Response to Referee Comment

We greatly thank the reviewer for the careful and thorough reading of this manuscript and for the thoughtful comments and constructive suggestions, which help to improve the quality of our manuscript. The comments have been carefully considered and responded. Please find below our response to each comment.

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 to our attention that we missed to show clearly enough the orographic dependence of the biases. 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. Therefore, we have included in the revised manuscript a more balanced discussion of how our approach is removing the biases.

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.

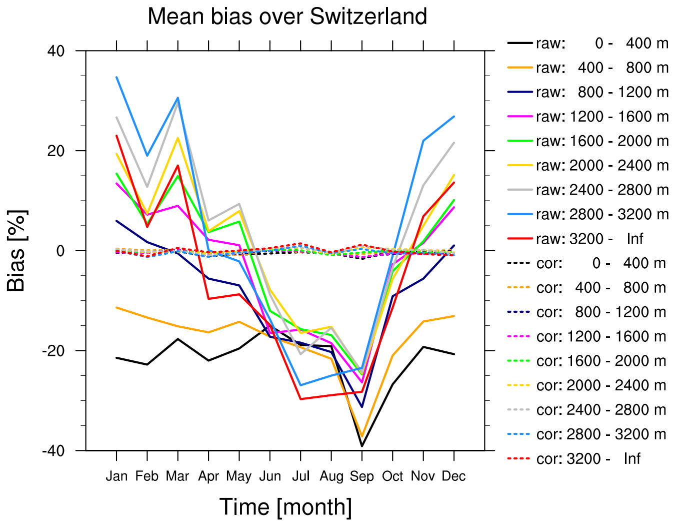


Figure R1. Mean bias over Switzerland for different height-classes.

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 the 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 substantially by using orographic features as additional variables for the correction. 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 from 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 and precipitation-related processes.

To clarify this, we will present a broader discussion on the general shortcomings of bias correction methods in the introduction and the conclusion section in the revised manuscript.

Furthermore, we also believe 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 added the definition of the precipitation in the manuscript and how days without precipitation are treated.

- 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 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 test because there is not any modified observational dataset that considers changes in the topography.

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. With such a method, the correction could be much better verified than with the method suggested by the reviewer.

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 will be 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. We also will include a broader discussion on the limitations of bias correction methods.

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 (2 km) …

- 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 may not be 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.

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 considered as censored values when they fall below a certain threshold. Many studies use a static threshold that is between 0.01 and 1.00 mm day^{-1}, whereas in our study, we calculate different thresholds to be consistent with the differentiate 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 between 0.001 and 1.00 mm day^{-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 (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. The dry days are now considered as censored values when they fall below a certain threshold. Many studies use a static threshold that is between 0.01 and 1.00 mm day^{-1} (Piani et al., 2010a; Lafon et al., 2013;

Maraun, 2013), whereas in our study, we calculate different thresholds to be consistent with the different biases-treatment across the groups (or subgroups) and months of the year. Then, we carry out the first part of 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. 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?

<u>RESPONSE</u>:

We thank you 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. Moreover, the correction is afterwards applied to the daily precipitation in 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.

<u>RESPONSE</u>:

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}]

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

…, the annual cycle and the monthly distributions of daily precipitation are …

…, the annual cycle and the distributions of monthly mean precipitation intensity are …

- 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 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.

<u>RESPONSE</u>:

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. 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); then, we calculate two sets of TFs. The first one is obtained from the same Swiss region called Internal TFs (Int-TF) using the RhiresD data set (at 2 km resolution), and the second one 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) …

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?

RESPONSE:

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. To clarify this, a discussion will be presented in the revised manuscript.

Still, 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 will discuss the uncertainty issue in particular for the results in high altitudes.

