Reply to the reviewers comments

The authors present a new dataset based on a combination of long station records and the analogue resampling method for daily temperature and precipitation. These fields are then adjusted using ensemble Kalman filtering or quantile matching. Both the non-adjusted as the adjusted dataset are validated using a leave-one-out approach and against independent station observations. Finally, an application of the dataset is given in a reconstruction of snowfall and the altitude of the 0 degree line to better understand a historic avalanche winter.

The study is sound and - as far as I can tell - no methodologic errors have been made. The study is a pleasure to read and the application, presented like it is the cherry on the cake, makes a compelling case for the dataset. Although I am quite enthusiastic about this study, there are three aspects which the authors may want to look into. One is the need for some additional explanation, one relates to an issue with the post-processing and the last one relates to the analogue method and a suggestion to overcome the drawback of the limited number of suitable analogues

We thank the reviewer for the very positive feedback and appreciate the valuable suggestions and comments.

1. On page 7 (line 189) is is argued that ’reconstructions are often affected by biases in the mean, an increased number in wet days and underestimation of extreme events’. This statement is corroborated by a reference to Piani et al. This study works with global climate model data and a global dataset of hydrological forcing data. It is common knowledge that such global datasets suffer from the problems described on line 189, but one of the appealing aspects of the analogue method is that it has the potential to avoid these ’smoothing’ problems. After all, it are observed situations that are used to build the reconstruction (including observed extremes) rather than a watered-down statistical interpolation. A more clear view on WHAT the reason is that the ARM provides estimates that have too many wet days, lack real extremes and suffer from a bias. After all, much of the study is devoted to adjusting for these problems.
This is an excellent suggestion. As the reviewer states, the analogue method generally has the advantage over e.g. statistical interpolation to reproduce natural variability and mean values. However, given the assumptions made in the setup of the method (e.g. similarity criterion, coupled reconstruction of temperature and precipitation) and limitations of available data (e.g. size of analogue pool, coverage of station data), also analogue reconstructions can suffer from the problems described in line 189. In the revised manuscript, we will state that more clearly and go further into detail about possible consequences of methodological choices on resulting reconstructions in section 3.1 (see also reply to reviewer's comment RC1) and adopt this argumentation in section 3.2 instead of the mentioned reference to literature. In section 4, the limited size of the analogue pool and relatively sparse station coverage are identified as the main causes for problems regarding the reconstruction of extreme events and the related bias in the mean. Tests for the period 1961-2017 revealed that the size of the analogue pool is limited by the restrictions of the analogue method (seasonal window, weather types) to 1772 on average, with 21% of the days having less than 1000 and about 1% less than 500 possible analogues. Whereas for problems regarding the discrimination between wet and dry days, no detailed assessment has been carried out in order to limit the scope of the manuscript. It could be shown however, that the analogue method is prone to such problems and that especially for moderate precipitation events it fails to correctly reproduce precipitation areas (figure 6).

2. On page 8, line 213, the authors state that the assumption in the post-processing method is that the precipitation distribution is not subject to changes in time. The period the authors use to calculate the parametric transfer functions is 1961-2017. Obviously, this period includes the climate change effects on the precipitation which are also evident in the Swiss climate. Examples of time

series with steep trends and/or decadal variability of e.g. RR1 (number of wet days) are Andermatt and Altdorf, extreme precipitation has changed as well, as evident in e.g. R95p in Basel-Binnigen. Can the authors comment on how climate change and decadal variability affects the effectiveness of the adjustment for precipitation?
Thank you for this important question. While the effects of climate change and decadal variability on precipitation are captured by the analogue method to the extent where they can be found in station data or changes in the occurrence of weather types, post-processing does not take such effects into account. This very simple setting of quantile mapping was chosen to avoid over-fitting to the period 1961-2017, as the correction is applied to the whole dataset back to 1864. However, whether adjustments by quantile mapping show a pattern that can be related to climatic changes or decadal variability has not been analysed. Nonetheless, as quantile mapping does not correct the number of wet days and an increase in the number of extreme events related to climate change is already captured by the analogue method (from station data and weather types), the impact of climate change or decadal variability on the effectiveness of the chosen post-processing approach is limited to the intensity of extreme precipitation. Considering the large uncertainties in the reconstruction of extremes compared to the magnitude of corrections by quantile mapping (see e.g. figure 7), the adjustment can be considered very effective albeit being calibrated for a period subject to climatic changes.

3. A problem with the analogue method, which the authors mention several times in the study, is the limited number of analogues. Earlier, Van den Dool (1994, his section 5) stumbled upon this problem as well and he suggests a way out. He suggests to construct an analogue having greater similarity than the best natural analogue. He considers linear combinations of naturally occuring analogues. There are a few differences between the Van den Dool study and the current study (monthly vs. daily fields for instance), but it may be worth looking into this suggestion as it may make the dataset presented in this study stronger.
Thank you for pointing out this interesting approach by Van den Dool. As our study has the advantage to dispose of a much larger pool of analogues and to cover a smaller area of study than the Van den Dool study, it is easier to find better matching analogues. Together with post-processing, reconstructions show very satisfying results. Nonetheless, it would be worth examining Van den Dool's method for the reconstruction of daily precipitation and temperature fields in Switzerland and we will definitely consider this suggestion for future work.
We will add the following sentence to the conclusions in the revised manuscript: "Another option to address the problem of small analogue pools as proposed by Van den Dool (1994) is to construct more similar analogues by linear combination of several possible analogue dates."

Other (minor) things the authors may want to look into
• page 5, line 120. What is the motivation to set this window to 60 days (and not e.g. 90 of 30)?
For the analogue method, we tested different seasonal windows. In order not to constrain the analogue pool too much but still to have reconstructions with similar seasonal patterns, an optimum was found at about ±60 days. This value is also in line with literature (e.g. Horton et al., 2017; Caillouet et al., 2019; Ben Daoud et al., 2016; all cited in the manuscript).

• page 6, line 172, an observation error of 1C is quick steep - is there a sound reason for taking it that large?
As station measurements can be affected by micro-climatic conditions that are not captured by gridded data and due to larger uncertainties of the earlier observations, a rather conservative observation error of 1°C was chosen.

very very minor remarks
• line 185, in my humble view, observations are not corrected but adjusted (as I think that an observation is not ’wrong’)

• line 223, change ’chapter’ to ’section’
• line 490, the family name of the 2nd author is ’van Leeuwen’ and his initials are P.J.
• caption figure 8, in my print out, the snow precipitation bars are grey and the avalanche acitvity periods are brownish

Thank you very much for these remarks; we will adjust them accordingly in the manuscript.
As for the colors in figure 8: they seem to be matching the description on screen and in my printout.

Reference

Van den Dool, H. M. (1994). Searching for analogues, how long must we wait?. Tellus A, 46(3), 314-324.