Review of

'A new bias-correction method for precipitation over complex terrain suitable for different climate states'

by P. Velasquez, M. Messmer, and C. Raible

Recommendation: major revisions

This manuscript presents a bias correction method that is based on Empirical Quantile Mapping (EQM) with corrections depending on elevation and slope orientation. It is tested using a 2km-resolution RCM simulation for the Alps driven by a GCM. The simulations are 30 years long and represent perpetual 1990 conditions. The method is validated against precipitation observations for Switzerland, with the EQM fitted either using data for Switzerland, or a larger Alpine area. The intended application is bias correction of paleoclimate simulations, in which the topography over glaciated areas is different from current conditions.

The approach is novel, interesting and potentially useful. However the manuscript has substantial shortcomings with respect to explaining the conceptual basis and technical details. Many of my concerns are essentially identical to those already raised by the first reviewer. I will thus not list them all, and comment only where I would like to emphasize issues or add some detail. Major rewriting addressing these comments is required before the manuscript can be considered for publication.

Specific comments

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.

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.

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.

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.

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.

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.

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).