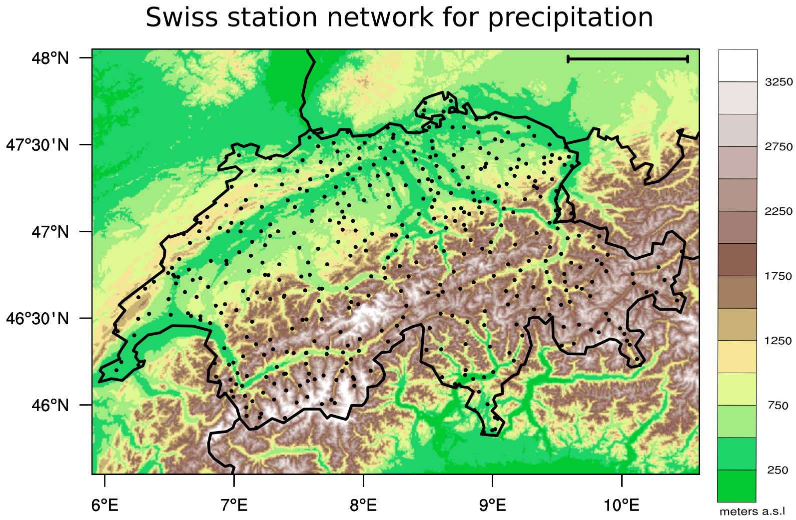
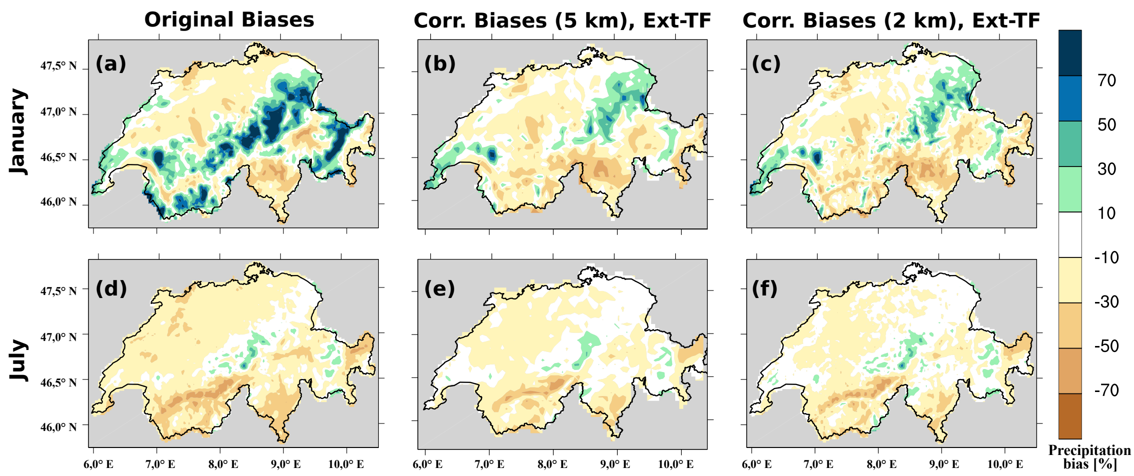


Figure R2. Swiss stations are integrated in RhiresD.

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 first mention that the method uses different observational data sets, which both mostly describe the topography. Moreover, we did it to 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. The other reason of carrying out the application at 2 km resolution is that the application at 5 km show minimal differences on the results, as is shown in the next Figure R3.



R2. 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.

<u>Minor points:</u>

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

We thank you 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, i.e., it rains less but more frequently …

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

We really thank you for pointing this out. To make the argumentation clear, we have re-structured the paragraphs and the change will be 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 mostly 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 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 these lines 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 important 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…

Chen, Di, and Aiguo Dai. 2019. 'Precipitation characteristics in the ommunity 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.

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.

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)

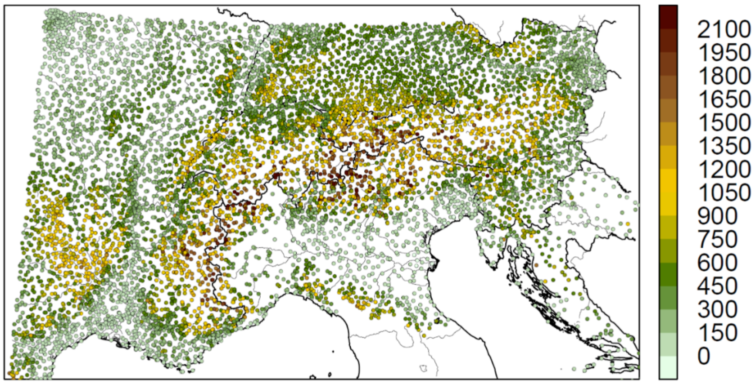


Figure R4. Each point corresponds to a rain-gauge station for which data was available in the the spatial analysis. The color is the height (m) of the station. Source: https://www.meteoswiss.admin.ch/home/search.subpage.html/en/data/products/2015/alpine-precipitation.html)

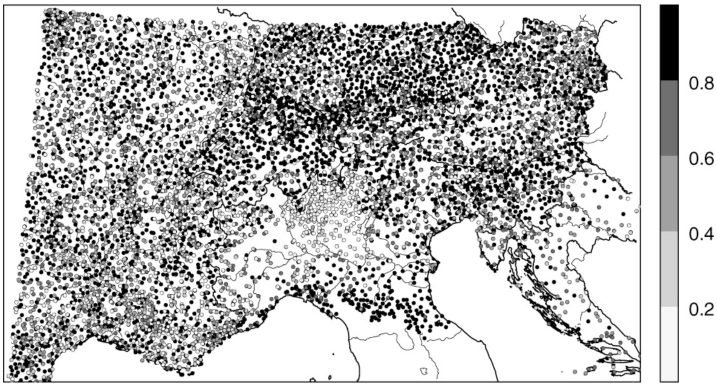


Figure R5. Distribution of stations from which records of daily precipitation are integrated in APGD dataset. Shading represents the fraction of the full period (1971–2008) covered by the respective record. (Isotta et al. 2014)

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.

We fully agree that this remains somewhat unclear. 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.

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 expectations.

Best regards,
On behalf of the co-authors,

Patricio Velasquez