Final Response to Referees

We greatly thank the reviewers for the careful and thorough reading of our manuscript. The additional clarifications and constructive suggestions have certainly helped to improve the quality of our manuscript. The comments have been carefully considered and responded. Please find below our response to each comment.

Response to Referee #1

Mayor points:

1. The underlying assumption of the presented exercise is that orographically defined classes are informative for the model's precipitation bias. In my opinion, this has not yet been convincingly shown. What would be required, for instance, is an analysis of the range of model biases WITHIN the individual orographic classes. Do classes separate from each other in such an analysis? Figure 7 provides an indication that this is not the case, as the spatial correlation does not systematically improve after application of the bias correction.

RESPONSE:

We thank the reviewer for bringing up this concern. We agree that the main purpose of the correction method might still be a bit unclear and we would like to clarify this in more details in the following. With the present study, we would like to obtain a flexible correction that can be applied to several different climate states at the same time. To obtain this, the correction method should not be constrained to the actual climate too much, this is, because circulation changes and atmospheric characteristics may be variable between different climates. We agree that cluster analysis of precipitation and its errors should be applied, so that errors can be grouped accordingly and to keep the error within classes as small as possible, to obtain an optimal correction result. This has for example been performed by Gomez et al. (2018) for Switzerland. The drawback of such a correction for our purpose is that such a cluster analysis is always based on the characteristics and circulation of the current climate and this is what we would like to avoid as much as possible. To be as much independent from current climates as possible and to still provide a correction that still touches upon important characteristics in the Alpine climate, we came up with "static" characteristics, i.e. topography height and orientation. Both, topography and orientation will remain similar during different climate states, even if we are aware of the fact that in any correction the effect of topography is implicitly included. Nevertheless, we would like to show here that biases have some orographic dependence. To clarify this, we have attached a figure that presents the monthly mean biases for each height-class before and after the correction (Fig. R1). Figure R1 illustrates an overestimation at high elevations and an underestimation at the lower ones during the colder months. Moreover, different levels of underestimation are observed across the height-classes during the warmer months. Thus, the splitting into different height-classes is appropriate to be used in the bias correction. Moreover, we would like to mention that we explicitly present the model biases within two classes in Fig. 3 (of the manuscript), and implicitly for all the height classes in Fig. 4 and 5. Therefore, we have included a more balanced discussion about our approach in the results part of the revised manuscript.

Furthermore, we agree that the spatial correlation is only weakly improved. However, we would like to highlight here that we do not only consider the spatial correlation to assess the performance of the different corrections, but we also include the spatial standard deviation and the spatial root-mean-square-error. The Taylor-diagram in Fig. 7 (of the manuscript) shows all three parameters and thus, provides wider criteria than just considering spatial correlation.

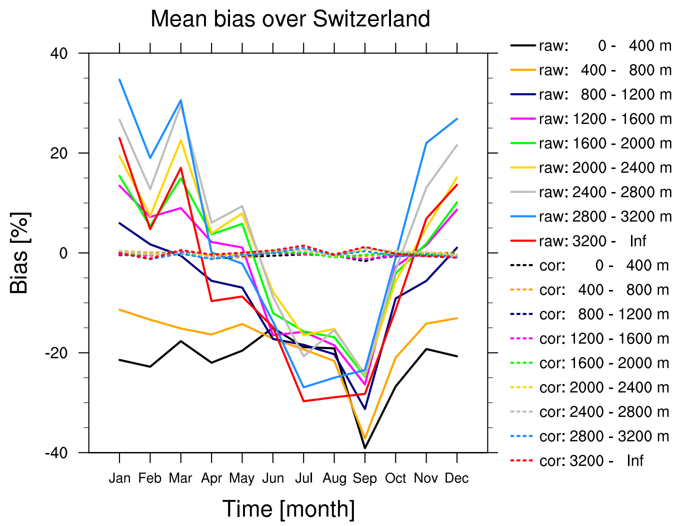


Figure R1. Mean bias over Switzerland for different height-classes.

2. As stated by the authors, the rationale behind the newly developed method is that bias correction would be possible for paleo climatic states subject to a different land surface topography (Alpine ice shield, for instance). There is a considerable danger that applying a correction method that is trained in today's climate does not hold for such a climatic state even if orography is considered as a co-variate in the bias correction. Large scale flow conditions, for instance, could be strongly different from today's conditions leading to a completely different bias structure even for the same orography class. Also, in a much colder climate the relation of snowfall to liquid precipitation would increase which might, in turn, lead to completely different model biases even for the same orographic class. To show that the assumption is valid, one would have to go much further with the modelling exercise. One could, for instance, carry out a second simulation with the very same GCM forcing but a modified Alpine topography in the RCM, and then apply the bias correction calibrated in the standard simulation with true orography. Would the bias-correction produce a realistic precipitation pattern in such a disturbed simulation?

RESPONSE:

We appreciate this comment and recognize that the manuscript might lead to misunderstandings about the application of our bias-correction method to other climate states. The danger of correcting biases in a simulated climate with a method that has been trained with a climate that does not correspond to the simulated climate is well-known in statistical downscaling methods. These are likewise calibrated with today's climate and applied to past and future climate states. Many statistical downscaling and correction methods suffer basically from the assumption of stationary biases, which implies that their algorithms trained with today's climate are considered to be also valid for different climate states. Thus, our work aims at presenting a new bias-correction that attempts to decrease this danger by using orographic features, which are less likely characteristics of the current climate only. Moreover, precipitation biases are not only produced by initial and boundary conditions provided by the global climate models, but also by parametrisations, physical and numerical formulations that are described in both global and regional climate models. The main goal of the presented work is to correct wet or dry biases that stem either from global or regional models or both. These biases can be produced by parametrisations and numerical formulations, but those that are mainly associated with orographic effects, namely, vertical motion leading to precipitation. To clarify this, we extended the discussion on the general shortcomings of bias correction methods in the introduction and the conclusion section in the revised manuscript.

Furthermore, we agree that the relation of snowfall – liquid precipitation would change in a much colder climate. However, this relation plays a negligible role in our correction method because the observational dataset and the model output, which are used in this work, consider both solid and liquid precipitation together. To clarify these points, we have included the definition of the precipitation and the days without precipitation in the manuscript as follows.

- Additional text on page 4 line 17

…Note that all data sets consider daily precipitation as total precipitation, i.e., both solid and liquid precipitation, and convective and non-convective precipitation. Moreover, days without precipitation are treated as censored values, i.e., not considered in analysis, when daily precipitation is equal to 0 mm day^{-1}, although in the case of observations this is equivalent to 0.1 mm day^{-1} due to gauge precision …

The suggested sensitivity simulation would provide several problems. First, the global simulation would have to be rerun with an adapted alpine topography, as a circulation change should be expected when the Alps are reduced and increased. If inconsistent boundary conditions are given to the regional model this might lead to further errors that cannot be corrected by the proposed correction method. Second, this correction cannot be validated as there are no observations for such a climate, so the same problem as for past and future climates remains. Thus, we have used a different Alpine region to calibrate the correction method, which is considered as a different climate state due to its different precipitation pattern compared to the one from Switzerland (Frei and Schär, 1998). In addition, the corrected results can be easily evaluated using the gridded Swiss observational dataset, which is not the case in the suggested sensitivity.

Still, we agree that the method should be evaluated in a different climate state but this is beyond the scope of this publication. An idea to validate the proposed correction may be to simulate e.g. Last Glacial Maximum conditions and compare them to proxy data like alpine ice sheet extent. Such a validation would include some collaboration with glacier modellers that are able to use raw and corrected precipitation to predict glacier extents. We think that such a method could provide a good way to verify the presented method as proxies could mimic the missing observations during glaciated times.

Frei, C., and C. Schär. 1998. 'A Precipitation Climatology of the Alps from High-Resolution Rain-Gauge Observations'. International Journal of Climatology 18 (8): 873–900. https://doi.org/10.1002/(SICI)1097-0088(19980630)18:8<873::AID-JOC255>3.0.CO;2-9.

3. The introduction definitely needs to be worked on and be streamlined. It currently includes quite some repetition, and the line of argumentation is not always straight. Some basic references (for instance on the evaluation of CORDEX experiments in Europe and over the Alps) are missing.

RESPONSE:

We greatly thank you for bringing to our attention that the introduction needs to be worked on. An improved introduction is presented in the revised manuscript avoiding repetitions.
Regarding the basic references, we would like to clarify that we point out the CORDEX experiments twice in the manuscript. First, it was brought up on page 2 line 14 when linking the precipitation biases with regional climate simulations. Second, we cited the work of Casanueva et al. (2016) on page 11 line 5, which is about an approach of correcting precipitation biases from some EURO-CORDEX RCMs. They mainly focus on Spain and the Alpine region.
Nevertheless, we agree that the CORDEX experiments are not fully mentioned in the manuscript and that they could be better introduced. Thus, we have included them more explicitly in the next version of the manuscript.

4. At several points in the paper the authors mention that the traditional QM approach would calibrate one correction function for the entire domain. This is certainly not true. In a pure bias correction setting (raw grid = target grid) a separate correction function is calibrated for each individual grid cell.

RESPONSE:

We fully agree that this statement needs to be considered for a reformulation, although a pure bias correction setting as mentioned by the reviewer (separate correction function calibrated for each grid point) would be also a statistical downscaling. Still, we have rephrased “commonly used method” into “simple approach” at various places throughout the manuscript and deleted some citations as follows:

- Page 6 lines 4 – 6:

…To demonstrate the improvement of using the new method, we further compare it to a commonly used method that is carried out without orographic features and uses TFs deduced for the entire region of Switzerland (2 km) (similar to Berg et al., 2012; Maraun, 2013; Fang et al., 2015) …

…To demonstrate the improvement of using the new method, we further compare it to a simple method that is carried out without orographic features and uses TFs deduced for the entire region of Switzerland (at 2 km resolution, 12 TFs in total) …

- Page 8 lines 5 – 7:

…We assess in the following, which of these characteristics are necessary to improve the simple approach of applying one EQM to the entire domain, often used in studies for present day and future climate change (e.g., Evans et al., 2017; Li et al., 2017; Ivanov et al., 2018) …

… We assess in the following, which of these characteristics are necessary to improve a simple approach of applying one EQM to the entire domain, where orographic features are not considered …

- Page 11 lines 11 – 12

…Clearly, the new method outperforms the standard method of applying one EQM transfer function deduced for the entire region of interest, which is commonly used (Berg et al., 2012; Maraun, 2013; Fang et al., 2015) …

…Clearly, the new method outperforms the simple method of applying one EQM transfer function that is deduced for the entire region of interest and does not consider any orographic features …

5. The reason for the second bias correction step (first part of local intensity scaling) remains completely unclear to me. The third step (QM) would account for this already (by adjusting the percentiles).

RESPONSE:

We agree that the reason for the local intensity scaling method was not fully explained. To clarify this point, it is necessary to mention the similarities and differences in the treatment of the very low intensity values between two quantile mapping techniques, namely, the parametric quantile mapping (QM) and the empirical quantile mapping (EQM). Both techniques treat days without precipitation as censored values and consider only days with precipitation. The QM obtains the quantiles and transfer functions (TFs) from a cumulative distribution function (CDF) that is previously fitted, and thus it could properly handle the very low values with an adequate distribution fitting. Whereas in our study, an empirical CDF is used to directly calculate the quantiles and TFs, which is the core of the EQM. The reason of using an EQM is because we do not assume any known distribution either in our data sets or in the possible application to other climate states. However, the results of the EQM can become unrealistic if the very low intensity values are not adjusted previously. The reason for this is that these values can produce inappropriate TFs due to an important shift in the distribution, i.e., the quantiles (Teutschbein and Seibert, 2012; Lafon et al. 2013).

To adjust these very low values, an additional parameter is included in the definition of days without precipitation that has been mentioned before in the response to the second major point. The days without precipitation are not considered for calculating the TFs when they fall below a certain threshold. Many studies use a static threshold for the entire data set which is between 0.01 and 1.00 mm day^{-1}, whereas in our study, we calculate a static threshold for each group (or subgroup) and months of the year. This allows to be the consistent with the different biases-treatment across the groups (or subgroups) and months of the year. The threshold is calculated using the local intensity scaling method and can vary in our study from 0.001 to 1.00 mm day^{-}

Changes in the manuscript are presented as follows:

- Page 5 lines 13 – 14

…2010). To correct precipitation with very low-intensity the first part of the local intensity scaling method is used (Schmidli et al., 2006). It consists …

…2010), which can distort the precipitation distribution substantially, i.e., shifting the quantiles, producing inappropriate corrections in the third step when EQM is applied (Teutschbein and Seibert, 2012; Lafon et al., 2013). To correct precipitation with very low intensity, an additional parameter is included in the definition of dry days related with the uncorrected precipitation that is described in the section of model and data before. Dry days are not considered for calculating the TFs when they fall below a certain threshold. Many studies use a static threshold for the entire data set which is between 0.01 and 1.00 mm day^{-1} (Piani et al., 2010a; Lafon et al., 2013; Maraun, 2013). We calculate a static threshold for each group (or subgroup) and months of the year. This allows to be the consistent with the different biases-treatment across the groups (or subgroups) and months of the year. Then, we carry out the local intensity scaling method (Schmidli et al., 2006) that is also used by Teutschbein and Seibert (2012) before using the quantile mapping technique. This method consists …

- Page 5 lines 16 – 17

…The threshold can vary from group to group, but it is often close to or smaller than 1 mm day^{-1} Schmidli et al., 2006).

…In our work, the threshold can vary from group to group and from month to month between 0.001 and 1 mm day^{-1}, similar to Schmidli et al. (2006) …

6. The general setup of the bias correction remains unclear. Is the correction carried out grid cell by grid cell, or in a bulk manner for each orographic class?

RESPONSE:

We thank the reviewer for bringing to our attention that the general setup of the bias correction remains unclear. To clarify it we have changed lines 31 – 32 on page 5 as follows:

…To combine all steps, the EQM is applied to each (sub-) group and each month of the year, separately. This results in a set of TFs for each (sub-) group and each month of the year. Thus…

…To combine all steps, the local intensity scaling method and the EQM are applied to each (sub-) group defined in the first step and each month of the year, separately, by pooling all grid points that belong to it and handling them as a single distribution of daily precipitation. This results in a set of TFs for each (sub-) group and each month of the year. For instance, when the correction is carried out using height-classes of 400 m, a TF is defined for each height group, resulting in nine TFs for each month and in total 108 TFs throughout the year. Moreover, the correction is afterwards applied to the daily precipitation at every grid point using the TFs that are common to all elements within the same group (or sub-group) and month. Thus…

7. Figure 3 is unclear. What do the boxplots represent and what is the true y-axis scale? Do the boxplots cover the spatial variability of monthly mean precipitation for the entire domain (a) or the elevation classes (b,c)? The text mentions that daily precipitation variability is shown, but how does this aggregate to monthly precipitation (y-axis label) then?

If boxplots really show the distribution of daily precipitation values does it really make sense to use the IQR? Depending on the wet day frequency more than 25% of the days might be dry, for instance.

RESPONSE:

We appreciate that you bring to our attention that the y-axis, the caption and the text are confusing. To clarify this, we would like to mention that the boxplots illustrate the spatial distribution of monthly mean values of precipitation intensity across a specific area within 30 years. Thus, we have modified them as follows:

- The y-axis

Monthly precipitation [mm day^{-1}]

Precipitation intensity [mm day^{-1}]

- The text in the caption

Boxplots are illustrating the annual cycle and monthly distribution of daily precipitation: (a) entire Switzerland, (b) all grid points in the height class of 400 – 800 m, and (c) of 2.800 – 3.200 m. Black box-plots represent the observations (RhiresD data), blue and red ones the raw and corrected simulation, respectively. Top and bottom ends of the dashed lines represent the maximum and minimum values, respectively. Dots represent the mean.

Boxplots illustrate the spatial distribution of monthly mean values of precipitation intensity across a specific area within 30 years: (a) the area covers all grid points over entire Switzerland, (b) the grid points in the height class of 400 – 800 m, and (c) the grid points in the height class of 2.800 – 3.200 m. Black box-plots represent the observations (RhiresD data), blue and red ones the raw and corrected simulation, respectively. Top and bottom ends of the dashed lines represent the maximum and minimum values, respectively. Dots represent the spatial climatological mean value.

- Text, page 6 line 19 – 20 modified and moved to the beginning of the paragraph.

…, the annual cycle and the monthly distributions of daily precipitation are estimated for different height-classes …

…The annual cycle and the distributions of monthly mean precipitation intensity are for different height-classes to…

- Text, page 6 line 32 – 33

… For these example months, we present the patterns of biases in precipitation, changes in the distribution of daily precipitation, illustrated by the interquartile range as well as biases in wet-day frequency …

… For these example months, we present the spatial patterns of the biases in the monthly mean precipitation intensity, in the variability illustrated by the interquartile range, and in the wet-day frequency …

8. Also the general validation setup remains unclear to some extent, the validation technique and the respective reference datasets used needs to be better described. It is sometimes unclear whether the Swiss 2 km serves as reference or the Alpine 5 km grid.

RESPONSE:

We agree that the validation technique and the data sets used are not fully described. To clarify it, we have modified it as follows:

- Page 5 lines 33 – 35 and page 6 lines 1 – 6

…To come up with a final method for the Alpine region we first test the influence of the different orographic characteristics (step 1). To be consistent with former studies (e.g., Sun et al., 2011; Themessl et al., 2012; Wilcke et al., 2013; Rajczak et al., 2016), the evaluation of the new method first uses the same region where the TFs are estimated. To be more rigorous, we additionally apply a cross-validation: Thereby, Switzerland is defined as the area to be corrected; then, we calculate two different TFs; namely, from the same Swiss region called Internal TFs (Int-TF), and from the corresponding Alpine region of Germany, France, and Austria altogether called External TFs (Ext-TF) (Fig. 1c). Note that Ext-TFs are carried out at 5 km horizontal resolution. To demonstrate the improvement of using the new method, we further compare it to a commonly used method that is carried out without orographic features and uses TFs deduced for the entire region of Switzerland (2 km) (similar to Berg et al., 2012; Maraun, 2013; Fang et al., 2015) …

… To come up with a final method for the Alpine region, we first evaluate the influence of the different orographic characteristics (step 1). To be consistent with former studies (e.g., Sun et al., 2011; Themessl et al., 2012; Wilcke et al., 2013; Rajczak et al., 2016), the evaluation uses the same region where the TFs are estimated. Explicitly, this means that the Swiss region in the WRF output (2 km) is defined as the area to be corrected and the RhiresD data set (at 2 km resolution) is used to obtain the TFs and to evaluate the different correction methods. These TFs are called Internal TFs (Int-TF) during the cross-validation process later on. Once the final method is determined, we additionally apply a cross-validation to test the method more rigorously: Thereby, Switzerland is defined as the area to be corrected (WRF output at 2 km resolution); in addition to the Int-TF (see above), which uses the same region to define TFs and to apply the correction, we also calculate a second set of TFs. The second one is obtained from the corresponding Alpine region of Germany, France, and Austria altogether called External TFs (Ext-TF) using the APGD data set (at 5 km resolution; Fig. 1c). Note that Ext-TFs are carried out at 5 km horizontal resolution and applied to Switzerland at 2 km resolution. To demonstrate the improvement of using the new method, we further compare it to a simple method that is carried out without orographic features and uses TFs deduced for the entire region of Switzerland (at 2 km resolution, 12 TFs in total) …

9. Any kind of bias correction will only be as good and as appropriate as the observational reference. The validity of an analysis of elevation dependencies and slope

dependencies at regional scales in the gridded observational precipitation datasets needs to be discussed. Does the reference grid really represent such dependencies?

