type classifications, *Phys. Chem. Earth, Parts A/B/C*, *35*(9–12), 360–373, doi:10.1016/j.pce.2009.12.010.

Philipp, A., C. Beck, R. Huth, and J. Jacobeit (2014), Development and comparison of circulation type classifications using the COST 733 dataset and software, *Int. J. Climatol.*, doi:10.1002/joc.3920.

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, *J. Hydrol.*, *395*(3–4), 199–215, doi:10.1016/j.jhydrol.2010.10.024.

Plavcová, E., and J. Kyselý (2011), Evaluation of daily temperatures in Central Europe and their links to large-scale circulation in an ensemble of regional climate models, *Tellus A*, *63*(4), 763–781, doi:10.1111/j.1600-0870.2011.00514.x.

Prudhomme, C., and M. Genevier (2011), Can atmospheric circulation be linked to flooding in Europe?, *Hydrol. Process.*, *25*(7), 1180–1190, doi:10.1002/hyp.7879.

Racherla, P. N., D. T. Shindell, and G. S. Faluvegi (2012), The added value to global model projections of climate change by dynamical downscaling: A case study over the continental U.S. using the GISS-ModelE2 and WRF models, *J. Geophys. Res.*, *117*, D20118, doi:10.1029/2012JD018091.

Riediger, U., and A. Gratzki (2014), Future weather types and their influence on mean and extreme climate indices for precipitation and temperature in Central Europe, *Meteorol. Zeitschrift*, *23*(3), 231–252, doi:10.1127/0941-2948/2014/0519.

Rohrer, M. (2013), Climate change and circulation types in the alpine region, Scientific Report MeteoSwiss 91. [Available at http://www.meteoschweiz.admin.ch/content/dam/meteoswiss/en/Ungebundene-Seiten/Publikationen/Scientific-Reports/doc/sr91rohrer.pdf.]

Schiemann, R., and C. Frei (2010), How to quantify the resolution of surface climate by circulation types: An example for Alpine precipitation, *Phys. Chem. Earth, Parts A/B/C*, *35*(9–12), 403–410, doi:10.1016/j.pce.2009.09.005.

Shepherd, T. G. (2014), Atmospheric circulation as a source of uncertainty in climate change projections, *Nat. Geosci.*, *7*(10), 703–708, doi:10.1038/ngeo2253.

Stahl, K., and S. Demuth (1999), Linking streamflow drought to the occurrence of atmospheric circulation patterns, *Hydrol. Sci. J.*, *44*(3), 467–482, doi:10.1080/02626669909492240.

Teng, J., N. J. Potter, F. H. S. Chiew, L. Zhang, B. Wang, J. Vaze, and J. P. Evans (2015), How does bias correction of regional climate model precipitation affect modelled runoff?, *Hydrol. Earth Syst. Sci.*, *19*, 711–728, doi:10.5194/hess-19-711-2015.

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, *J. Hydrol.*, *456–457*, 12–29, doi:10.1016/j.jhydrol.2012.05.052.

Teutschbein, C., and J. Seibert (2013), Is bias correction of regional climate model (RCM) simulations possible for non-stationary conditions?, *Hydrol. Earth Syst. Sci.*, *17*(12), 5061–5077, doi:10.5194/hess-17-5061-2013.

Themeßl, M. J., A. Gobiet, and A. Leuprecht (2011), Empirical-statistical downscaling and error correction of daily precipitation from regional climate models, *Int. J. Clim.*, *31*(10), 1530–1544, doi:10.1002/joc.2168.

Thrasher, B., E. P. Maurer, C. McKellar, and P. B. Duffy (2012), Technical note: Bias correcting climate model simulated daily temperature extremes with quantile mapping, *Hydrol. Earth Syst. Sci.*, *16*, 3309–3314, doi:10.5194/hess-16-3309-2012.

Turco, M., A. Sanna, S. Herrera, M. C. Llasat, and J. M. Gutiérrez (2013), Large biases and inconsistent climate change signals in ENSEMBLES regional projections, *Clim. Change*, *120*(4), 859–869, doi:10.1007/s10584-013-0844-y.

Uppala, S. M., et al. (2005), The ERA-40 re-analysis, *Q. J. R. Meteorol. Soc.*, *131*(612), 2961–3012, doi:10.1256/qj.04.176.

van der Linden, P., and J. F. B. Mitchell (Eds) (2009), ENSEMBLES: Climate change and its impacts: Summary of research and results from the ENSEMBLES project, Met Office Hadley Centre, UK.

van Ulden, A., G. Lenderink, B. van den Hurk, and E. van Meijgaard (2007), Circulation statistics and climate change in Central Europe: PRUDENCE simulations and observations, *Clim. Change*, *81*(S1), 179–192, doi:10.1007/s10584-006-9212-5.

van Ulden, A. P., and G. J. van Oldenborgh (2006), Large-scale atmospheric circulation biases and changes in global climate model simulations and their importance for climate change in Central Europe, *Atmos. Chem. Phys.*, *6*(4), 863–881, doi:10.5194/acp-6-863-2006.

Wilby, R. L., and N. W. Quinn (2013), Reconstructing multi-decadal variations in fluvial flood risk using atmospheric circulation patterns, *J. Hydrol.*, *487*, 109–121, doi:10.1016/j.jhydrol.2013.02.038.

Wilcke, R. A. I., T. Mendlik, and A. Gobiet (2013), Multi-variable error correction of regional climate models, *Clim. Change*, *120*(4), 871–887, doi:10.1007/s10584-013-0845-x.