Manuscript: A Combined Statistical Bias Correction and Stochastic Downscaling Method for Precipitation

Major remarks

The authors present a bias correction approach for GCM or RCM precipitation where this post processing is separated into two steps. Step 1 comprises the pure bias correction using a quantile-mapping method at the same scale as the RCM data that are corrected in the study. In Step 2, a downscaling method is applied from the grid scale to station locations. Using vector generalized linear gamma model (VGLM), downscaling parameters are derived from gridded and station observations, and then these calibrated VGLM is applied to the RCM data. This separation into bias correction and downscaling is an innovative and promising approach and, hence, of interest for a wider scientific community.

There are a few points that should be improved before publication:

The selection procedure and evaluation in step 1 is not well described, e.g. on p.7 = Sect. 3.1. Usually, I would expect that only bias-corrected data are used in step 2, which is not the case. The selection procedure is well described in the appendix A2 (especially lines 37-40 on p. 16) but not in the main text. Also I would expect that a bias corrected precipitation map is compared with the uncorrected precipitation data and observations. But suddenly a predictor is mentioned instead of precipitation, and a mixed map is shown only for the station locations. This is rather confusing when first reading the paper.

In step 1, the quantile mapping bias correction improves RCM precipitation in 73 of 86 cases in DJF, but only for 49 of 86 cases in JJA. As quantile mapping can be a rather powerful approach, it seems that the chosen gamma function for the transfer functions fails in a lot of cases especially in JJA. Concluding from this it may be suitable to point to approaches (in the discussion section) where several functions can be used as candidate for the transfer function, such it has been done, e.g. by (Piani, C., G.P. Weedon, M. Best, S.M. Gomes, P. Viterbo, S. Hagemann, and J. O. Haerter, 2010: Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models. J. Hydrol., 395, doi:10.1016/j.jhydrol.2010.10.024, 199–215)

For step 2, it seems even worse. Here, the combined approach provides the best precipitation estimate only for 25 (45) of 86 stations in DJF (JJA). Thus on a first glance, the application of the chosen VLGM does not appear to be a suitable method for the downscaling that can be recommended. What would happen if you use the quantile mapping to directly bias correct the RCM data to the station observations, i.e. using an approach such as it commonly done in bias correction literature? How this would compare to your results?

Actually, I really like the approach of separating bias correction and downscaling. But the rather poor results of the downscaling step may obscure the innovation of the used approach. Thus, some of my remarks aim at that this obscuration does not happen. In summary I suggest accepting the paper for publication after some revisions have been conducted.

I don’t wish do stay anonymous, Stefan Hagemann

Minor remarks

In the following suggestions for editorial corrections are marked in *Italic*.

p.4 – line 1
Please provide an explanation for readers that are not familiar with the "five-fold cross validation". Either short in the main text or long in the appendix with a reference to this in the main text.

p.4 – line 19-20
Although E-OBS is probably *not an appropriate reference in some regions* it …

p.6 – line 28
The term "logit link function" is not common knowledge. Please explain!

p.7 – line 26
…version *since the* calibration …

p.8 – line 31
… performance *of the VGLM gamma for different climates*, …

p.9 – line 5-7
It is written:
To evaluate the goodness of fit we use residual QQ-plots (Fig. 6 for DJF and Fig. 8 for JJA). As a QQ-plot requires quantiles of an unconditional distribution we standardized the from day-to-day varying distribution to a stationary gamma distribution (Coles, 2001; Wong et al., 2014). Thereby the effect of the predictor is approximately removed.

What do you mean with "the effect of the predictor is approximately removed"? You are using E-OBS as predictor. If the effect of E-OBS is removed, would you get the same results with any other predictor? I don't understand.

p.10 – line 10
…correction, *section* 5.1) …

p.11 – line 29
This *raises* the question …

p.12 – line 30-32
It is written: " The spatial …improved by the stochastic downscaling step."

Obviously this statement is correct for DJF. If half of the regions get worse with the downscaling, this questions the general usefulness of the chosen VGLM method (see also major remarks).
In DJF, precipitation is generally strongly determined by the large-scale circulation. Here, the QM bias correction already yields quite good corrected precipitation values. But why the VGLM makes it worse in the majority of cases? This implies a strong weakness of the chosen downscaling method. You provide some potential reasons, but I suggest also coming up with some more details on how this may be improved. You may even undermine this with examples for single stations. For example, you mention "For instance, another distribution in the VGLM…." on p. 15 – line 3-5. Is it possible to provide a plot for one station where another distribution/function is used that improves the downscaling for this particular station?

Figure 2

I suggest including the PRUDENCE regions in the plot as a major part of your evaluation is based on these regions. For example increase the size of the figure and include PRUDENCE regions as boxes with another colour, e.g. red.

Figure 3 caption

… QM *corrected RCM*, triangles: *uncorrected* RCM.

Figure 7 and 9

What benefit do Fig. 7 and 9 provide? Are they necessary or can they be removed?