<u>RESPONSE</u>:

We appreciate this comment. We agree that we missed to show the validity of the elevation and slope dependencies in the gridded observational data sets. Note that the observational data sets have a height dependence on its quality. As mentioned by (Isotta, 2014), the gridded observational data sets do not only present errors due to the interpolation methods, but they also show errors that may differ in quantity from one to the other station (Sevruk, 1985; Richter, 1995) and are related to the "gauge undercatch", whose magnitudes range from 5% over the flatland regions to 30% above 1500 m a.s.l..To clarify this, a discussion is presented in results section of the revised manuscript.

Note that the observational data sets are considered generally reliable and represent orographic features well, although at high altitudes less data sets are available (Fig. R2; Isotta et al. 2014). Note that in this study we do not explicitly consider any uncertainty, and instead assume that these observations represent the true precipitation without errors. Still, we have discussed the uncertainty issue in particular for the results in high altitudes.

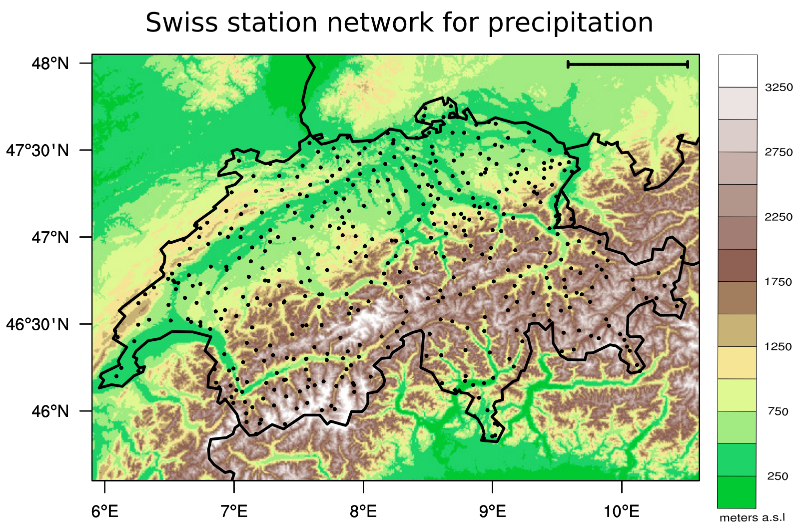


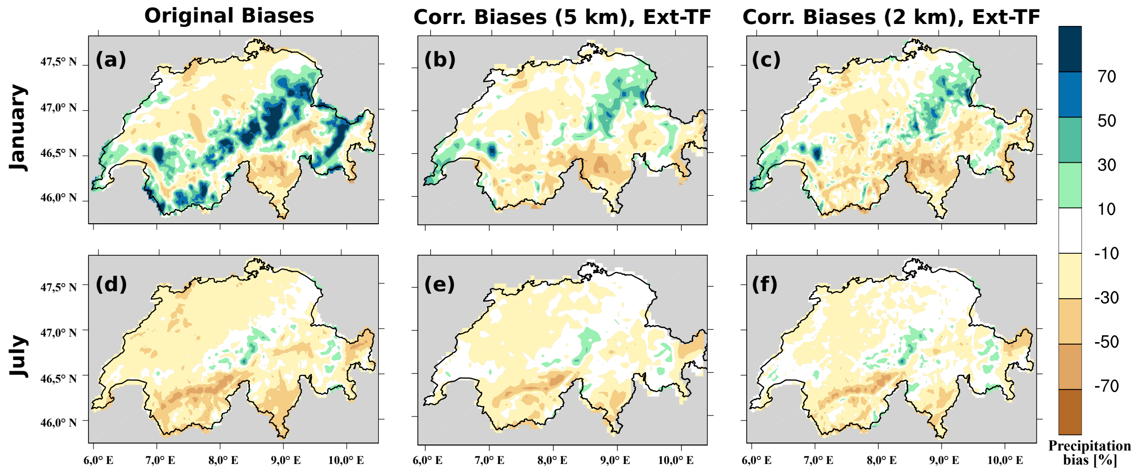
Figure R2. Swiss stations are integrated in RhiresD.

10. The application of the Ext-TFs mixes spatial scales (classes based on 5 km orography vs. classes based on 2 km orography). This is potentially dangerous and the effects of this mismatch should be shown. Why is the validation, in this case, not carried out on the 5 km scale as well?

<u>RESPONSE</u>:

We thank you for highlighting this point. To clarify it, we would like to mention that the method uses different observational data sets. We used the 5 km classes applied to 2 km target as we directly compare the results with the ones obtained from the application of Int-TFs and to avoid any additional uncertainty produced by interpolating between the two grids. Another reason is that the application at 5 km show minimal differences on the results, as is shown in the next

Figure R3. Therefore, we have mentioned this experiment in the results part of the revised manuscript but its figures are not shown because of the minimal differences.



R3. Biases in the climatological mean value of precipitation intensity over Switzerland. (a) represents the original biases in January, (b) the biases after being corrected at 5 km using Ext-TFs in January, (c) the biases after being corrected at 2 km using Ext-TFs in January, (d), (e), and (f) as (a), (b), and (c) but in July, respectively.

Minor points:

a) ***page 1 line 19: "is" instead of "has been"***

We thank the reviewer for the suggestion. We have changed it in the manuscript.

b) ***page 2 line 20: What is meant by "weaker intensity" here? Unclear***

It means that the simulated precipitation intensity is weaker than the observational one. As an example, instead of 20 mm day^{-1} the simulated precipitation intensity is 5 mm day^{-1}. To make this point clear, we have modified it as follows.

… with a weaker intensity …

… with a lower intensity …

c) ***page 2 lines 16-19: Line of argumentation unclear. RCMs were already referred to just above (line 12ff)***

We agree and to make the argumentation clearer, we have re-structured the paragraphs as also suggested in Major point #3 and the change is presented in the revised manuscript.

d) ***page 4 lines 1-2: No true in general. Ban et al. for instance show that mean precipitation can also be much worse in convection resolving experiments. Certain aspects (such as the diurnal cycle) are improved, but not all.***

We agree that the statement in these lines is not in general true. To correct it, we have modified it as follows:

…Convection permitting model resolutions are preferred as recent studies show a better performance in simulating precipitation (e.g., Ban et al., 2014; Prein et al., 2015) …

…Convection permitting model resolutions are in general preferred as many recent studies show a better performance in simulating precipitation (e.g., Ban et al., 2014; Prein et al., 2015; Kendon et al., 2017; Berthou et al., 2018; Finney et al., 2019). However, we shall keep in mind that some biases in temperature and cloud formation may be produced by this set up, which may lead to additional biases in precipitation as shown in Ban et al. (2014) …

e) ***page 4 lines 4-7: I don't really understand the reason behind this splitting in ten single 3-year simulations. 2 months spin up is certainly not enough for soil parameters and snow. Some more information on the setup and on the rationale behind it needs to be provided.***

Splitting up the simulations can be explained by the time-consuming setup to run a simulation over the Alps at 2 km resolution over 30 years. Namely, 3 model years are equivalent to 1 month in real time, which means that a 30-years simulation in a single piece would have taken at least 10 months in real time without any interruption.

Regarding the spin-up, we would like to mention that WRF has only an atmospheric component that is fed by initial and boundary conditions obtained from the GCM. Moreover, we consider the ice cover and soil in a quasi-stable state, as they are initially provided by the GCM and because of its long simulation these variables are in equilibrium there and because the interactions with the atmosphere are fully parametrised in WRF. Thus, the spin-up time was considered only for the atmosphere, which requires a much shorter spin-up period that certainly does not exceed two months.

f) ***page 4 lines 19-20: I guess this is hardly true. In areas where no observations are available gridded products can be subject to very high uncertainties as inter- and extrapolation are required here.***

We agree that gridded products can be subject to important uncertainties in areas where there is no observation. To avoid misunderstandings, we have modified on page 4 the lines 18 – 20 as follows:

…The observational gridded data sets provide valuable insights, in particular in areas where observations are not possible due to extreme weather conditions or insufficient accessibility, such as mountain peaks. However, they also contain some discrepancies and uncertainties, e.g., high precipitation intensities are systematically underestimated and low intensities overestimated. …

…The observational gridded data sets provide valuable insights. However, they also contain some discrepancies and uncertainties due to inter- and extrapolation methods, e.g., high precipitation intensities are systematically underestimated and low intensities overestimated, especially in areas where observations are not available …

g) ***page 5 lines 4-9: It remains unclear how these classes are computed. Based on the relation of a grid cell to its 8 direct neighbour grid cells? Please clarify.***

We thank you for bringing to our attention that this parameter remains unclear. To make it clear, we would like to mention that the slope-orientation is obtained by a simple trigonometric function using the two variables that are directly calculated by WRF. Namely, we sum two vectors: the slope north-south vector and the slope west-east vector, which both come directly from WRF. Thus, we have added additional information in the manuscripts follows:

- Page 5 line 8

  …< 315). Note that this characteristic is obtained by summing the two slope vectors that are directly provided by WRF. Combining …

h) ***page 5 lines 15-17: Which threshold is then used in the present work?***

The threshold varies from group to group (or sub-group to sub-group) and from month to month. See major point 5.

i) ***page 7 lines 30-32: This explanation seems to be not very likely given the turnaround time of atmospheric water vapor (a couple of days only). Water vapor should also frequently be resupplied by the boundary forcing of the RCM. Can you back this up by some reference?***

We appreciate that you bring this point to the discussion and we agree that the explanation needs to be improved. To achieve that, we would first like to mention that the drizzle effect is mainly caused by the horizontal resolution and the physics in the model (e.g. Gutowski et al. 2003; Chen and Dai 2019), and it can be independent of resupplying by the boundary conditions. Moreover, we have modified the explanation as follows:

… wet-day frequency may also explain the underestimation of the extreme precipitation (Fig. 3) as moisture necessary for extreme precipitation events is removed via the drizzle effect …

…wet-day frequency may slightly contribute to the underestimation of the extreme precipitation (Fig. 3) as precipitable water necessary for extreme precipitation events is removed via the drizzle effect. Namely, the precipitable water available for a daily extreme precipitation event may be distributed over several days due to problems in the parameterisations of the cloud microphysical and precipitation processes as found in Knist et al. (2018). …

Chen, Di, and Aiguo Dai. 2019. ‘Precipitation characteristics in the Community Atmosphere Model and Their Dependence on Model Physics and Resolution’. Journal of Advances in Modeling Earth Systems 11 (7): 2352–74. https://doi.org/10.1029/2018MS001536.

Knist, Sebastian, Klaus Goergen, and Clemens Simmer. 2018. ‘Evaluation and Projected Changes of Precipitation Statistics in Convection-Permitting WRF Climate Simulations over Central Europe’. Climate Dynamics, February. https://doi.org/10.1007/s00382-018-4147-x.

Gutowski, William J., Steven G. Decker, Rodney A. Donavon, Zaitao Pan, Raymond W. Arritt, and Eugene S. Takle. 2003. ‘Temporal–spatial scales of observed and simulated precipitation in Central U.S. climate’. Journal of Climate 16 (22): 3841–47. https://doi.org/10.1175/1520-0442(2003)016<3841:TSOOAS>2.0.CO;2.

j) ***Figure 1: Why are Italy and Slovenia excluded from the Ext-TF analysis? They are part of the APGD dataset.***

Italy and Slovenia are excluded from the Ext-TF because of their poor station density covering the period 1979 – 2008 compared to the ones we used, especially over a complex topography and at high altitudes. This poor density could lead to more uncertainties in the dataset when representing the precipitation over complex topography, which could diminish the ability of the correction method. Therefore, we have included an explanation about this in the models and data section of the revised manuscript.

To clarify this, we show here two figures published in the website of Meteoswiss and in Isotta et al. (2014), respectively (Fig R4 and R5). Figure R4 and R5 show the station density used for creating the APGD data set. Moreover, Figure R4 presents the altitude of each station and Fig. R5 the time-covering fraction of the period 1971–2008 (Isotta et al. 2014).

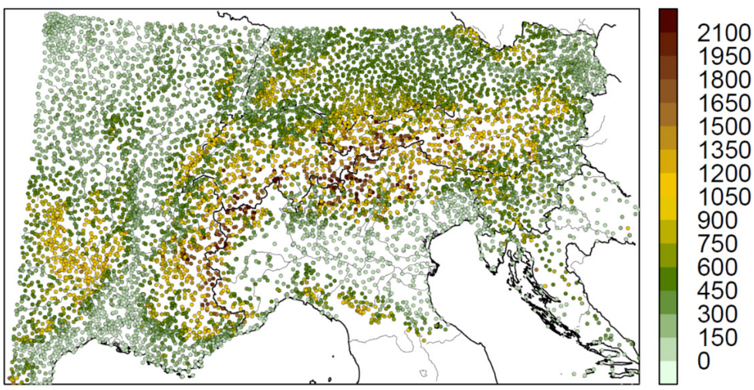


Figure R4. Each point corresponds to a rain-gauge station for which data was available in the the spatial analysis. The color is the height (m) of the station. Source: https://www.meteoswiss.admin.ch/home/search.subpage.html/en/data/products/2015/alpine-precipitation.html)

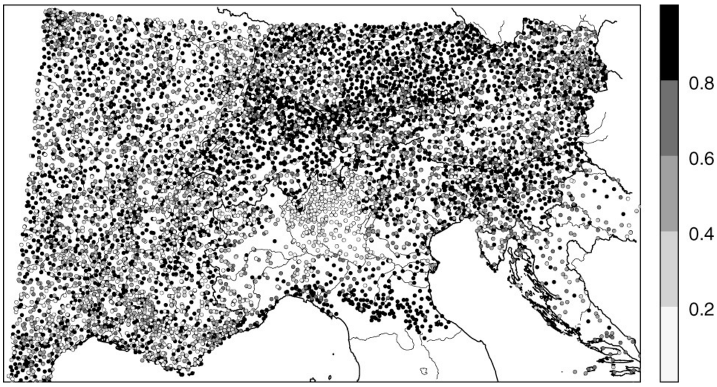


Figure R5. Distribution of stations from which records of daily precipitation are integrated in APGD dataset. Shading represents the fraction of the full period (1971–2008) covered by the respective record. (Isotta et al. 2014)

k) Figures 4 and 5: Sorry, but it is unclear to me which bias is shown in these two figures. Bias of the IQR of daily precipitation amount sin Figure 5? Which intensity in Figure 4? Mean wet day intensity? Needs to be better explained.

To clarify that, we have modified the captions of the three Figures as follow:

- Figure 4

Biases of precipitation in terms of intensity over Switzerland. (a) represents the original biases in January, (b) the biases after being corrected using Int-TFs in January, (c) the biases after being corrected using Ext-TFs in January, (d), (e), and (f) as (a), (b), and (c) but in July, respectively.

Biases in the climatological mean value of precipitation intensity over Switzerland. (a) represents the original biases in January, (b) the biases after being corrected using Int-TFs in January, (c) the biases after being corrected using Ext-TFs in January, (d), (e), and (f) as (a), (b), and (c) but in July, respectively.

- Figure 5

Biases of precipitation in terms of interquartile range over Switzerland. (a) represents the original biases in January, (b) the biases after being corrected using Int-TFs in January, (c) the biases after being corrected using Ext-TFs in January, (d), (e), and (f) as (a), (b), and (c) but in July, respectively.

Biases in the interquartile range of monthly mean precipitation intensity over Switzerland. (a) represents the original biases in January, (b) the biases after being corrected using Int-TFs in January, (c) the biases after being corrected using Ext-TFs in January, (d), (e), and (f) as (a), (b), and (c) but in July, respectively.

- Figure 6

Biases of precipitation in terms of wet-day frequency over Switzerland. (a) represents the original biases in January, (b) the biases after being corrected using Int-TFs in January, (c) the biases after being corrected using Ext-TFs in January, (d), (e), and (f) as (a), (b), and (c) but in July, respectively.

Biases in the wet-day frequency within the 30-year period over Switzerland. (a) represents the original biases in January, (b) the biases after being corrected using Int-TFs in January, (c) the biases after being corrected using Ext-TFs in January, (d), (e), and (f) as (a), (b), and (c) but in July, respectively.

<u>Response to Referee #2</u>

<u>Specific comments:</u>

1. ***The setup for the EMQ is completely unclear. Standard EQM is local, i.e. it would apply a different correction for each location for which observations are available, in this case for each gridcell of the observational datasets. There is no explanation of how the corrections for the subclasses (elevation and slope) are obtained. Are the local corrections averaged, or is the precipitation averaged prior to fitting the EQM?***
This is obviously a key aspect of the method and it is surprising that it is not explained.
The statement that standard bias correction methods do not include the effect of topography is wrong, as the observations, which are the basis for the fitting, do include these effects. What is presumably meant is that standard bias correction does not include these effects explicitly, which means it cannot be applied when the topography changes

<u>RESPONSE</u>:

We agree that the setup of the bias correction remains unclear. Still, we would like to point out that one strength of our method is that it is not local (the standard EQM described by the reviewer is a bias correction plus statistical downscaling). The simple reason is that a localized correction would fail in different states like the Last Glacial Maximum as valleys are filled with ice. To make the suggested method clearer, we have modified the manuscript as follows:

- Page 5 lines 31 – 32

…To combine all steps, the EQM is applied to each (sub-) group and each month of the year, separately. This results in a set of TFs for each (sub-) group and each month of the year. Thus…

…To combine all steps, the local intensity scaling method and the EQM are applied to each (sub-) group defined in the first step and to each month of the year, separately, by pooling all grid points that belong to each group and handling them as a single distribution of daily precipitation. This results in a set of TFs for each (sub-) group and each month of the year. For instance, when the correction is carried out using height-classes of 400 m, a TF is defined for each group, resulting in nine TFs for each month and in total 108 TFs throughout the year. Moreover, the correction is afterwards applied to the daily precipitation at every grid point using the TFs that are common to all elements within the same group (or sub-group) and month. Thus…

We also agree that the observational data sets implicitly include effects of topography. Changes regarding this point are presented in the following lines of the manuscript:

- Page 2 line 33

…correction methods do not consider orographic features that…

…correction methods only implicitly consider orographic features that…

- Page 3 line 6

…time includes orographic characteristics…

… explicitly combined orographic characteristics…

2. ***As already pointed out by the first reviewer, determining joint bias corrections for the subclasses defined by topography and slope only makes sense if the local bias corrections within a class are more similar than those between the classes. This needs to be shown***

RESPONSE:

We appreciate this comment and we agree that we missed to show clearly enough the argumentation for using different classes. As reviewer 1 asked a similar question we present there the same answer: We thank the reviewer for bringing up this concern. We agree that the main purpose of the correction method might still be a bit unclear and we would like to clarify this in more details in the following. With the present study, we would like to obtain a flexible correction that can be applied to several different climate states at the same time. To obtain this, the correction method should not be constrained to the actual climate too much, this is, because circulation changes and atmospheric characteristics may be variable between different climates. We agree that a cluster analysis of precipitation and its errors should be applied, so that errors can be grouped accordingly and to keep the error within classes as small as possible, to obtain an optimal correction result. This has for example been performed by Gomez et al. (2018) for Switzerland. The drawback of such a correction for our purpose is that such a cluster analysis is always based on the characteristics and circulation of the current climate and this is what we would like to avoid as much as possible. To be as much independent from current climates as possible and to still provide a correction that still touches upon important characteristics in the Alpine climate, we came up with “static” characteristics, i.e. topography height and orientation. Both, topography and orientation will remain similar during different climate states, even if we are aware of the fact that in any correction the effect of topography is implicitly included. Nevertheless, we would like to show here that biases have some orographic dependence. To clarify this, we have attached a figure that presents the monthly mean biases for each height-class before and after the correction (Fig. R1). Figure R1 illustrates an overestimation at high elevations and an underestimation at the lower ones during the colder months. Moreover, different levels of underestimation are observed across the height-classes during the warmer months. Thus, the splitting into different height-classes is appropriate to be used in the bias correction. Moreover, we would like to mention that we explicitly present the model biases within two classes in the Fig. 3 (of the manuscript), and implicitly for all the height classes in Fig. 4 and 5. Note that the biases within the classes are much smaller than between the classes. Therefore, we have included a more balanced discussion about our approach in results section of the revised manuscript

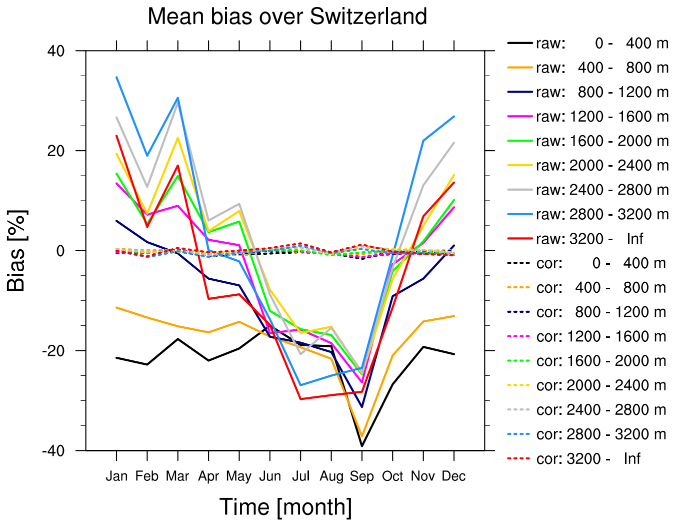


Figure R6. Mean bias over Switzerland for different height-classes.

3. The justification for the intended application is superficial and ignores key problems. In turn this means that the justification for the new approach itself is weak. As pointed out already by the first reviewer, many things in addition to the topography are different in a glacial climate, for instance the large-scale circulation or the moisture content. It is thus highly questionable whether applying a bias correction that is based on present climate, even if it explicitly accounts for topography, would yield meaningful results.
This problem is closely related to the distinction of different types of errors and to the issue of propagation of GCM errors through dynamical downscaling. There are a few statements in the paper that mention that discrepancies of RCM simulations and observations might be caused by the driving GCM. However, there is no systematic discussion of what kind of errors bias correction could correct in a meaningful way. A discussion of these issues can be found for instance in
Maraun, et al., 2017: Towards process-informed bias correction of climate change simulations. Nature Climate Change, 7(11), 764-773
Maraun and Widmann, 2018: Statistical downscaling and bias correction in climate research. Cambridge University Press, ISBN 1107066050
Eden, J.M., Widmann, M., Grawe, D, and Rast. S., 2012: Reassessing the skill of GCM-simulated precipitation. J. Climate, 25(11), 3970-3984.

RESPONSE:

We appreciate that the reviewer brings up the point that it might be misleading to what extent the presented bias-correction can be applied to other climate states. As already responded to reviewer 1, we would like to mention that the danger of correcting biases in a simulated climate with a method that has been trained with a climate that does not correspond to the simulated one is well-known in the statistical downscaling and correction methods. Statistical downscaling and correction methods suffer basically from the assumption of stationary biases, which implies that their algorithms trained with today's climate are considered to be also valid for different climate states. Thus, our work aims at presenting a new bias-correction that attempts to decrease this danger by using orographic features, which are less likely characteristics of the current climate only. Moreover, precipitation biases are not only produced by initial and boundary conditions provided by the global climate models, but also by parametrisations, physical and numerical formulations that are described in both global and regional climate models. The main goal of the presented work is to correct wet or dry biases that stem either from global or regional models or both. These biases can be produced by parametrisations and numerical formulations, but those that are mainly associated with orographic effects, namely, vertical motion leading to precipitation. To clarify this, we extended the discussion on the general shortcomings of bias correction methods in the conclusion section of the revised manuscript. Note that the presented correction is only applicable in regions where the topography is rather complex and where topography has certainly an influence on the local atmospheric circulation.

4. ***The fact that EQM leads to correct distributions for the fitting data is trivially true by construction. The informative part of the validation of statistical models is related to the aspects that are not trivially in agreement with observations. For each aspect of the validation it should be discussed to what extent a good skill can be expected by construction. For instance, given the unclear setup for fitting and application of the bias correction, it is not clear what causes the differences between observed and corrected distributions in Fig.3, or the differences in Fig. 4 and Fig. 5.***
Some problems related to the validation of bias correction methods are discussed in Maraun, D. and M. Widmann, 2018, 'Cross-validation of bias-corrected climate simulations is misleading', HESS, 22(9), 4867-4873.

RESPONSE:

We thank the reviewer for this comment and agree that the validation discussion can be improved. As noted by Bennett et al. (2014), the importance of cross-validation methods is that they can test the ability of bias-correction techniques on a different climate state. However, this might not be reasonable as the biases of the other climate state may not remain unchanged and the method's accomplishment relies on the biases caught during the period the method is trained on. We also recognise that recent studies by Maraun et al. (2017) and Maraun and Widmann (2018) have argued against carrying out a cross-validation for evaluating bias corrections. The authors remarked that the observational and simulated data sets do not have a synchronised internal climate variability. Thus, this asynchronism in the internal climate variability may be one of the sources of the biases in free-running models.

Furthermore, as mentioned by Maraun and Widmann (2018), our cross-validation method does not compare the correction to the observations on the validation period (future or past climate state), which can produce false positive or true negative results due to internal variability in the model or observations, but the method assesses whether the statistical evolution of the model is kept.

Moreover, one of the reasons that may explain the remaining difference between the observational and the corrected data sets, as mentioned in the manuscript, can be traced back to the fact that some height classes sample over regions with slightly different biases. Hence, biases of one area could be diminished by the biases that are shared by the other areas. For instance, the strong negative biases observed in the Rhone Valley and Ticino are not fully corrected because the slight underestimation across the Swiss Plateau dominates the bias in this height-class.

Nevertheless, we agree that the evaluation and the argumentation for the remaining biases is not discussed clearly enough in the manuscript and that this should be better explained. Thus, we have extended the discussion more explicitly in the results and conclusion section of next version of the manuscript.

Bennett, James C., Michael R. Grose, Stuart P. Corney, Christopher J. White, Gregory K. Holz, Jack J. Katzfey, David A. Post, and Nathaniel L. Bindoff. 2014. 'Performance of an empirical bias-correction of a high-resolution climate dataset'. International Journal of Climatology 34 (7): 2189–2204. https://doi.org/10.1002/joc.3830.

5. ***It is not clear why the wet-day frequency is adjusted prior to the fitting of the EQM. If EQM is applied to the whole distribution including dry days, this adjustment is included in the EQM fitting. The justification might be linked to the unexplained details in the fitting setup.***

RESPONSE:

We thank the reviewer for highlighting this point and recognize that this adjustment may not be clear enough. We would like to mention that the adjustment does not mainly focus on the wet-day frequency, but the very low intensity values. As clarified already in the answer for reviewer 1, we agree that the argumentation for this adjustment can be better explained. To make this clear, we would like to mention that, in our study, we use an empirical quantile mapping technique (EQM) that differs from the parametric quantile mapping technique (QM). The reason of using an EQM is because this technique uses an empirical cumulative distribution function and does not fit any parametric distribution to the sample, i.e, (sub-) groups, as it is done in the QM. Therefore, we do not assume any known distribution either in our data sets or in the possible application to other climate states. However, the results of the EQM can become unrealistic if the very low intensity values are not adjusted previously. The reason for this is that these values can produce inappropriate TFs due to an important shift in the distribution, i.e., the quantiles (Teutschbein and Seibert, 2012; Lafon et al. 2013).

To adjust these very low values, an additional parameter is included in the definition of days without precipitation that has been mentioned before in the respond of the second major point of reviewer 1. The days without precipitation are not considered for calculating the TFs when they fall below a certain threshold. Many studies use a static threshold for entire data set that is between 0.01 and 1.00 mm day^{-1}, whereas in our study, we calculate a static threshold for each group (or subgroup) and months of the year. This allows to be the consistent with the different biases-treatment across the groups (or subgroups) and months of the year. The threshold is calculated using the local intensity scaling method and can vary vary in our study from 0.001 to 1.00 mm day^{-1}. To clarify this, we have made some changes that are presented in the revised manuscript and also in response to the fifth major comment of reviewer 1.

Changes in the manuscript are presented as follows:

- Page 5 lines 13 – 14

…2010). To correct precipitation with very low-intensity the first part of the local intensity scaling method is used (Schmidli et al., 2006). It consists …

…2010), which can distort the precipitation distribution substantially, i.e., shifting the quantiles, producing inappropriate corrections in the third step when EQM is applied

(Teutschbein and Seibert, 2012). To correct precipitation with very low intensity, an additional parameter is included in the definition of dry days related with the uncorrected precipitation. Dry days are not considered for calculating the TFs when they fall below a certain threshold. Many studies use a static threshold for the entire data set which is between 0.01 and 1.00 mm day^{-1} (Piani et al., 2010a; Lafon et al., 2013; Maraun, 2013). We calculate a static threshold for each group (or subgroup) and months of the year. This allows to be the consistent with the different biases-treatment across the groups (or subgroups) and months of the year. Then, we carry out the local intensity scaling method (Schmidli et al., 2006) that is also used by Teutschbein and Seibert (2012) before using the quantile mapping technique. This method consists …

- Page 5 lines 16 – 17

…The threshold can vary from group to group, but it is often close to or smaller than 1 mm day^{-1} Schmidli et al., 2006).

…In our work, the threshold can vary from group to group and from month to month between 0.001 and 1 mm day-1 as in Schmidli et al. (2006) …

6. ***Although it is mentioned that the errors in the observations should be taken into account when interpreting the results, there is no substantial effort to actually do this. For instance, it would be instructive to do a rough correction for the substantial undercatch of precipitation falling as snow, which strongly affects the high elevations, and assess to what extent the validation results are sensitive to this error.***

RESPONSE:

We appreciate this comment. We agree that we missed to show a wider discussion about the error in the observational data sets when interpreting the results of the correction method. As reviewer 1 asked a similar question we present there the same answer: As mentioned by (Isotta, 2014), the gridded observational data sets do not only present errors due to the interpolation methods, but they also show errors that may differ in quantity from one to the other station (Sevruk, 1985; Richter, 1995) and are related to the "gauge undercatch", whose magnitudes range from 5% over the flatland regions to 30% above 1500 m a.s.l.. Therefore, we have included a better discussion of these errors when analysing the correction, which is presented in the results discussion part of the revised manuscript.

Sevruk B. 1985. Systematischer Niederschlagmessfehler in der Schweiz. Der Niederschlag in der Schweiz, Beitr¨age zur. Geologischen Karte der Schweiz-Hydrologie 31: 65–75.

Richter D. 1995. Ergebnisse methodischer Untersuchungen zur Korrektur des systematischen Messfehlers des Hellmann-Niederschlagsmessers. Bericht Deutschen Wetterdienstes 194, 93 pp. (To be obtained from German Weather Service, Offenbach a.M., Germany.)

7. ***As the realization of internal variability is different the observations and in a free-running GCM (as opposed to a reanalysis) some differences between observations and simulations will be due to internal variability. This effect should be roughly quantified, for***

instance by showing fitting and validating the method for 10 or 15 year sub-periods (which would lead to 9 or 4 possible combinations of fitting and validation subperiods).

RESPONSE:

We thank the reviewer for bringing to our attention the approach to quantify the biases that may be caused by differences between the internal variability of the observational data set and the simulated one. Furthermore, we would like to mention that correction methods are sensitive to the period the methods are trained on, and their accuracies would increase as more information from the observational data sets is taken into account (Lafon et al., 2013). Therefore, since the accuracy of our correction method needs to be kept as high as possible, we have carried out the suggestion made by the reviewer by splitting the data sets into two sub-periods, which is explained and analysed in the following paragraphs.

To quantify any difference that may be caused by using data sets with different internal variabilities, we have calculated two additional sets of Int-TFs using the first and last 15 years, separately. Note that we avoid shorter periods (like the suggested 10 yrs) as less data is available to estimate the TFs. Each set of Int-TFs is then applied to the 30-year simulated precipitation over Switzerland, to be comparable with the 30-yr period used so far in the manuscript. Thus, we obtain two newly corrected precipitation data sets (15yr-A and 15yr-B, respectively) that are compared to the data set that was obtained by the correction trained with the 30-year period (30yr). To assess the difference related to colder and warmer months, we select, as in the manuscript, two months that mainly represent each period; namely, January and July.

Focusing on the biases in the climatological mean value of precipitation intensity, and comparing the original biases with the three approaches, we observe that the methods carried out with 15yr-A and 15yr-B illustrate a correction similar to the method with 30yr. Namely, they reduce the overestimation over high mountain regions during colder months and the general underestimation during warmer months. In addition, the regions with remaining biases agree with the remaining biases of the correction with 30yr. Still, some differences between the 15yr-A and 15yr-B and the method using 30yr are evident: During January, the method using 15yr-A shows a better performance over the high altitudes but not over the flatlands and in the Ticino, and inversely, the method using 15yr-B outperforms the latter areas but not over the mountains (Fig. R7). During July, the method using 15yr-A outperforms over the flatlands and the Ticino but not the high altitudes, and inversely, the method using 15yr-B shows a better performance over the latter area but not over the flatlands and in the Ticino (Fig. R8). This demonstrates that the method calibrated with the two sub-periods can slightly influence the correction method but its effects can be considered minimal when the work by Lafon et al. (2013) is taken into account. As described before, Lafon et al. (2013) found that the accuracy of the correction methods is sensitive to the period the methods are trained on, which could explain some of the remaining biases when using 15yr-A and 15yr-B Therefore, we have mentioned this experiment in the results part of the revised manuscript but its figures are not shown due to the minimal effects.

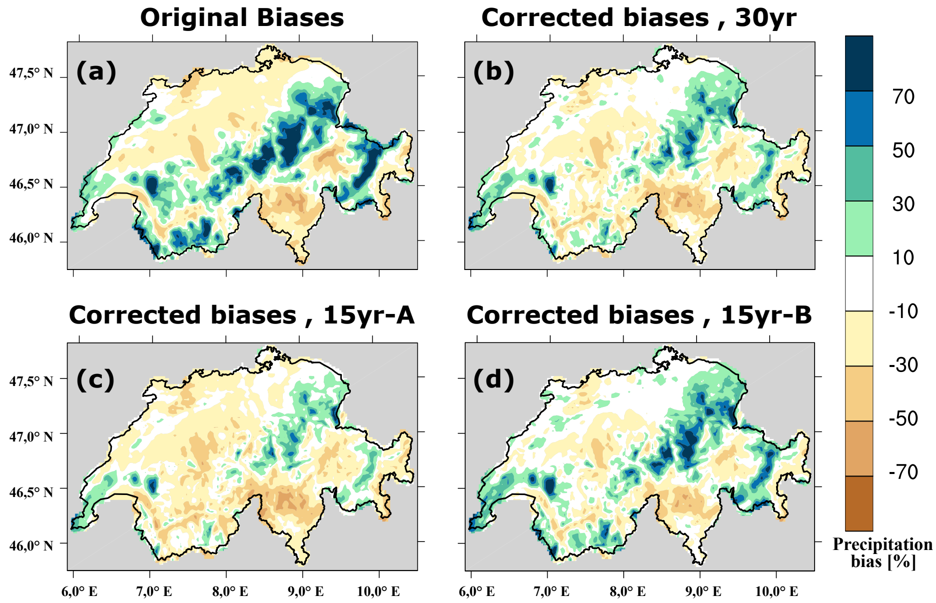


Figure R7. Biases in the climatological mean value of precipitation intensity in January over Switzerland. (a) represents the original biases, (b) the biases after being corrected using Int-TFs obtained from the 30-year period, (c) as in (b) but from the first 15-year period, (d) as (c) but the second 15-year period.

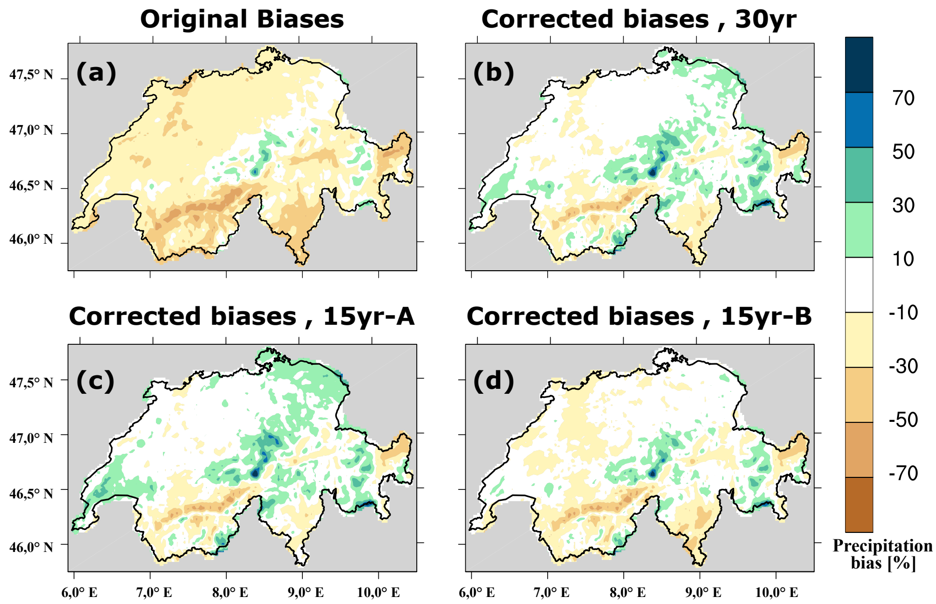


Figure R8. Biases in the climatological mean value of precipitation intensity in July over Switzerland. (a) represents the original biases, (b) the biases after being corrected using Int-TFs obtained from the 30-year period, (c) and (d) as in (b) but Int-TFs obtained from the first and second 15-year period, respectively.

Once again, we would like to thank the reviewer for the time invested to review our paper so carefully and we are looking forward to meeting the reviewers' expectations.

Best regards,

Patricio Velasquez

A new bias-correction method for precipitation over complex terrain suitable for different climate states: a case study using WRF (version 3.8.1)

Patricio Velasquez[1,2], Martina Messmer[1,2,3], and Christoph C. Raible[1,2]

[1]Climate and Environmental Physics Institute, University of Bern, Switzerland
[2]Oeschger Centre for Climate Change Research, University of Bern, Switzerland
[3]School of Earth Sciences, The University of Melbourne, Melbourne, Victoria, Australia

Correspondence to: Patricio Velasquez (velasquez@climate.unibe.ch)

Abstract. This work presents a new bias-correction method for precipitation that considers orographic characteristics, which makes it flexible to be used under highly different climate conditions, e.g., glacial conditions. The new bias-correction and its performance are presented for Switzerland using a regional climate simulation under perpetual 1990 conditions at 2-km resolution driven by a simulation performed with a global climate model. Comparing the regional simulations with observations, we find a strong seasonal and height dependence of the bias in precipitation commonly observed in regional climate modelling over complex terrain. Thus, we suggest a 3-step correction method consisting of (i) a separation into different orographic characteristics, (ii) correction of very low intensity precipitation, and finally (iii) the application of empirical quantile mapping, which is applied to each month separately. Testing different orographic characteristics shows that separating in 400-m height-intervals provides the overall most reasonable correction of the biases in precipitation and additionally at the lowest computational costs. The seasonal precipitation bias induced by the global climate model is fully corrected, whereas some regional biases remain, in particular positive biases in winter over mountains and negative biases in winter and summer in deep valleys and Ticino. The biases over mountains are difficult to judge, as observations over complex terrain are afflicted with uncertainties, which may be more than 30 % above 1500 m a.s.l. A rigorous cross validation, which trains the correction method with independent observations from Germany, Austria and France, exhibits a similar performance compared to just using Switzerland as training and verification region. This illustrates the robustness of the new method. Thus, the new bias-correction provides a flexible tool which is suitable in studies where orography strongly changes, e.g., during glacial times.

1 Introduction

The hydrological cycle ~~has been~~ is an important component in the Earth's climate system, because of its capability to transport and redistribute mass and energy around the world. Changes in the hydrological cycle can on one hand lead to droughts or floods and thus impact the ecosystem services, but it has also been shown that it plays an important role in shaping the Earth's climate history (Mayewski et al., 2004). The latter is because the hydrological cycle shows a strong response to different external forcing functions and to changes in atmospheric compositions (Ganopolski and Calov, 2011; Stocker et al., 2013). Namely, hydrology and water resources are strongly influenced by changes in precipitation patterns (Stocker et al., 2013; Raible et al., 2016). In consequence of this, important modelling tools have been developed, e.g., global atmospheric climate models and hydrological models. These offer valuable information to improve the understanding of the Earth's system responses and feedbacks to internal and external forcings on time scales longer than some centuries (e.g., Xu, 2000; Andréasson et al., 2004; Xu et al., 2005; Fowler et al., 2007a; Yang et al., 2010; Chen et al., 2012).

Still, uncertainties remain, in particular in the hydrological cycle, as not all relevant processes are explicitly simulated by the models (e.g., Ban et al., 2014; Giorgi et al., 2016). This is especially true for global models, which ~~still have a relatively~~ have a comparably coarse spatial resolution. Hence, most processes governing regional- to local-scale precipitation are not resolved ~~yet~~ and need to be parameterised (Leung et al., 2003; Su et al., 2012), resulting in a strong parameter dependence when simulating regional-scale precipitation (Rougier et al., 2009).

~~To avoid~~ To overcome some of the uncertainties, regional climate models (RCMs) are used to ~~dynamically~~ further downscale global climate models ~~. Still, precipitation patterns~~ dynamically. Many RCM simulations are carried out within the framework of the Coordinated Regional Downscaling Experiment (CORDEX), which defines one of the premier goals to better understand relevant phenomena at finer scales (Moss et al., 2010). Even though regional climate models can solve atmospheric equations on a much finer scale than global models, the simulated precipitation patterns still show large biases for present day climate ~~show large biases~~ when comparing them to observations~~, as~~ . This has for example been illustrated by the ~~simulations performed by , e. g., the Coordinated Regional Downscaling Experiment (CORDEX) (e.g., Rajczak and Schär, 2017). The biases are mainly related to the~~ CORDEX simulations analysed by Casanueva et al. (2016); Rajczak and Schär (2017). These biases are not only produced by initial and boundary conditions provided by GCMs, but they are also related to regions characterized by complex topography and to processes that correspond to a finer scale~~and are still insufficiently described due to the model resolution (Boer, 1993; Zhang and McFarlane, 1995; Fu, 1996; Yang et al., 2013)~~, such as cloud microphysical processes. These processes need to be parameterised as they cannot be explicitly resolved because of the RCM resolution used in CORDEX (Boer, 1993; Zhang and McFarlane, 1995; Fu, 1996; Haslinger et al., 2013; Yang et al., 2013; Warrach-Sagi et al., 2013; Mar
. To overcome these shortcomings, RCMs ~~are run to~~ need to be run at a resolution where they can explicitly resolve some of the relevant processes, e.g. convection (e.g., Giorgi et al., 2016; Messmer et al., 2017). Even though the convection-resolving RCMs can describe precipitation much more precisely, biases are still evident (e.g., Ban et al., 2014; Gómez-Navarro et al., 2018). ~~Hence, one important problem in regional climate modelling is that precipitation is simulated more frequently than observed but for most of the RCMs with a weaker intensity (Murphy, 1999; Fowler et al., 2007b; Maraun, 2013). A second problem is that the precipitation is still biased over complex topography by most of the RCMs, even though they are carried out with a higher resolution than the GCMs (Haslinger et al., 2013; Warrach-Sagi et al., 2013; Maraun and Widmann, 2015; Hui et al., 2016).~~ These inconsistencies and uncertainties impact, e.g., the results obtained through hydrological and glacier modelling that follow next in the modelling chain (Allen and Ingram, 2002; Seguinot et al., 2014).

Some climate change studies try to correct parts of these errors in precipitation patterns and ~~amounts~~ intensities by so-called bias-correction methods (Maraun et al., 2010). So far, several correction methods are suggested in the literature, e.g., linear scaling, local intensity scaling, or power transformation (e.g., Berg et al., 2012; Fang et al., 2015; Lafon et al., 2013). Another important bias-correction method is the empirical quantile mapping (EQM) known as one of the best techniques to correct ~~the precipitation biases (e.g., Lafon et al., 2013; Teutschbein and Seibert, 2013; Teng et al., 2015)~~precipitation biases (e.g., Lafon et al., 2013; Teutschbein and Seibert, 2012, 2013; Teng et al., 2015). All these methods ~~have in common that~~ suffer basically from the assumption of stationary biases because they are trained with a climate that does not correspond to the simulated climate that is afterwards bias-corrected. Namely, statistical relationships between observations and model output are used to estimate transfer functions in the observed period and are then applied to different climate states, e.g., past and future climate change scenarios. ~~Besides the strong~~ For additional reviews of bias correction methods see Maraun (2016) and the book by Maraun and Widmann (2018). Besides the assumption of stationarity of the transfer functions, these correction methods ~~do not~~ only implicitly consider orographic features that strongly affect precipitation and its biases (e.g., Piani et al., 2010b; Amengual et al., 2011; Berg et al., 2012; Chen et al., 2013; Cannon et al., 2015; Fang et al., 2015). Hence, the applicability ~~to a different climate state may not be justified for climate states where orography has strongly changed, e.g., during~~ of bias corrections to different climate states such as the Last Glacial Maximum (LGM) ~~where the European Alps were covered with an icecap (Kleman et al., 2013; Ludwig et al., 2019)~~may not be justified because of the before mentioned assumptions and limitations.

This calls for a flexible method ~~, which is able to correct biases also for highly different climatestates~~that can decrease the danger of assuming stationary biases when correcting precipitation errors. One possibility is to apply a cluster analysis to precipitation and its biases to identify classes with similar bias behaviour. An example for Switzerland of such an approach is presented by Gómez-Navarro et al. (2018). The drawback of such an approach for our purpose is that the cluster analysis still relies on the characteristics and circulation of the current climate. Thus, ~~the purpose of this study is to fill this gap and develop a~~ to be as much independent from current climates as possible and to provide a correction that includes important characteristics of the Alpine climate, we came up with "static" characteristics, i.e. topography height and slope orientation and the assumption that relationships to these static characteristics remain unchanged in different climate states. Thus, our work aims at presenting a new bias-correction method ~~for RCMs~~ that fills this gap by using orographic features as variables for the correction. Such a correction avoids the usage of current atmospheric circulation, which makes the technique better applicable to highly different climate states. The new method is based on EQM ~~(Lafon et al., 2013; Teutschbein and Seibert, 2013; Teng et al., 2015) and at the same time includes orographic characteristics.~~ (Lafon et al., 2013; Teutschbein and Seibert, 2012, 2013; Teng et al., 2015) explicitly combined with orographic characteristics, and attempts to correct wet or dry biases that are introduced by parameterisations and numerical formulations in either global or regional models or both. Such biases include especially those that are mainly associated with orographic effects, namely, vertical motion leading to precipitation. The data to be corrected stems from a present day climate simulation performed with the high-resolution RCM Weather Research and Forecasting (WRF) model (Skamarock and Klemp, 2008) that is driven by a simulation under perpetual 1990 conditions using the Community Climate System Model version 4 (CCSM4, Gent et al., 2011). To estimate the transfer functions of the EQM we use two observation data sets, separately; one for Switzerland (MeteoSwiss, 2013) and one for the Alpine region (Isotta et al., 2014). The focus of the presented study is on the method itself and its evaluation over the Alps.

The paper is structured as follows. Section 2 describes the models and data sets used to construct the method. Section 3 presents the new bias-correction method. Section 4 evaluates the new method. Finally, conclusive remarks are given in Sect. 5.

2 ~~Model~~ Models and data

The global climate simulation is performed with the Community Climate System Model (version 4; CCSM4; Gent et al., 2011). The model's atmospheric component is calculated by the Community Atmosphere Model version 4 (CAM4, Neale et al., 2010) and the land component by the Community Land Model version 4 (CLM4, Oleson et al., 2010). We only use these two components and so-called data models are used for the ocean and sea ice, i.e., the atmospheric component is forced by time-varying sea surface temperatures and sea ice cover obtained from a coarser resolved fully coupled 1990 AD simulation with CCSM3 (Hofer et al., 2012a). The atmosphere land-only model was run with a horizontal resolution of 1.25° $\times$ 0.9° (longitude $\times$ latitude) and with 26 vertical hybrid sigma-pressure levels. The global climate simulation covers 31 years using perpetual 1990 AD conditions, i.e., the orbital forcing and atmospheric composition (Table 1). The time resolution of the output is 6-hourly. More detailed information on this simulation and its ~~setting~~ settings are presented in Hofer et al. (2012a, b) and Merz et al. (2013, 2014a, b, 2015).

To investigate the climate over central Europe and in particular over Switzerland in more detail, an RCM is used for the dynamical downscaling. Note that Switzerland is only covered by 12 grid points and the Alps are represented with a maximum height of approximately 1400 m a.s.l. in CCSM4. As RCM, we use the WRF version 3.8.1 (Skamarock and Klemp, 2008). The model is set up with four two-way nested domains with a nest ratio of 1:3. The domains have a horizontal resolution of 56, 18, 6 and 2 km, respectively, and 40 vertical eta levels. The outermost domain includes an extended westward and northward area that takes as midpoint the Alpine region, which allows to capture the influence of the North Atlantic and Scandinavia on the central European and Alpine climate (Fig. 1a). Moreover, the innermost domain focusses on the Alpine region. The fine resolution of 2 km over this area is important as it covers a highly complex terrain. The resolution in the two innermost domains permits the explicit resolution of convective processes, thus, the parameterisation for convection can be switched off in these two domains. Convection permitting model resolutions are ~~preferred as~~ in general preferred as many recent studies show a better performance in simulating precipitation ~~(e.g., Ban et al., 2014; Prein et al., 2015).~~ (e.g., Ban et al., 2014; Prein et al., 2015; Kendon et al., 2017; Berthou . However, we shall keep in mind that some biases in temperature and cloud formation may be produced by this set up, which may lead to additional biases in precipitation as shown in Ban et al. (2014). The relevant parameterisation schemes chosen to run WRF with are listed in Table 2.

WRF is driven by the global simulation and is run for 30 years using perpetual 1990 AD conditions (Table 1). Note that the RCM is not nudged to the global simulation. The ~~simulation is~~ 30-years simulation is split up into ten single 3-years simulations and carried out with ~~adapting~~ adaptive time-step in order to increase the throughput on the available computer facilities. Furthermore, ~~the 30-years simulation is split up into ten single 3-years simulations that have a spin-up of~~ 2-months ~~each~~spin-up time is considered for each 3-years simulation because WRF has only an atmospheric component and its interaction with surface variables, e.g., ice cover and soil, is fully parametrised. Note that the surface variables are provided by the GCM and they are in equilibrium.

Two gridded observational data sets for daily precipitation are used: daily precipitation RhiresD (MeteoSwiss, 2013) and the Alpine Precipitation Grid Dataset (APGD; Isotta et al., 2014). Both data sets cover more than 35 years. In this study, we use only the 30-years period 1979–2008. The RhiresD has a spatial resolution of approximately 2 $\times$ 2 km and covers only Switzerland (MeteoSwiss, 2013). This data set is based on rain gauge measurements distributed across Switzerland ~~(for more details see, Isotta et al., 2014; Güttler et al., 2015)~~(for more details see; Isotta et al., 2014; Güttler et al., 2015). These point measurements are spatially interpolated to obtain a gridded data set, which is described in more detail in Frei and Schär (1998), Shepard (1984) and Schwarb et al. (2001). The APGD encompasses the entire Alpine region with a spatial resolution of 5 $\times$ 5 km (Isotta et al., 2014). For our analysis, the Alpine areas of Italy and Slovenia are excluded because of their poor station density covering the period 1979 – 2008 compared to RhiresD, especially over a complex topography and at high altitudes. It was developed in the framework of EURO4M (European Reanalysis and Observations for Monitoring) by using a distance-angular weighting scheme that integrates climatological precipitation using the local orography and the rain gauge measurements (Isotta et al., 2014). Note that all data sets consider daily precipitation as total precipitation, i.e., both solid and liquid precipitation, and convective and non-convective precipitation. Moreover, days without precipitation are treated as censored values, i.e., not considered in the analysis, when daily precipitation is equal to 0 mm day^{-1}, although in the case of observations this is equivalent to 0.1 mm day^{-1} due to gauge precision.

The observational gridded data sets provide valuable insights~~, in particular in areas where observations are not possible due to extreme weather conditions or insufficient accessibility, such as mountain peaks~~. However, they also contain some discrepancies and uncertainties due to inter- and extrapolation methods, e.g., high precipitation intensities are systematically underestimated and low intensities overestimated~~.~~, especially in areas where observations are not available, i.e. on high elevated areas, such as mountain peaks. The magnitude of these errors depends on the season and the altitude. In regions above 1500 m a.s.l., the error can reach higher values than 30 % because of ~~an undercatch~~a “gauge undercatch” induced by strong winds and the ~~interpolation method~~ applied interpolation method carried out with a distance-angular weighting scheme (Frei and Schär, 1998; Nešpor and Sevruk, 1999; Auer et al., 2001; Ungersböck et al., 2001; Schmidli et al., 2002; Frei et al., 2003; MeteoSwiss, 2013; Isotta et al., 2014). Note that the limitations of the observational data sets are not included in the analysis of this study, i.e., we consider the observational gridded data sets as truth. Nevertheless, one shall keep the limitations of the observational data in mind, in particular when discussing the remaining biases in areas and seasons where the observational data sets also have problems.

~~For the analysis, in particular the comparison between the observational and simulated data, a bilinear interpolation method is used to convert the original grid of WRF into the corresponding one of the observational data sets.~~

3 Bias correction

The correction method, developed in this study, consists of three steps: (i) separation with respect to different orographic characteristics, (ii) adjustment of low-intensity daily precipitation, and (iii) application of the EQM. Each of these three steps ~~are~~ is described in more detail in the following paragraphs.

In a first step, three orographic characteristics are used to separate the region of interest into several groups. These characteristics are height, slope-orientations, and a combination of both. The height ranges from circa 200 m a.s.l. to a maximal value of 3.800 m a.s.l. over the area of interest. Thus, the groups are selected by height-intervals, which cover the range from 400 to 3.200 m a.s.l. Two height intervals are tested separately: 100 or 400 m (e.g., height-intervals of 400 m are shown in Fig. 1c). The heights below 400 and above 3.200 m a.s.l. are considered as two additional height-intervals. The second characteristic, used to group the region of interest, are four slope-orientations: north ($315^\circ \leq$ slope-orientation $< 45^\circ$), east ($45^\circ \leq$ slope-orientation $< 135^\circ$), south ($135^\circ \leq$ slope-orientation $< 225^\circ$) and west ($225^\circ \leq$ slope-orientations $< 315^\circ$). Note that this characteristic is obtained by summing the two slope vectors that are directly provided by the RCM. Combining both characteristics, the groups are selected by height-intervals and then separated into sub-groups by the slope-orientations.

In a second step, we correct the daily simulated precipitation with very low-intensity in each group (or sub-group) and each month of the year, separately. The reason for this is that the frequency of precipitation with very low-intensity is often strongly overestimated due to the drizzle effect produced by the RCM (Murphy, 1999; Fowler et al., 2007b; Maraun et al., 2010)~~.~~, which can distort the precipitation distribution substantially, i.e., shifting the quantiles, producing inappropriate corrections in the third step when EQM is applied (Teutschbein and Seibert, 2012; Lafon et al., 2013). To correct precipitation with very low-intensity~~the first part of the~~, an additional parameter is included in the definition of dry days related with the uncorrected precipitation that is described in the section of model and data before. Dry days are not considered for calculating the TFs when they fall below a certain threshold. Many studies use a static threshold for the entire data set which is between 0.01 and 1.00 mm day^{-1} (Piani et al., 2010a; Lafon et al., 2013; Maraun, 2013). We calculate a static threshold for each group (or subgroup) and months of the year. This allows to be consistent with the different biases-treatment across the groups (or subgroups) and months of the year. Then, we carry out the local intensity scaling method ~~is used (Schmidli et al., 2006). It consists of deleting precipitation values that~~ (Schmidli et al., 2006) that is also used by Teutschbein and Seibert (2012) before using the quantile mapping technique. This method consists of censoring precipitation values by setting them zero when they are below a specific threshold ~~. This threshold is~~ determined from the daily simulated precipitation such that the threshold exceedance coincides with the precipitation-day occurrence from the observations. ~~The~~ In our work, the threshold can vary from group to group ~~, but it is often close to or smaller than~~ and from month to month between 0.001 and 1 mm day^{-1}~~(Schmidli et al., 2006)~~, similar to Schmidli et al. (2006).

In a third step, we correct the daily precipitation rate using an EQM method (Themessl et al., 2011; Lafon et al., 2013; Fang et al., 2015; Teng et al., 2015). EQM is based on the assumption that all probability distribution functions are unknown, i.e. non-parametric (Wilks, 2011). The method consists of adjusting the quantile values from a simulation (Q-SIM) with those from observations (Q-OBS) through a transfer function (TF; Fig. 2). The method is implemented by splitting each cumulative distribution function, i.e., observed and modelled, into 100 discrete quantiles. For each quantile value, the adjustment is carried out with a linear correction, where Q-SIM is transformed into Q-SIM* (corrected quantile) so that Q-SIM* = TF × Q-SIM and TF = Q-OBS / Q-SIM (Lafon et al., 2013). This linear correction is akin to the ~~‘~~factor of change ~~’ or ‘delta change ’~~ or delta change used in Hay et al. (2000). For values that are between quantiles, the same linear correction is used, but the TF is approximated by using a linear interpolation between the TFs related to the two nearest quantiles. In cases where values are below (above) the first (last) quantile, the TF related to the first (last) quantile is used for the adjustment. Similar methods were successfully applied to correct biases in precipitation simulated by RCMs (e.g., Sun et al., 2011; Themessl et al., 2012; Rajczak et al., 2016; Gómez-Navarro et al., 2018).

To combine all steps, the ~~EQM is~~ local intensity scaling method and the EQM are applied to each (sub-) group ~~and~~ defined in the first step and to each month of the year, separately, by pooling all grid points that belong to each group and handling them as a single distribution of daily precipitation. This results in a set of TFs for each (sub-) group and each month of the year. For instance, when the correction is carried out using height-classes of 400 m, a TF is defined for each height group, resulting in nine TFs for each month and in total 108 TFs throughout the year. Moreover, the correction is afterwards applied to the daily precipitation at every grid point using the TFs that are common to all elements within the same group (or sub-group) and month. Thus, the new correction method guarantees that seasonality and height are taken into account making the method flexible for climate states with a changed orography, e.g., the LGM.

To come up with a final method for the Alpine region~~we first test~~, we first evaluate the influence of the different orographic characteristics (step 1). To be consistent with former studies (e.g., Sun et al., 2011; Themessl et al., 2012; Wilcke et al., 2013; Rajczak et al., 2016), the evaluation ~~of the new method first~~ uses the same region where the TFs are estimated. ~~To be more rigorous~~Explicitly, this means that the Swiss region in the WRF output (at 2 km resolution) is defined as the area to be corrected and the RhiresD data set (at 2 km resolution) is used to obtain the TFs and to evaluate the different correction methods. These TFs are called Internal TFs (Int-TF) during the cross-validation process later on. Once the final method is determined, we additionally apply a cross-validation to test the method more rigorously: Thereby, Switzerland is defined as the area to be corrected ~~; then, we calculate two different TFs; namely, from the same Swiss region called Internal TFs (~~(WRF output at 2 km resolution); in addition to the Int-TF ~~), and~~ (see above), which uses the same region to define TFs and to apply the correction, we also calculate a second set of TFs. The second set of TFs is obtained from the corresponding Alpine region of Germany, France, and Austria altogether called External TFs (Ext-TF) ~~(Fig. 1c~~using the APGD data set (at 5 km resolution; Fig. 1c). Note that Ext-TFs are carried out at 5 km horizontal resolution and applied to Switzerland at 2 km resolution. To demonstrate the improvement of using the new method, we further compare it to a ~~commonly used~~ simple method that is carried out without orographic features and uses TFs deduced for the entire region of Switzerland (at 2 km ~~) (similar to Berg et al., 2012; Maraun, 2013; Fang et al., 2015).~~ resolution, 12 TFs in total). Note that our approach mainly focuses on correcting biases caused by parameterisations and systematic errors related to the topography.

4 Results

4.1 Evaluation of WRF: Seasonality and bias

To obtain insights into the performance of the RCM over complex topography, we compare the spatial and temporal representation of the simulated precipitation (the raw model output) with the RhiresD data. Focusing on monthly ~~precipitation~~ mean precipitation intensity across Switzerland, the box plots illustrate biases in the climatological annual mean cycle (Fig. 3a). ~~Mean~~ The climatological mean values are slightly overestimated during colder months, i.e., between November and March, and are underestimated during warmer months, i.e., between April and October, but especially in September. In addition to the climatological mean values, Fig. 3a also shows the distributions of ~~daily precipitation~~ monthly mean precipitation intensity and their interquartile ranges. In colder months, the simulated distributions ~~of daily precipitation~~ are wider and shifted to higher values than the observed ~~distribution~~distributions, whereas during warmer months a clear shift to less precipitation is found compared to the observed ones. Overall the interquartile ranges are reasonably simulated, which means that WRF realistically represents the variability of ~~daily precipitation~~ monthly mean precipitation intensity. Extreme precipitation, however, is strongly underestimated.

~~To~~ The annual cycle and the distributions of monthly mean precipitation intensity are estimated for different height-classes to get additional understanding of the behaviour of the simulated precipitation ~~, the annual cycle and the monthly distributions of daily precipitation are estimated for different height-classes. Figure~~ and also to explicitly illustrate the relation of the precipitation biases to the topography. This is summarised in Figs. 3b and 3c ~~show the boxplots~~ for the height class 400–800 m and 2800–3200 m ~~, respectively, to illustrate the precipitation bias and its relation to the topography of Switzerland~~that mostly represent the low and high altitudes, respectively. The climatological monthly means of the colder months, i.e., from November to March, are generally underestimated in the lower height-classes ~~,~~ but overestimated at high altitudes. Hence, we identify a positive correlation between the main biases and the topography during these colder months. In the warm months, i.e., April to October, the height-classes 400–800 m and 2800–3200 m both reveal an underestimation in the climatological monthly means compared to the observations. Therefore, the simulated annual cycle changes from a weak cycle at low altitudes, in agreement with the one of the observations, to a strong and inverse seasonal cycle at high altitudes (Fig. 3b and 3c). An inverse annual cycle is also identified by Gómez-Navarro et al. (2018), where they carried out WRF simulations using a similar global climate model ~~as~~ for initial and boundary conditions as used in this study. These authors found that the inversed annual cycle in precipitation is caused by the driving global climate model. Furthermore, we observe positive biases in the interquartile ranges during colder months, and a slight underestimation during warmer months (Fig. 3b and 3c). Thus, the splitting into different height-classes demonstrates to be appropriate for being used in the bias correction.

To better describe the spatial biases related to colder and warmer months, we select two months that mainly represent each period; namely, January and July. For these example months, we present the ~~patterns of biases~~ spatial patterns of the biases in the monthly mean precipitation intensity, in ~~precipitation, changes in the distribution of daily precipitation,~~ the variability illustrated by the interquartile range~~as well as biases in~~, and in the wet-day frequency. Note that ~~we consider~~ the observational data sets are considered generally reliable and represent orographic features well, although at high altitudes less observations are available (Isotta et al., 2014). Furthermore, these spatial patterns implicitly illustrate the relation between the precipitation biases and the topography considering an uncertainty of around 30 % acceptable in the simulated precipitation due to the uncertainty in the observational data sets (Sect. 2).

The biases in the climatological mean precipitation intensity at each grid point (Fig. 4a and 4d) confirms the strong height dependence and seasonality already shown in Fig. 3, which demonstrates that the splitting into different height-classes is appropriate to be used in the bias correction. The strongest positive biases are mainly observed over mountains and during colder months, whereas the Swiss Plateau seems to be reasonably well simulated (Fig. 4a). Note that also the observations tend to underestimate precipitation in mountain regions so that a part of the strong positive bias is related to observational uncertainties (Isotta et al., 2014). In warmer months, the strongest negative biases are found in the north-western part of Switzerland, Ticino and in the steep valleys, where the Rhone Valley is marked by the strongest biases~~, whereas in~~ . In high mountain regions smaller positive biases are identified during warmer months than during colder months (Fig. 4d). The strongest biases over mountains and in steep valleys seem to be induced by an amplification of different observed precipitation climatologies that govern those areas; namely, the mountains are known as wet regions and the steep valleys as dry areas ~~(for more details see, Frei and Schär, 1998; Schwarb et al., 2001)~~(for more details see; Frei and Schär, 1998; Schwarb et al., 2001) . This gives a first hint that different processes may lead to the biases. The positive precipitation bias over mountains in colder months may be mainly related to wet bias of the global simulation and synoptic transport, which is also overestimated in the global simulation (Hofer et al., 2012a, b). ~~Note also that the observations have the strongest measurement errors over the mountains, i.e., they tend to underestimate precipitation.~~ The resolution of the RCM seems to be important as this affects the representation of steep valleys, especially during convective processes in warmer months. The same is also true for colder months, but to a lesser extent, as convective processes only play a minor role in these months.

The biases in the interquartile range of the ~~daily precipitation distribution~~ distribution of monthly mean precipitation intensity at each grid point (Fig. 5a and 5d) are strongly overestimated to a large extent over the Alps during colder months, whereas during warmer months the interquartile range is generally smaller compared to the observations. The biases are stronger than the ones observed in the climatological mean value (Fig. 4a and 4d), which means that the variability simulated by WRF is strongly season-dependent (Fig. 5a and 5d). The simulated increase in variability during colder months is a hint that processes common during winter, e.g., the overestimated synoptic atmospheric systems in the global simulation, may be too efficient in producing precipitation compared to the observations. The reduced variability in the warmer months hints to remaining problems in convective processes as these are more relevant during summer. Also observations do not perfectly estimate the range due to their incertainty whose magnitudes range from 5% over the flatland regions to more than 30% in high altitudes (Isotta et al., 2014).

Another important measure to ~~characterize~~ characterise precipitation is the occurrence of precipitation at each grid point, defined by the wet-day frequency (the number of days with precipitation rate of at least 1 mm day^{-1}). The wet-day frequency is strongly overestimated during colder months, but shows only a slight overestimation during warmer months (Fig. 6a and 6d). This overestimation can be also related to the well-known problem in regional climate modelling, which is defined as a higher frequency in precipitation but at the same time with a lower intensity than observed (Murphy, 1999; Fowler et al., 2007b; Maraun, 2013)

. The overestimation in wet-day frequency, so-called drizzle effect, can be mainly related to the occurrence of synoptic atmospheric systems commonly observed during colder months and not to local convective processes that are frequently observed during summer (for climatology see Frei and Schär, 1998; Isotta et al., 2014). Furthermore, the positive bias in the wet-day frequency may ~~also explain~~ slightly contribute to the underestimation of the extreme precipitation (Fig. 3) as ~~moisture~~ precipitable water necessary for extreme precipitation events is removed via the drizzle effect. Namely, the precipitable water available for a daily extreme precipitation event is distributed over several days due to problems in the parameterisations of the cloud microphysical and precipitation processes as found in Knist et al. (2018).

4.2 Influence of different orographic characteristics on the performance of the bias-correction method

Different orographic characteristics are suggested to be used as classification in the new bias-correction method (step 1 in Sect. 3): the height-intervals (100 m and 400 m), the slope-orientations, and a combination of both using the height interval of 400 m (combined-features). Note that the results are not affected by interchanges in the order of the orographic characteristics in the combined-features (therefore not shown). We assess in the following, which of these characteristics are necessary to improve ~~the~~ a simple approach of applying one EQM to the entire domain, ~~often used in studies for present day and future climate change (e.g., Evans et al., 2017; Li et al., 2017; Ivanov et al., 2018)~~where orographic features are not considered. Therefore, we use Taylor diagrams (Fig. 7) for four months namely January, April, July, and September, as the biases show a strong seasonality (see previous section). The evaluation is carried out with three statistics: the spatial correlation, the spatial root-mean-square-error and the spatial standard deviation.

Figure 7a shows that the correction methods using height-intervals of both, 100 and 400 m, and the combined-features have a better performance during the colder months than the other methods, using just orientation or one EQM for the entire domain: the standard deviation is better adjusted, especially by using height-intervals of 100 m, the root-mean-square-error is reduced by roughly 32 %, and the correlation is slightly increased (Fig. 7b). During the cold-to-warm transition months (here illustrated by April), the correction using height-intervals of 400 m and the combined-features have a better performance than the other settings. This is because the standard deviation is fully adjusted, the root-mean-square-error is reduced by 17 %, and the correlation is increased to $r = 0.75$ (Fig. 7b). During the warmer months, all correction methods except the one using height-intervals of 100 m show a similar good performance, i.e., the standard deviation is fully adjusted, the root-mean-square-error is slightly reduced, and the correlation is slightly increased (Fig. 7c). During the warm-to-cold transition months (September, Fig. 7d) all correction methods show a similar performance increase compared to the observations, correlation and root-mean-square-error are only slightly improved. The method using height-intervals of 100 m often reduces the standard deviation, which ~~may~~ can be explained by a ~~weak data coverage in~~ reduced data coverage and thus less variability within some height classes.

Even though, all the settings mostly show a good performance, the one using height-intervals of 400 m outperforms in most measures and months. In addition, the correction method using the height-intervals of 400 m needs less computational time compared to the similarly good correction method using height-intervals of 400 m and slope-orientations. Therefore, the method using height-intervals of 400 m seems to be the most appropriate and is used in the following analysis.

4.3 Application of the bias-correction method and cross-validation

The bias-correction method using height-intervals of 400 m is now assessed in more details. First, we focus on results where the TFs in the method are estimated in the domain of Switzerland (Int-TFs) and then results obtained by the cross-validation are discussed, i.e., estimating the TFs with the surrounding Alpine region, excluding Switzerland (Ext-TFs).

To illustrate the improvement by the correction method using Int-TFs, we compare the spatial and temporal representation of the corrected precipitation with the RhiresD data set. Focusing on the monthly ~~precipitation~~ mean precipitation intensity across Switzerland, we find that the climatological annual cycle of mean precipitation intensity fully coincides with the one of the observations (Fig. 3a). Also, the ~~monthly distributions of daily precipitation~~ distributions of monthly mean precipitation intensity are fully adjusted and the corresponding interquartile ranges mainly correspond to the ones of the observations when using the new bias-correction method. Still, the extreme precipitation events are underestimated with the new method, which is expected as the TF of the extreme values is poorly constrained in the EQM approach (e.g., Themessl et al., 2011). The segregation into the height-classes (Fig. 3b and 3c) ~~show~~ shows that the climatological monthly means and the ~~monthly distributions of daily precipitation~~ distributions of monthly mean precipitation intensity are also well adjusted compared to the observations. This illustrates that the bias-correction method using height-intervals of 400 m works.

To further describe the spatial improvements of the new bias-correction method, we select here, as in the Sect. 4.1, two months that mainly represent the colder and warmer months, e.g., January and July. We again focus on biases in ~~precipitation, changes in the distribution of daily precipitation,~~ the monthly mean precipitation intensity, in the variability illustrated by the interquartile range~~as well as biases in~~, and in the wet-day frequency.

~~Comparing~~ A comparison between Fig. 4a and 4d with Fig. 4b and 4e, shows that the ~~mean precipitation biases~~ biases in climatological mean precipitation intensity are substantially reduced, especially the overestimation over high mountain regions during colder months and the general underestimation during warmer months. Still, regions with positive and negative biases remain over the eastern part of the mountains in colder months and in the steep valleys like the Rhone Valley in warmer months. Also, the negative bias in the Ticino during colder months remains, albeit it is slightly ameliorated. The rather moderate performance in these regions can be traced back to the fact that some height classes sample over regions with different biases. Hence, biases of one area are ~~strongly~~ diminished by the biases that are shared by the other areas. For instance, the strong negative biases observed in the Rhone Valley and Ticino are not fully decreased because the slight underestimation from the Swiss Plateau dominates this height-class (Fig. 4b and 4e).

To assess the improvements with respect to precipitation variability, we focus on the interquartile range of the ~~daily precipitation distribution~~ distribution of monthly mean precipitation intensity at each grid point (Fig. 5b and 5e compared to Fig. 5a and 5d). The biases of the interquartile range improve only moderately, i.e., the strong overestimation over the mountains is partly corrected during colder months but not during warmer months. The underestimation over the flatlands and steep valleys is corrected during warmer months and poorly during colder months.

For the wet-day frequency, we find that the positive biases are mostly reduced, especially the strong overestimation over the mountains during colder months (Fig. 6b and 6e). However, the regions of Rhone Valley and Ticino, which show no biases in the raw model output, are slightly underestimated during colder months. The negative biases observed in the region of Grisons become stronger during colder months and in the region of Rhone Valley during warmer months (Fig. 6b and 6e). This effect is again caused by sampling different regions with different biases in the height classes.

Recent studies by Maraun et al. (2017) and Maraun and Widmann (2018) remark that the observational and simulated data sets do not have a synchronised internal climate variability and, thus, this may be one of the sources of the remaining biases in free-running model. To assess these remaining biases, two additional tests are carried out with different sets of Int-TFs that are calculated from the first and last 15 years of the 30-yr period, separately. Note that the accuracy of the correction method is sensitive to the length of the calibration period (Lafon et al., 2013). The two tests are compared to the correction method that is trained on the entire 30-year period. The tests perform similar to the correction method using 30 years, which demonstrates that the calibration length has only a weak and negligible effect on the resulting corrected precipitation data set (therefore not shown).

To check the robustness of the new bias-correction method, a cross-validation is performed. As noted by Bennett et al. (2014), the importance of cross-validation is to test the transferability of a bias-correction method to a different climate state. Thereby, the TFs are estimated from an independent data set of the Alpine region (the APGD in coarser resolution of 5 km) excluding Switzerland (Ext-TFs) and then these TFs are applied to the Swiss region ~~.~~ (at 2 km resolution) to directly compare the results with the ones obtained from the application of Int-TFs (at 2 km resolution) and to avoid any additional uncertainty produced by interpolation. Additionally, we also evaluate the performance of the correction when using Ext-TFs trained at 5 km and then applied to the Swiss region at 5 km resolution, which shows minimal differences on the results (therefore not shown). To have insights into the effects of the correction method using Ext-TFs, we compare the spatial and temporal representation of the corrected precipitation with the results obtained by the Int-TFs. Note that ~~for the bias calculation always the~~ the RhiresD data set is always used as observations for the bias calculation. Again, to describe the spatial effects, we select here two months that mainly represent the colder and warmer months, i.e., January and July.

~~Comparing~~ A comparison between Fig. 4c with 4b shows almost the same pattern, i.e., the improvement in mean precipitation achieved by using Ext-TFs is similar to the Int-TFs during colder months. Still, some positive biases over the mountains seem to be smaller using Ext-TFs than Int-TFs, whereas the remaining negative biases are slightly stronger than the ones after using Int-TFs (Fig. 4b and 4c). The reason for the latter could lie in the coarser resolution of APGD data set used to estimate the Ext-TFs or the inclusion of larger regions in the north and west of the Alps mixing different climate conditions and thus bias behaviours. The slightly better performance in the mountain regions is probably due to the fact that for these height classes more data are available, i.e., more grid-points at high altitudes (Fig. 1c), and thus a better constraint of the TFs is possible. In the warmer months, we find that the method using Ext-TFs ~~show~~ shows slightly more negative biases than with Int-TFs, in particular over the Swiss plateau. Again, we hypothesise that the inclusion of larger regions in the north and west of the Alps is responsible for this bias behaviour.

The interquartile ranges of the ~~monthly distribution of daily precipitation~~ distribution of monthly mean precipitation intensity are similar when using either Ext-TFs or Int-TFs for the colder months (Fig. 5c compared to 5b). During warmer months, the negative biases in the western part of Switzerland are less improved using Ext-TFs than Int-TFs, again a hint that the inclusion of larger regions in the north and west of the Alps in the lower height classes plays a role in the bias of the interquartile range.

The wet-day frequencies are very similarly corrected as in the approach using Ext-TFs compared to Int-TFs (Fig. 6c and 6f compared to Fig. 6b and 6e). Thus, the wet-day frequency seems to be insensitive to the region where the TFs are estimated from.

In summary, the new correction method reasonably well corrects biases in ~~different precipitation variables~~the monthly mean precipitation intensity, in the variability illustrated by the interquartile range, and in the wet-day frequency. The cross-validation shows that using different observational data sources from independent regions have only a minor effect on the improvement obtained by the method and thus demonstrates its robustness.

5 Conclusions

In this study, we present a new bias-correction method for precipitation over complex topography, which takes orographic characteristics into account. To illustrate the performance of the new method, a simulation under perpetual 1990 AD conditions is carried out with the regional climate model WRF at 2-km resolution over Switzerland. This simulation is driven by the general circulation model CCSM4.

The comparison between the dynamically downscaled simulation and the observations over Switzerland shows that the biases are ~~seasonal~~ season dependent and strongly related to the complexity of the topography. Colder months (November to March) exhibit positive biases over mountains and negative biases in steep valleys, whereas during the warmer months (April to October) negative biases dominate, especially in the Rhone Valley and Ticino. Parts of the biases are introduced by the global climate model, in particular the seasonal biases as shown by Gómez-Navarro et al. (2018). Moreover, the large scale atmospheric circulation of the global climate model is too zonal – a known problem in many models (e.g., Raible et al., 2005, 2014; Hofer et al., 2012a, b; Mitchell et al., 2017) – which cannot be fully compensated by the regional climate model. Thus, the wet bias present in the global simulation (Hofer et al., 2012a, b) may be transported into the regional model domain rendering especially the colder months with more precipitation. Still, observations are also not perfect and underestimate precipitation in particular in high altitudes by up to 30% (Isotta et al., 2014). Other biases are potentially induced by the regional climate model, e.g., a WRF simulation using a similar setting but driven by ERA-Interim (Gómez-Navarro et al., 2018) shows also a similar overestimation of precipitation over mountain regions as the simulation used in this study. In addition, we find that the extreme precipitation values are underestimated. This is due to the drizzle effect (Murphy, 1999; Fowler et al., 2007b) that can remove moisture needed for the extreme precipitation, which mainly comes from physical parameterisations of the model itself (Solman et al., 2008; Menéndez et al., 2010; Gianotti et al., 2011; Carril et al., 2012; Jerez et al., 2013). A hint for this is given by the fact that the wet-day frequency in the simulation is enhanced compared to the observations.

Although numerous approaches to correct biases exist (e.g, Maraun, 2013; Teng et al., 2015; Casanueva et al., 2016; Ivanov et al., 2018), a new method is needed, which can decrease the danger of assuming stationarity biases and is flexible enough to be applicable to different climate states like glacial times which are ~~characterized~~ characterised by a strongly changed topography.

The new method consists of three steps: the orographic characteristics differentiation, the adjustment of very low precipitation intensity~~adjustment~~, and the EQM. Different orographic characteristics, i.e., the height-intervals, the slope-orientations, and the combination of both, are tested showing that the method using height-intervals of 400 m is generally the most skilful correction compared to other orographic characteristics and at the same time is computationally the most efficient one. Clearly, the new method outperforms the ~~standard~~ simple method of applying one EQM transfer function that is deduced for the entire region of interest ~~, which is commonly used (Berg et al., 2012; Maraun, 2013; Fang et al., 2015)~~and does not consider any orographic features.

Applying the new bias-correction method to the Swiss region exclusively shows that the biases are mostly corrected. In particular, the distribution of the monthly precipitation across Switzerland is mainly adjusted, the mean precipitation biases are substantially reduced, and the biases in the wet-day frequency are mostly reduced. The method better corrects the positive biases during colder than warmer months, and reversely, the negative biases during warmer than colder months. However, some biases are still observed, which is explained by the fact that some height classes sample over regions with different biases and that the deficient constraint of the TFs in uttermost quantiles poorly corrects extreme values, i.e., below the first quantile and above the last quantile. Furthermore, part of the remaining biases may also be interpreted as possible error propagation, which initially comes from the interpolation methods and “gauge undercatch” in the gridded observational data sets, especially at high altitudes where less data is available (for more details see; Sevruk, 1985; Richter, 1995; Isotta et al., 2014).

The cross-validation presented in this work might not be reasonable as the biases of the other climate state may not remain unchanged and the method’s accomplishment relies on the biases caught during the period the method is trained on. In addition, Maraun et al. (2017) and Maraun and Widmann (2018) have argued against carrying out a cross-validation for evaluating bias corrections due to the asynchronism in the internal climate variability of the data sets. Maraun and Widmann (2018) argued that cross-validation methods shall compare the correction with the observations on different climate states, i.e., the future or past climate state, otherwise they can produce false positive or true negative results. To overcome these possible limitations, we first check the transferability of the bias-correction method to a different climate state by selecting an independent data set of the Alpine region (APGD) excluding Switzerland. The cross-validation using independent data to estimate the transfer functions (Ext-TFs) shows a similar improvement as the correction performed with data over the Swiss region exclusively (Int-TFs). Even though, the positive biases are slightly better corrected compared to using the Int-TFs, the remaining negative biases are slightly stronger than using the Int-TFs. We find that the inclusion of larger mountainous regions in the east and west of the Swiss Alps may be responsible for the improvement in positive bias-correction. The less efficient correction of the negative biases is related to the inclusion of larger areas of grid points in lower height classes in the north and west of Switzerland mixing different climate conditions and bias behaviours. Moreover, we evaluate the influence of the different internal variabilities on the correction performance by using different periods to calibrate the TFs. The evaluation performs similar to the correction method using 30 years, which demonstrates that the calibration length has only a weak and negligible effect on the resulting corrected precipitation data set. Thus, the cross-validation shows that the new bias-correction method is less dependent on different internal variability and on the region which the ~~region, which is used for fitting the TFs, than other~~

~~methods commonly used~~ method is trained on than commonly used methods (e.g., Berg et al., 2012; Maraun, 2013; Fang et al., 2015). This demonstrates the robustness of the new method.

Still, some of the limitations could be improved in a future work by using additional features; e.g. a two-dimensional concavity index that ~~can not~~ cannot only describe the form and orientation of the valleys, but also distinguish the flatlands from the valleys that are located in the middle of the Alps. Besides, one of the next steps will be the application of this new method to other climate states that have a different complex topography, e.g., the LGM. Glaciologists can benefit from a better accuracy of precipitation data that is used as input data by their models, whose results may provide an alternative method for the cross-validation when evaluating the prediction and proxy data of the glacier extents.

Code and data availability. WRF is a community model that can be downloaded from its web page (http://www2.mmm.ucar.edu/wrf/users/code_admin.php). The two climate simulations (global: CCSM4 and regional: WRF) occupy several terabytes and thus are not freely available. Nevertheless, they can be accessed upon request to the contributing authors. The post-processed daily precipitation that is used to perform the bias-correction is archived on Zenodo (Velasquez et al., 2019). The RhiresD and APGD data set can be requested from MeteoSwiss. Simple calculations carried out at a grid point level are performed with Climate Data Operator (CDO, Schulzweida, 2019) and NCAR Command Language (NCL, UCAR/NCAR/CISL/TDD, 2019). The figures are performed with NCL (UCAR/NCAR/CISL/TDD, 2019) and RStudio (RStudio Team, 2015). The codes to perform the bias-correction, the simple calculations and the figures are archived on Zenodo (Velasquez et al., 2019).

Author contributions. PV, MM, and CCR contributed to the design of the experiments. PV carried out the simulations and wrote the first draft. All authors contributed to the internal review of the text previous to the submission.

Competing interests. The authors declare no competing interests.

Acknowledgements. This work is supported by the Swiss National Science Foundation (grant 200021_162444). MM acknowledges support by the SNF (Early Postdoc.Mobility). The CCSM4 and WRF simulations were performed on the supercomputing architecture of the Swiss National Supercomputing Centre (CSCS). Thanks are due to European Reanalysis and Observations for Monitoring for providing the APGD data set.

References

Allen, M. R. and Ingram, W. J.: Constraints on future changes in climate and the hydrologic cycle, https://doi.org/10.1038/nature01092, 2002.

Amengual, A., Homar, V., Romero, R., Alonso, S., and Ramis, C.: A statistical adjustment of regional climate model outputs to local scales: Application to Platja de Palma, Spain, Journal of Climate, 25, 939–957, https://doi.org/10.1175/JCLI-D-10-05024.1, 2011.

Andréasson, J., Bergström, S., Carlsson, B., Graham, L. P., and Lindström, G.: Hydrological change – climate change impact simulations for Sweden, Journal of the Human Environment, 33, 228–234, https://doi.org/10.1579/0044-7447-33.4.228, 2004.

Auer, I., Böhm, R., and Schöner, W.: Austrian long-term climate 1767–2000, Osterreichische Beiträge zu Meteorologie und Geophysik, 25, 147, 2001.

Ban, N., Schmidli, J., and Schär, C.: Evaluation of the convection-resolving regional climate modeling approach in decade-long simulations, Journal of Geophysical Research: Atmospheres, 119, 7889–7907, https://doi.org/10.1002/2014JD021478, 2014.

Bennett, J. C., Grose, M. R., Corney, S. P., White, C. J., Holz, G. K., Katzfey, J. J., Post, D. A., and Bindoff, N. L.: Performance of an empirical bias-correction of a high-resolution climate dataset, International Journal of Climatology, 34, 2189–2204, https://doi.org/10.1002/joc.3830, 2014.

Berg, P., Feldmann, H., and Panitz, H. J.: Bias correction of high resolution regional climate model data, Journal of Hydrology, 448-449, 80–92, https://doi.org/10.1016/j.jhydrol.2012.04.026, 2012.

Berthou, S., Kendon, E. J., Chan, S. C., Ban, N., Leutwyler, D., Schär, C., and Fosser, G.: Pan-European climate at convection-permitting scale: a model intercomparison study, Climate Dynamics, https://doi.org/10.1007/s00382-018-4114-6, 2018.

Boer, G. J.: Climate change and the regulation of the surface moisture and energy budgets, Climate Dynamics, 8, 225–239, https://doi.org/10.1007/BF00198617, 1993.

Cannon, A. J., Sobie, S. R., and Murdock, T. Q.: Bias correction of GCM precipitation by quantile mapping: How well do methods preserve changes in quantiles and extremes?, Journal of Climate, 28, 6938–6959, https://doi.org/10.1175/JCLI-D-14-00754.1, 2015.

Carril, A. F., Menéndez, C. G., Remedio, A. R. C., Robledo, F., Sörensson, A., Tencer, B., Boulanger, J.-P., de Castro, M., Jacob, D., Le Treut, H., Li, L. Z. X., Penalba, O., Pfeifer, S., Rusticucci, M., Salio, P., Samuelsson, P., Sanchez, E., and Zaninelli, P.: Performance of a multi-RCM ensemble for south eastern South America, Climate Dynamics, 39, 2747–2768, https://doi.org/10.1007/s00382-012-1573-z, 2012.

Casanueva, A., Kotlarski, S., Herrera, S., Fernández, J., Gutiérrez, J. M., Boberg, F., Colette, A., Christensen, O. B., Goergen, K., Jacob, D., Keuler, K., Nikulin, G., Teichmann, C., and Vautard, R.: Daily precipitation statistics in a EURO-CORDEX RCM ensemble: Added value of raw and bias-corrected high-resolution simulations, Climate Dynamics, 47, 719–737, https://doi.org/10.1007/s00382-015-2865-x, 2016.

Chen, H., Xu, C.-Y., and Guo, S.: Comparison and evaluation of multiple GCMs, statistical downscaling and hydrological models in the study of climate change impacts on runoff, Journal of Hydrology, 434-435, 36–45, https://doi.org/10.1016/j.jhydrol.2012.02.040, 2012.

Chen, J., Brissette, F. P., Chaumont, D., and Braun, M.: Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America, Water Resources Research, 49, 4187–4205, https://doi.org/10.1002/wrcr.20331, 2013.

Evans, J. P., Argueso, D., Olson, R., and Di Luca, A.: Bias-corrected regional climate projections of extreme rainfall in south-east Australia, Theoretical and Applied Climatology, 130, 1085–1098, https://doi.org/10.1007/s00704-016-1949-9, 2017.

Fang, G., Yang, J., Chen, Y. N., and Zammit, C.: Comparing bias correction methods in downscaling meteorological variables for a hydrologic impact study in an arid area in China, Hydrology and Earth System Sciences, 19, 2547–2559, https://doi.org/10.5194/hess-19-2547-2015, 2015.

Finney, D. L., Marsham, J. H., Jackson, L. S., Kendon, E. J., Rowell, D. P., Boorman, P. M., Keane, R. J., Stratton, R. A., and Senior, C. A.: Implications of improved representation of convection for the East Africa water budget using a convection-permitting model, Journal of Climate, 32, 2109–2129, https://doi.org/10.1175/JCLI-D-18-0387.1, 2019.

Fowler, H. J., Blenkinsop, S., and Tebaldi, C.: Linking climate change modelling to impacts studies: Recent advances in downscaling techniques for hydrological modelling, International Journal of Climatology, 27, 1547–1578, https://doi.org/10.1002/joc.1556, 2007a.

Fowler, H. J., Ekström, M., Blenkinsop, S., and Smith, A. P.: Estimating change in extreme European precipitation using a multimodel ensemble, Journal of Geophysical Research: Atmospheres, 112, D18 104, https://doi.org/10.1029/2007JD008619, 2007b.

Frei, C. and Schär, C.: A precipitation climatology of the Alps from high-resolution rain-gauge observations, International Journal of Climatology, 18, 873–900, https://doi.org/10.1002/(SICI)1097-0088(19980630)18:8<873::AID-JOC255>3.0.CO;2-9, 1998.

Frei, C., Christensen, J. H., Déqué, M., Jacob, D., Jones, R. G., and Vidale, P. L.: Daily precipitation statistics in regional climate models: Evaluation and intercomparison for the European Alps, Journal of Geophysical Research: Atmospheres, 108, https://doi.org/10.1029/2002JD002287, 2003.

Fu, Q.: An accurate parameterization of the solar radiative properties of cirrus clouds for climate models, Journal of Climate, 9, 2058–2082, https://doi.org/10.1175/1520-0442(1996)009<2058:AAPOTS>2.0.CO;2, 1996.

Ganopolski, A. and Calov, R.: The role of orbital forcing, carbon dioxide and regolith in 100 kyr glacial cycles, Climate of the Past, 7, 1415–1425, https://doi.org/10.5194/cp-7-1415-2011, 2011.

Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne, S. R., Lawrence, D. M., Neale, R. B., Rasch, P. J., Vertenstein, M., Worley, P. H., Yang, Z.-L., and Zhang, M.: The Community Climate System Model Version 4, Journal of Climate, 24, 4973–4991, https://doi.org/10.1175/2011JCLI4083.1, 2011.

Gianotti, R. L., Zhang, D., and Eltahir, E. A. B.: Assessment of the regional climate model version 3 over the maritime continent using different cumulus parameterization and land surface schemes, Journal of Climate, 25, 638–656, https://doi.org/10.1175/JCLI-D-11-00025.1, 2011.

Giorgi, F., Torma, C., Coppola, E., Ban, N., Schär, C., and Somot, S.: Enhanced summer convective rainfall at Alpine high elevations in response to climate warming, Nature Geoscience, 9, 584–589, https://doi.org/10.1038/ngeo2761, 2016.

Gómez-Navarro, J. J., Raible, C. C., Bozhinova, D., Martius, O., García Valero, J. A., and Montávez, J. P.: A new region-aware bias-correction method for simulated precipitation in areas of complex orography, Geoscientific Model Development, 11, 2231–2247, https://doi.org/10.5194/gmd-11-2231-2018, 2018.

Güttler, I., Stepanov, I., Branković, Č., Nikulin, G., and Jones, C.: Impact of horizontal resolution on precipitation in complex orography simulated by the regional climate model RCA3, Monthly Weather Review, 143, 3610–3627, https://doi.org/10.1175/MWR-D-14-00302.1, 2015.

Haslinger, K., Anders, I., and Hofstätter, M.: Regional climate modelling over complex terrain: An evaluation study of COSMO-CLM hindcast model runs for the greater Alpine region, Climate Dynamics, 40, 511–529, https://doi.org/10.1007/s00382-012-1452-7, 2013.

Hay, L. E., Wilby, R. L., and Leavesley, G. H.: A Comparison of delta change and downscaled GCM scenarios for three mountainous basins in the United States, Journal of the American Water Resources Association, 36, 387–397, https://doi.org/10.1111/j.1752-1688.2000.tb04276.x, 2000.

Hofer, D., Raible, C. C., Dehnert, A., and Kuhlemann, J.: The impact of different glacial boundary conditions on atmospheric dynamics and precipitation in the North Atlantic region, Climate of the Past, 8, 935–949, https://doi.org/10.5194/cp-8-935-2012, 2012a.

Hofer, D., Raible, C. C., Merz, N., Dehnert, A., and Kuhlemann, J.: Simulated winter circulation types in the North Atlantic and European region for preindustrial and glacial conditions: Glacial circulation types, Geophysical Research Letters, 39, https://doi.org/10.1029/2012GL052296, 2012b.

Hui, P., Tang, J., Wang, S., Wu, J., Niu, X., and Kang, Y.: Impact of resolution on regional climate modeling in the source region of Yellow River with complex terrain using RegCM3, Theoretical and Applied Climatology, 125, 365–380, https://doi.org/10.1007/s00704-015-1514-y, 2016.

Isotta, F. A., Frei, C., Weilguni, V., Perčec Tadić, M., Lassègues, P., Rudolf, B., Pavan, V., Cacciamani, C., Antolini, G., Ratto, S. M., Munari, M., Micheletti, S., Bonati, V., Lussana, C., Ronchi, C., Panettieri, E., Marigo, G., and Vertačnik, G.: The climate of daily precipitation in the Alps: Development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data, International Journal of Climatology, 34, 1657–1675, https://doi.org/10.1002/joc.3794, 2014.

Ivanov, M. A., Luterbacher, J., and Kotlarski, S.: Climate model biases and modification of the climate change signal by intensity-dependent bias correction, Journal of Climate, 31, 6591–6610, https://doi.org/10.1175/JCLI-D-17-0765.1, 2018.

Jerez, S., Montavez, J. P., Jimenez-Guerrero, P., Gomez-Navarro, J. J., Lorente-Plazas, R., and Zorita, E.: A multi-physics ensemble of present-day climate regional simulations over the Iberian Peninsula, Climate Dynamics, 40, 3023–3046, https://doi.org/10.1007/s00382-012-1539-1, 2013.

Kendon, E. J., Ban, N., Roberts, N. M., Fowler, H. J., Roberts, M. J., Chan, S. C., Evans, J. P., Fosser, G., and Wilkinson, J. M.: Do convection-permitting regional climate models improve projections of future precipitation change?, Bulletin of the American Meteorological Society, 98, 79–93, https://doi.org/10.1175/BAMS-D-15-0004.1, 2017.

Kleman, J., Fastook, J., Ebert, K., Nilsson, J., and Caballero, R.: Pre-LGM Northern Hemisphere ice sheet topography, Climate of the Past, 9, 2365–2378, https://doi.org/10.5194/cp-9-2365-2013, 2013.

Knist, S., Goergen, K., and Simmer, C.: Evaluation and projected changes of precipitation statistics in convection-permitting WRF climate simulations over Central Europe, Climate Dynamics, https://doi.org/10.1007/s00382-018-4147-x, 2018.

Lafon, T., Dadson, S., Buys, G., and Prudhomme, C.: Bias correction of daily precipitation simulated by a regional climate model: A comparison of methods, International Journal of Climatology, 33, 1367–1381, https://doi.org/10.1002/joc.3518, 2013.

Leung, L. R., Mearns, L. O., Giorgi, F., and Wilby, R. L.: Regional climate research, Bulletin of the American Meteorological Society, 84, 89–95, https://doi.org/10.1175/BAMS-84-1-89, 2003.

Li, J., Evans, J., Johnson, F., and Sharma, A.: A comparison of methods for estimating climate change impact on design rainfall using a high-resolution RCM, Journal of Hydrology, 547, 413–427, https://doi.org/10.1016/j.jhydrol.2017.02.019, 2017.

Ludwig, P., Gómez-Navarro, J. J., Pinto, J. G., Raible, C. C., Wagner, S., and Zorita, E.: Perspectives of regional paleoclimate modeling, Annals of the New York Academy of Sciences, 1436, 54–69, https://doi.org/10.1111/nyas.13865, 2019.

Maraun, D.: Bias correction, quantile mapping, and downscaling: Revisiting the inflation issue, Journal of Climate, 26, 2137–2143, https://doi.org/10.1175/JCLI-D-12-00821.1, 2013.

Maraun, D.: Bias Correcting Climate Change Simulations - a Critical Review, Current Climate Change Reports, 2, 211–220, https://doi.org/10.1007/s40641-016-0050-x, https://doi.org/10.1007/s40641-016-0050-x, 2016.

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

Maraun, D. and Widmann, M.: Statistical downscaling and bias correction for climate research, Cambridge University Press, 2018.

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

Maraun, D., Shepherd, T. G., Widmann, M., Zappa, G., Walton, D., Gutiérrez, J. M., Hagemann, S., Richter, I., Soares, P. M. M., Hall, A., and Mearns, L. O.: Towards process-informed bias correction of climate change simulations, Nature Climate Change, 7, 764–773, https://doi.org/10.1038/nclimate3418, https://www.nature.com/articles/nclimate3418, 2017.

Mayewski, P. A., Rohling, E. E., Stager, J. C., Karlén, W., Maasch, K. A., Meeker, L. D., Meyerson, E. A., Gasse, F., Kreveld, S. v., Holmgren, K., Lee-Thorp, J., Rosqvist, G., Rack, F., Staubwasser, M., Schneider, R. R., and Steig, E. J.: Holocene climate variability, Quaternary Research, 62, 243–255, https://doi.org/10.1016/j.yqres.2004.07.001, 2004.

Menéndez, C. G., de Castro, M., Boulanger, J.-P., D'Onofrio, A., Sanchez, E., Sörensson, A. A., Blazquez, J., Elizalde, A., Jacob, D., Le Treut, H., Li, Z. X., Núñez, M. N., Pessacg, N., Pfeiffer, S., Rojas, M., Rolla, A., Samuelsson, P., Solman, S. A., and Teichmann, C.: Downscaling extreme month-long anomalies in southern South America, Climatic Change, 98, 379–403, https://doi.org/10.1007/s10584-009-9739-3, 2010.

Merz, N., Raible, C. C., Fischer, H., Varma, V., Prange, M., and Stocker, T. F.: Greenland accumulation and its connection to the large-scale atmospheric circulation in ERA-Interim and paleoclimate simulations, Climate of the Past, 9, 2433–2450, https://doi.org/10.5194/cp-9-2433-2013, 2013.

Merz, N., Born, A., Raible, C. C., Fischer, H., and Stocker, T. F.: Dependence of Eemian Greenland temperature reconstructions on the ice sheet topography, Climate of the Past, 10, 1221–1238, https://doi.org/10.5194/cp-10-1221-2014, 2014a.

Merz, N., Gfeller, G., Born, A., Raible, C. C., Stocker, T. F., and Fischer, H.: Influence of ice sheet topography on Greenland precipitation during the Eemian interglacial, Journal of Geophysical Research: Atmospheres, 119, 10,749–10,768, https://doi.org/10.1002/2014JD021940, 2014b.

Merz, N., Raible, C. C., and Woollings, T.: North Atlantic Eddy-Driven jet in interglacial and glacial winter climates, Journal of Climate, 28, 3977–3997, https://doi.org/10.1175/JCLI-D-14-00525.1, 2015.

Messmer, M., Gómez-Navarro, J. J., and Raible, C. C.: Sensitivity experiments on the response of Vb cyclones to sea surface temperature and soil moisture changes, Earth System Dynamics, 8, 477–493, https://doi.org/10.5194/esd-8-477-2017, 2017.

MeteoSwiss: Documentation of MeteoSwiss gridded data product, daily precipitation: RhiresD, http://www.meteoschweiz.admin.ch/content/dam/meteoswiss/de/service-und-publikationen/produkt/raeumliche-daten-niederschlag/doc/ProdDoc_RhiresD.pdf, 2013.

Mitchell, D., Davini, P., Harvey, B., Massey, N., Haustein, K., Woollings, T., Jones, R., Otto, F., Guillod, B., Sparrow, S., Wallom, D., and Allen, M.: Assessing mid-latitude dynamics in extreme event attribution systems, Climate Dynamics, 48, 3889–3901, https://doi.org/10.1007/s00382-016-3308-z, 2017.

Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., van Vuuren, D. P., Carter, T. R., Emori, S., Kainuma, M., Kram, T., Meehl, G. A., Mitchell, J. F. B., Nakicenovic, N., Riahi, K., Smith, S. J., Stouffer, R. J., Thomson, A. M., Weyant, J. P., and Wilbanks, T. J.: The next generation of scenarios for climate change research and assessment, Nature, 463, 747–756, https://doi.org/10.1038/nature08823, 2010.

Murphy, J.: An evaluation of statistical and dynamical techniques for downscaling local climate, Journal of Climate, 12, 2256–2284, https://doi.org/10.1175/1520-0442(1999)012<2256:AEOSAD>2.0.CO;2, 1999.

Neale, R. B., Richter, J. H., Conley, A. J., Park, S., Lauritzen, P. H., Gettelman, A., Rasch, P. J., and Vavrus, J.: Description of the NCAR community atmosphere model (CAM4), National Center for Atmospheric Research Tech. Rep. NCAR/TN+ STR, 2010.

Nešpor, V. and Sevruk, B.: Estimation of wind-induced error of rainfall gauge measurements using a numerical simulation, Journal of Atmospheric and Oceanic Technology, 16, 450–464, https://doi.org/10.1175/1520-0426(1999)016<0450:EOWIEO>2.0.CO;2, 1999.

Oleson, W., Lawrence, M., Bonan, B., Flanner, G., Kluzek, E., Lawrence, J., Levis, S., Swenson, C., Thornton, E., Dai, A., Decker, M., Dickinson, R., Feddema, J., Heald, L., Hoffman, F., Lamarque, J.-F., Mahowald, N., Niu, G.-Y., Qian, T., Randerson, J., Running, S., Sakaguchi, K., Slater, A., Stockli, R., Wang, A., Yang, Z.-L., Zeng, X., and Zeng, X.: Technical description of version 4.0 of the community land model (CLM), NCAR Technical Note NCAR/TN-478+STR, National Center for Atmospheric Research, National Center for Atmospheric Research, Boulder, CO, 2010.

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

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

Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., Keller, M., Tölle, M., Gutjahr, O., Feser, F., Brisson, E., Kollet, S., Schmidli, J., Lipzig, N. P. M. v., and Leung, R.: A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges, Reviews of Geophysics, 53, 323–361, https://doi.org/10.1002/2014RG000475, 2015.

Raible, C. C., Stocker, T. F., Yoshimori, M., Renold, M., Beyerle, U., Casty, C., and Luterbacher, J.: Northern hemispheric trends of pressure indices and atmospheric circulation patterns in observations, reconstructions, and coupled GCM simulations, Journal of Climate, 18, 3968–3982, https://doi.org/10.1175/JCLI3511.1, 2005.

Raible, C. C., Lehner, F., González-Rouco, J. F., and Fernández-Donado, L.: Changing correlation structures of the Northern Hemisphere atmospheric circulation from 1000 to 2100 AD, Climate of the Past, 10, 537–550, https://doi.org/10.5194/cp-10-537-2014, 2014.

Raible, C. C., Brönnimann, S., Auchmann, R., Brohan, P., Frölicher, T. L., Graf, H.-F., Jones, P., Luterbacher, J., Muthers, S., Neukom, R., Robock, A., Self, S., Sudrajat, A., Timmreck, C., and Wegmann, M.: Tambora 1815 as a test case for high impact volcanic eruptions: Earth system effects, Wiley Interdisciplinary Reviews: Climate Change, 7, 569–589, https://doi.org/10.1002/wcc.407, 2016.

Rajczak, J. and Schär, C.: Projections of future precipitation extremes over Europe: A multimodel assessment of climate simulations, Journal of Geophysical Research: Atmospheres, 122, 10,773–10,800, https://doi.org/10.1002/2017JD027176, 2017.

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

Richter, D.: Ergebnisse methodischer untersuchungen zur korrektur des systematischen messfehlers des hellmann-niederschlagsmessers, tex.publisher: Deutscher Wetterdienst Offenbach, 1995.

Rougier, J., Sexton, D. M. H., Murphy, J. M., and Stainforth, D.: Analyzing the climate sensitivity of the HadSM3 climate model using ensembles from different but related experiments, Journal of Climate, 22, 3540–3557, https://doi.org/10.1175/2008JCLI2533.1, 2009.

RStudio Team: RStudio: Integrated Development Environment for R, RStudio, Inc., Boston, MA, http://www.rstudio.com/, 2015.

Schmidli, J., Schmutz, C., Frei, C., Wanner, H., and Schär, C.: Mesoscale precipitation variability in the region of the European Alps during the 20th century, International Journal of Climatology, 22, 1049–1074, https://doi.org/10.1002/joc.769, 2002.

Schmidli, J., Frei, C., and Vidale, P. L.: Downscaling from GCM precipitation: A benchmark for dynamical and statistical downscaling methods, International Journal of Climatology, 26, https://doi.org/10.1002/joc.1287, 2006.

Schulzweida, U.: CDO User Guide (Version 1.9.6), https://doi.org/10.5281/zenodo.2558193, 2019.

Schwarb, M., Daly, C., Frei, C., and Schär, C.: Mean annual and seasonal precipitation in the European Alps 1971–1990, Hydrological Atlas of Switzerland, Landeshydrologie und Geologie, Bern, Switzerland, 2001.

Seguinot, J., Khroulev, C., Rogozhina, I., Stroeven, A. P., and Zhang, Q.: The effect of climate forcing on numerical simulations of the Cordilleran ice sheet at the Last Glacial Maximum, The Cryosphere, 8, 1087–1103, https://doi.org/10.5194/tc-8-1087-2014, 2014.

Sevruk, B.: Der niederschlag in der schweiz, Kümmerly und Frey, 1985.

Shepard, D. S.: Computer mapping: The SYMAP interpolation algorithm, in: Spatial Statistics and Models, Theory and Decision Library, pp. 133–145, Springer, Dordrecht, https://doi.org/10.1007/978-94-017-3048-8_7, 1984.

Skamarock, W. C. and Klemp, J. B.: A time-split nonhydrostatic atmospheric model for weather research and forecasting applications, Journal of Computational Physics, 227, 3465–3485, https://doi.org/10.1016/j.jcp.2007.01.037, 2008.

Solman, S. A., Nuñez, M. N., and Cabré, M. F.: Regional climate change experiments over southern South America. I: Present climate, Climate Dynamics, 30, 533–552, https://doi.org/10.1007/s00382-007-0304-3, 2008.

Stocker, T., Plattner, G.-K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P., eds.: Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of IPCC the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, doi:10.1017/CBO9781107415324, 2013.

Su, F., Duan, X., Chen, D., Hao, Z., and Cuo, L.: Evaluation of the global climate models in the CMIP5 over the Tibetan Plateau, Journal of Climate, 26, 3187–3208, https://doi.org/10.1175/JCLI-D-12-00321.1, 2012.

Sun, F., Roderick, M. L., Lim, W. H., and Farquhar, G. D.: Hydroclimatic projections for the Murray-Darling Basin based on an ensemble derived from Intergovernmental Panel on Climate Change AR4 climate models, Water Resources Research, 47, https://doi.org/10.1029/2010WR009829, 2011.

Teng, J., Potter, N. J., Chiew, F. H. S., Zhang, L., Wang, B., Vaze, J., and Evans, J. P.: How does bias correction of regional climate model precipitation affect modelled runoff?, Hydrology and Earth System Sciences, 19, 711–728, https://doi.org/10.5194/hess-19-711-2015, 2015.

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

Teutschbein, C. and Seibert, J.: Is bias correction of regional climate model (RCM) simulations possible for non-stationary conditions?, Hydrology and Earth System Sciences, 17, 5061–5077, https://doi.org/10.5194/hess-17-5061-2013, 2013.

Themessl, M. J., Gobiet, A., and Leuprecht, A.: Empirical-statistical downscaling and error correction of daily precipitation from regional climate models, International Journal of Climatology, 31, 1530–1544, https://doi.org/10.1002/joc.2168, 2011.

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

UCAR/NCAR/CISL/TDD: The NCAR Command Language (Version 6.6.2) [Software], 10.5065/D6WD3XH5, 2019.

Ungersböck, M., Auer, I., Rubel, F., Schöner, W., and Skomorowski, P.: Zur Korrektur des systematischen Fehlers bei der Niederschlagsmessung: Anwendung des Verfahrens für die ÖKLIM Karten, p. 5, 2001.

Velasquez, P., Messmer, M., and Raible, C. C.: Code and dataset, https://doi.org/10.5281/zenodo.3243797, 2019.

Warrach-Sagi, K., Schwitalla, T., Wulfmeyer, V., and Bauer, H.-S.: Evaluation of a climate simulation in Europe based on the WRF–NOAH model system: Precipitation in Germany, Climate Dynamics, 41, 755–774, https://doi.org/10.1007/s00382-013-1727-7, 2013.

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

Wilks, D. S.: Statistical methods in the atmospheric sciences, Academic Press, google-Books-ID: IJuCVtQ0ySIC, 2011.

Xu, C.-y.: Modelling the effects of climate change on water resources in central sweden, Water Resources Management, 14, 177–189, https://doi.org/10.1023/A:1026502114663, 2000.

Xu, C.-y., Widén, E., and Halldin, S.: Modelling hydrological consequences of climate change — Progress and challenges, Advances in Atmospheric Sciences, 22, 789–797, https://doi.org/10.1007/BF02918679, 2005.

Yang, B., Qian, Y., Lin, G., Leung, L. R., Rasch, P. J., Zhang, G. J., McFarlane, S. A., Zhao, C., Zhang, Y., Wang, H., Wang, M., and Liu, X.: Uncertainty quantification and parameter tuning in the CAM5 Zhang-McFarlane convection scheme and impact of improved convection on the global circulation and climate, Journal of Geophysical Research: Atmospheres, 118, https://doi.org/10.1029/2012JD018213, 2013.

Yang, W., Andréasson, J., Graham, L. P., Olsson, J., Rosberg, J., and Wetterhall, F.: Distribution-based scaling to improve usability of regional climate model projections for hydrological climate change impacts studies, Hydrology Research, 41, 211–229, https://doi.org/10.2166/nh.2010.004, 2010.

Zhang, G. J. and McFarlane, N. A.: Sensitivity of climate simulations to the parameterization of cumulus convection in the Canadian climate centre general circulation model, Atmosphere-Ocean, 33, 407–446, https://doi.org/10.1080/07055900.1995.9649539, 1995.

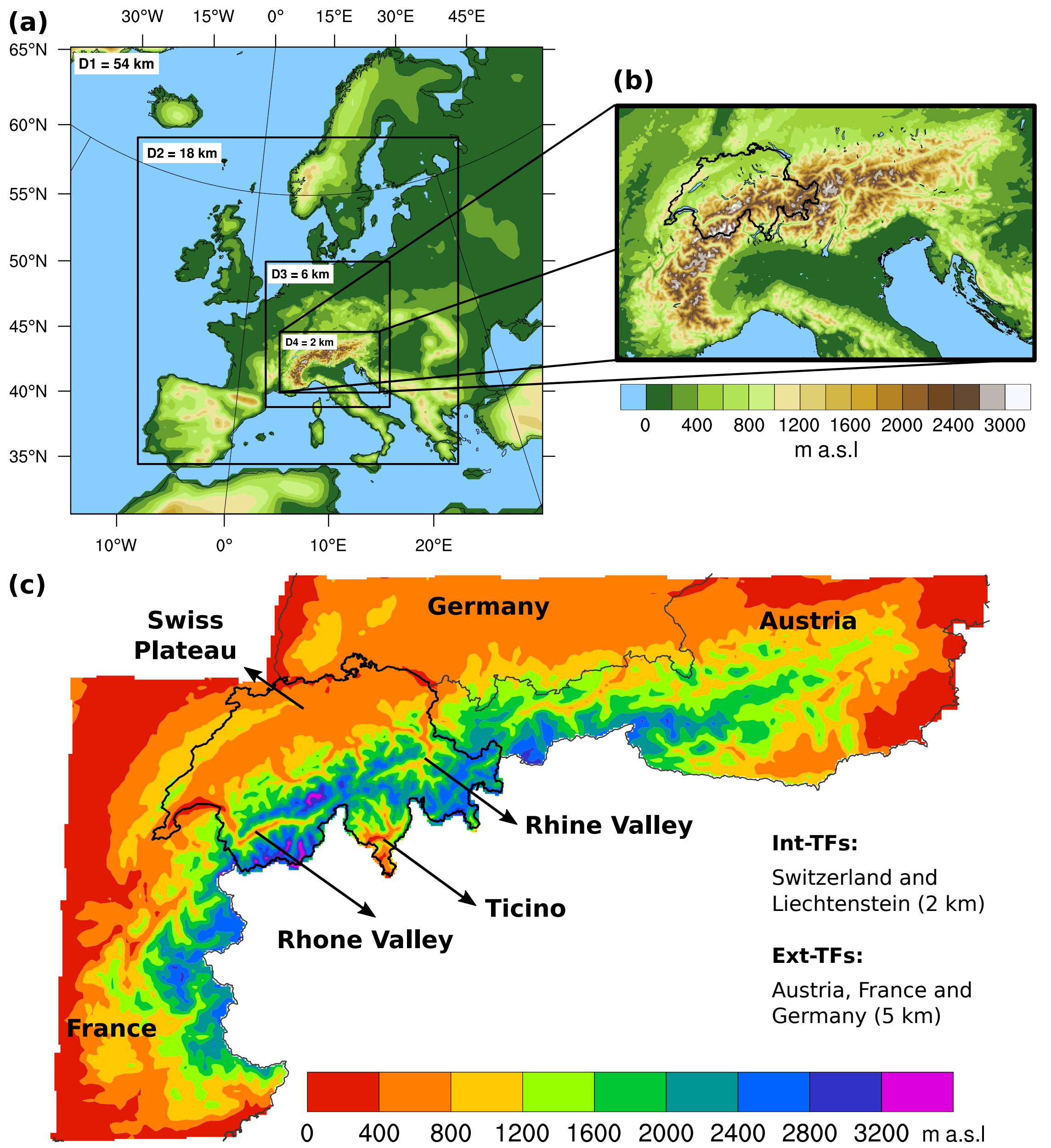


Figure 1. WRF domains and topography. (a) illustrates the topography and the four domains used by WRF. (b) shows the fourth domain including the area of interest (Switzerland) outlined by a black line. (c) indicates the height-classes used for the correction method (400 m interval) for the Int-TFs at 2-km resolution (Switzerland, black outline) and for the Ext-TFs at 5-km resolution (other shaded areas). Additionally, some labels are added to identify some specific areas in Switzerland that are used throughout the paper.

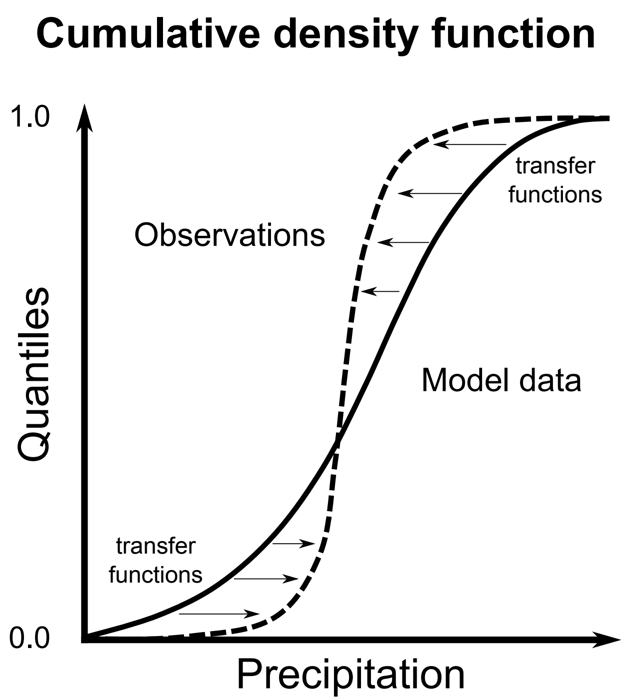


Figure 2. Diagram of empirical quantile mapping technique. Solid (dashed) line shows a schematic simulated (observed) cumulative distribution.

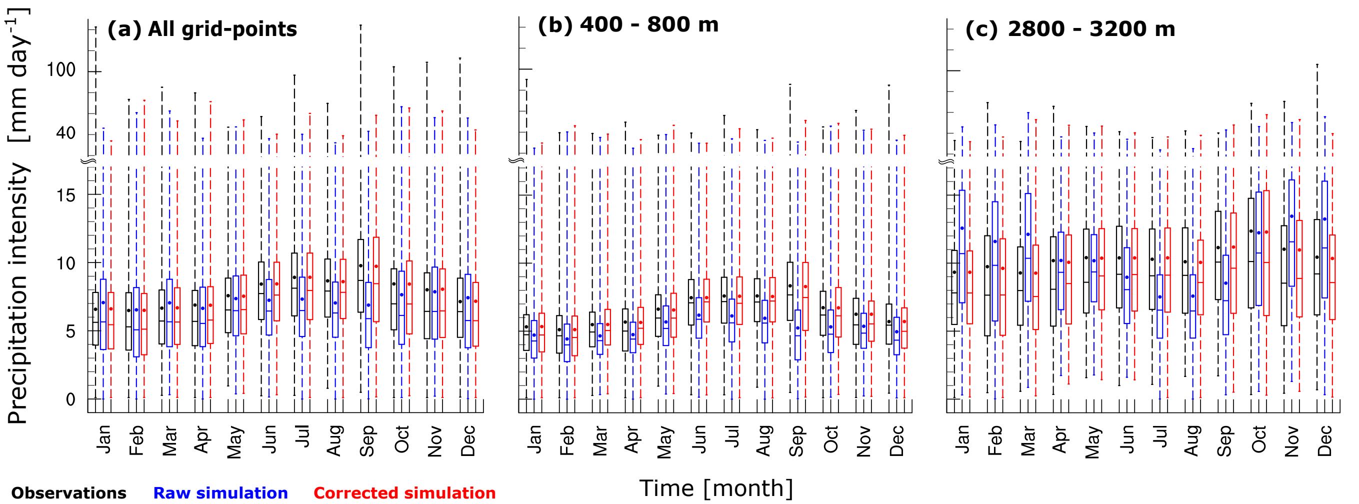


Figure 3. Boxplots ~~are illustrating~~ illustrate the ~~annual cycle and monthly~~ spatial distribution of ~~daily~~ monthly mean values of precipitation intensity across a specific area within 30 years: (a) the area covers all grid points over entire Switzerland, (b) ~~all~~ the grid points in the height class of ~~400–800~~ 400 – 800 m, and (c) the grid points in the height class of 2.800 ~~–3.200~~ – 3.200 m. Black box-plots represent the observations (RhiresD data), blue and red ones the raw and corrected simulation, respectively. Top and bottom ends of the dashed lines represent the maximum and minimum values, respectively. Dots represent the spatial climatological mean value.

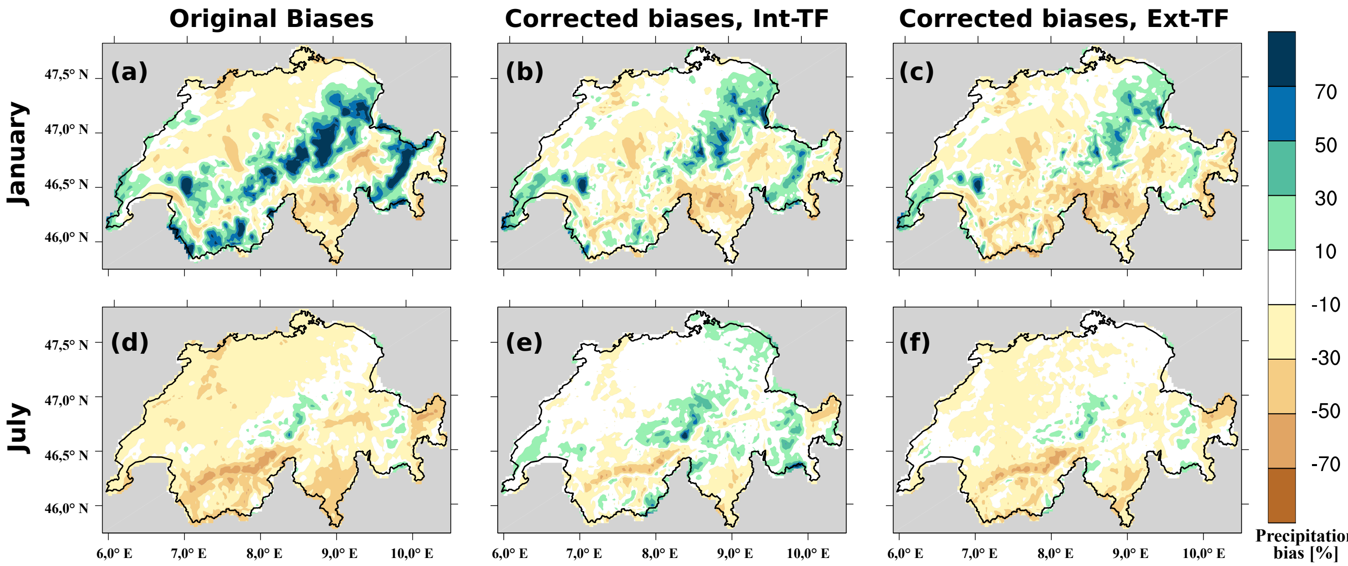


Figure 4. Biases ~~of precipitation~~ in ~~terms~~ the climatological mean value of precipitation intensity over Switzerland. (a) represents the original biases in January, (b) the biases after being corrected using Int-TFs in January, (c) the biases after being corrected using Ext-TFs in January, (d), (e), and (f) as (a), (b), and (c) but in July, respectively.

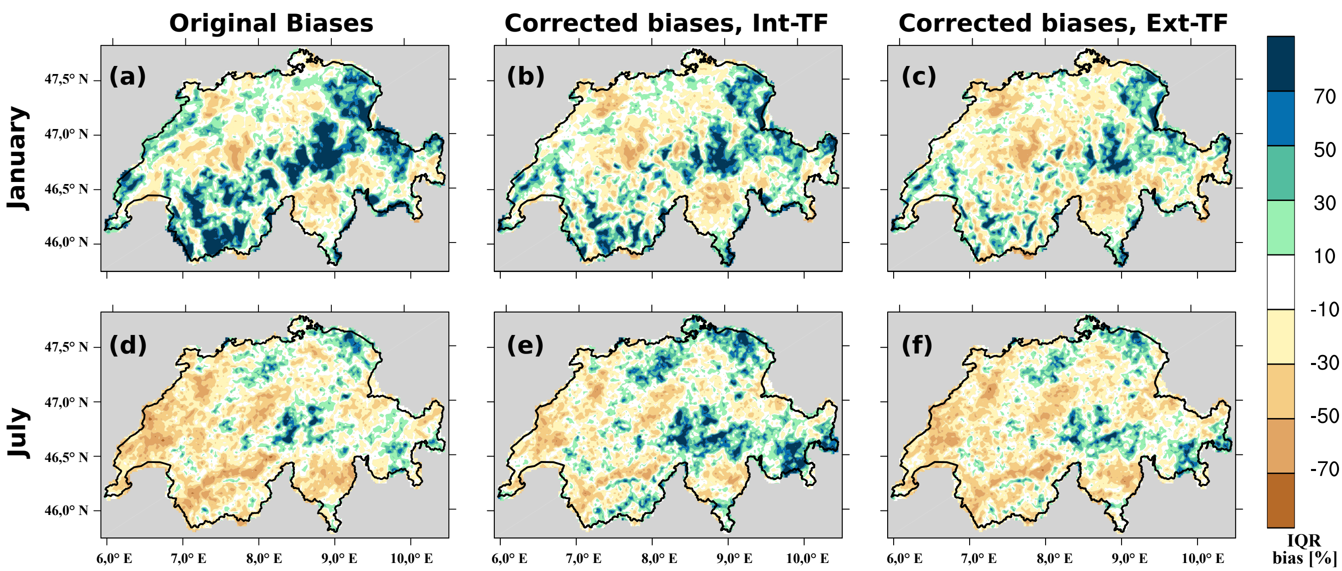


Figure 5. Biases ~~of precipitation~~ in ~~terms of~~ the interquartile range of monthly mean precipitation intensity over Switzerland. (a) represents the original biases in January, (b) the biases after being corrected using Int-TFs in January, (c) the biases after being corrected using Ext-TFs in January, (d), (e), and (f) as (a), (b), and (c) but in July, respectively.

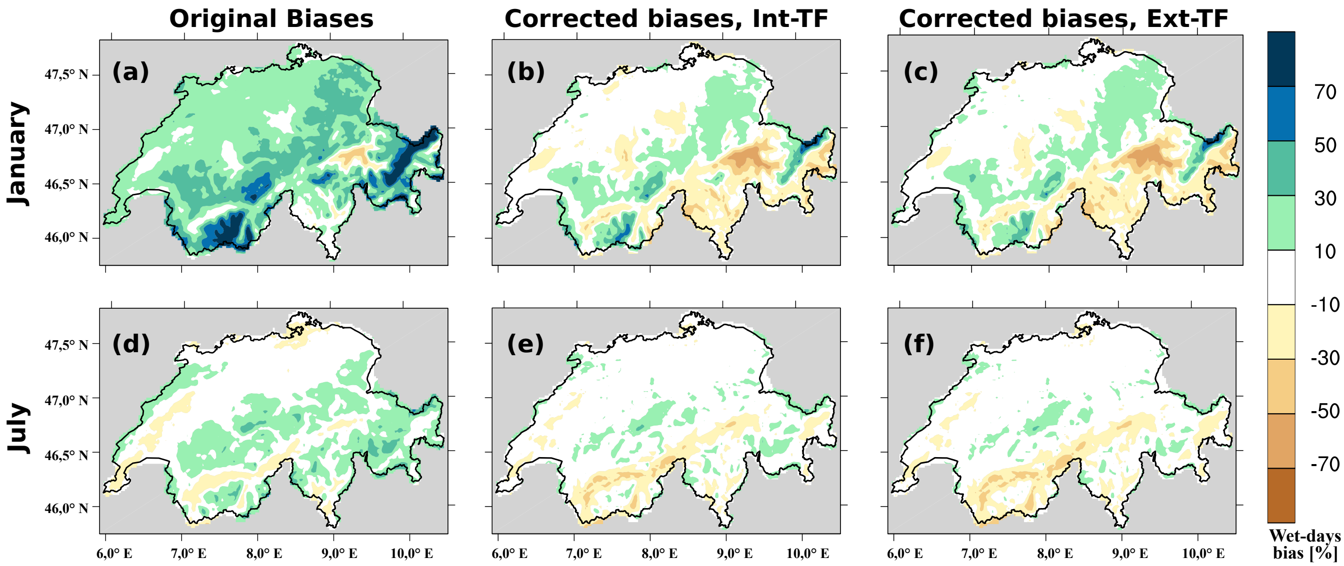


Figure 6. Biases ~~of precipitation~~ in ~~terms of~~ the wet-day frequency within the 30-year period over Switzerland. (a) represents the original biases in January, (b) the biases after being corrected using Int-TFs in January, (c) the biases after being corrected using Ext-TFs in January, (d), (e), and (f) as (a), (b), and (c) but in July, respectively.

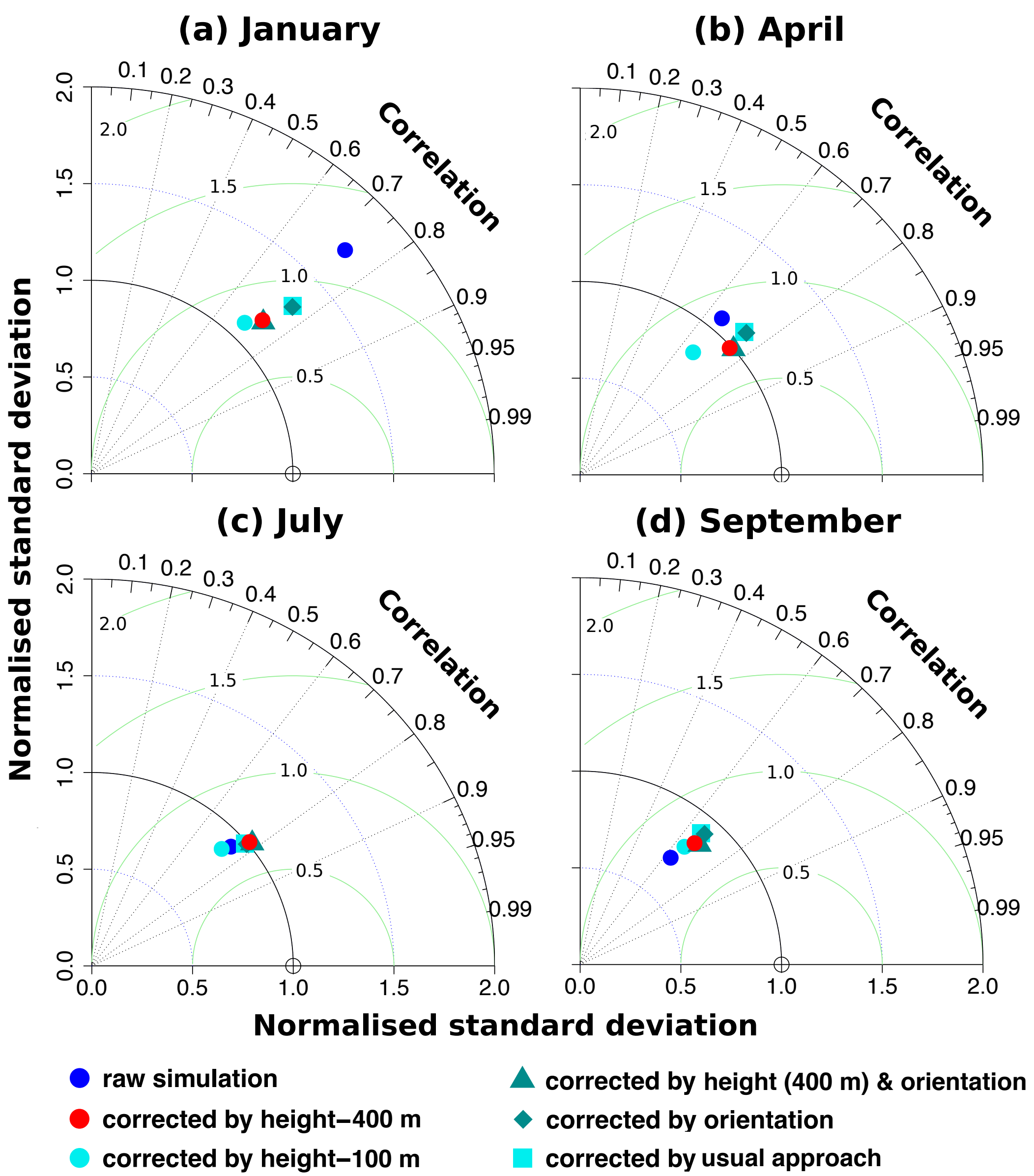


Figure 7. Performance of bias-correction with different settings. (a) shows a Taylor diagram for January, (b) for April, (c) for July and (d) for September. Blue dots represent the raw simulation, red dots the simulation corrected by using height-intervals of 400 m, cyan dots the simulation corrected by using height-intervals of 100 m, petrol triangles the simulation corrected by using height-intervals of 400 m and slope-orientations, petrol diamonds the simulation corrected by slope-orientations, and cyan squares the simulation corrected by the usual approach (the entire Swiss region). Note that in the Taylor diagram the spatial correlation, spatial root-mean-square-error and spatial standard deviation are shown.

Table 1. External forcing used in Hofer et al. (2012a, b) for 1990 AD conditions.

Parameter name	Value
TSI (W m^{-2})	1361.77
Eccentricity	1.6708×10^{-2}
Obliquity ($^\circ$)	23.441
Angular precession ($^\circ$)	102.72
CO2 (ppm)	353.9
CH4 (ppb)	1693.6
N2O (ppb)	310.1

Table 2. Important parameterisations used to run WRF.

Parameterisation	Parameter name	Chosen parameterisation	Applied to
Microphysics	mp_physics	WRF single moment 6-class scheme	Domain 1 – 4
Longwave radiation	ra_lw_physiscs	RRTM scheme	Domain 1 – 4
Shortwave radiation	ra_sw_physics	Dudhia scheme	Domain 1 – 4
Surface layer	sf_sfclay_pysics	MM5 similarity	Domain 1 – 4
Land/water surface	sf_surface_physics	Noah–Multiparameterization Land Surface Model	Domain 1 – 4
Planetary boundary layer	bl_pbl_physics	Yonsei University scheme	Domain 1 – 4
Cumulus	cu_physics	Kain–Fritsch scheme	Domain 1 – 2
		No parameterisation	Domain 3 – 4