Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2018-458-AC3, 2018

Hydrology and Earth System Sciences
Discussions

HESSD

Interactive comment

Interactive comment on "Subseasonal hydrometeorological ensemble predictions in small-and medium-size mountainous catchments: Benefits of the NWP approach" *by* Samuel Monhart et al.

Samuel Monhart et al.

samuel.monhart@wsl.ch

Received and published: 16 November 2018

Thank you for your valuable comments. Indeed, QM is not a universal method to bias correct and downscale meteorological predictions and does have its limitation (especially regarding the variance inflation as you stated in your comments). Hence, we agree with your critical points raised concerning the QM methodology and intend to adapt the manuscript accordingly by extending the discussions and point out the difficulties and limitations of the chosen method and providing suggestions for alternative approaches for the readers. Concerning the two recommended changes we plan to include it in the following way. First, as you propose, we extend the description about how we bridge the gap between the resolution of the meteorological forecasts (with 50km spatial resolution) and the gridded observations (with 2km spatial resolution). In short, as you assumed correctly, we use QM in a cross-calibration framework for this purpose. Hence, we use the same cross-calibration framework as proposed in the paper by Monhart et al. (2018) but the point observations are replaced with the gridded observation data. However, we perform a bilinear interpolation to the surrounding grid data of the coarse forecast (for both the stations and the 2km grid) before applying the QM approach. Second, the topic of variance inflation is indeed important and, as you suggested, it might be of greater importance in this study when the forecasts are aggregated again within the hydrological system, compared to the study by Monhart et al. 2018 where single locations are used for the bias-correction and verification. As mentioned above, we first perform a bilinear interpolation from the coarse resolution forecast grid onto the 2 km observational grid, prior to the bias correction and downscaling in the leave-one-year out cross-calibration framework using QM. Hence, some spatial variability is induced and each of the 2km grid point does provide a slightly different information to be downscaled. This might reduce the spatial effect of variance inflation compared to a bias correction (and downscaling) using a nearest neighbour interpolation technique. However, we plan to have a closer look at the variance inflation within our prediction setup and will extend the discussion in the revised version of the manuscript to point out potential effects on the results.

In the following, we give short answers to your general comments and how we plan to include your points raised in the revised version of the manuscript:

Page 4 L7-9 As mentioned above, we do use a cross-calibration framework to bias correct and downscale. This will be included in the revised manuscript.

Zero values are not handled in a specific way. The reason is, that we apply a multiplicative correction where zero values will not cause an issue. With a multiplicative correction, QM does not artificially produce rain, i.e. in case zero precipitation is forecasted the corrected precipitation still provides zero precipitation. In general, weather prediction models exhibit a drizzle effect due to their large grid sizes, meaning that raw model forecast predict too much rain compared to the observation (in case of very low precipitation rates). Hence, zero precipitation values do not need special treatment during the pre-processing. We missed to include this information in the present form of the manuscript and will add this in the revised version.

Page 5 Section 2.4

For each score we will include the range of values to make the figures easier to interpret.

L16

The spread to error ratio is defined as the ratio between the variance of the forecast ensemble (forecast spread) and the mean squared error (MSE) of the forecast ensemble (forecast error). We will explicitly mention how we calculated the spread to error ratio ratio in the revised version of the manuscript.

Page 8 L18 Figure 3

We will increase the readability of the figures according to your suggestions.

L22-23

We agree that the formulation of this sentence is too optimistic and will adapted the sentence by only highlighting positive skill up to three weeks in spring.

L28

As you assume correctly we calculate the NSE based on the mean of the ensemble and the bias corresponds to the mean ensemble bias. We will highlight this accordingly in the method section.

Page 9 L7-8

In general, we agree that the ESP predictions should be reliable by construction as it is found for the Thur catchment. The reduced reliability as well as the bias could be a result of the combination of the two following characteristics. First, the climatology is represented by the average conditions within 20 years, based on the reference simulation. If, for a specific forecast instance, the initial conditions strongly deviate from the climatological estimate, the resulting ESP predictions tend to show a bias especially at shorter lead times until the influence of the initial conditions diminishes. This effect is expected to be more pronounced in snow dominated and fast reacting catchments (as the Verzasca and the Klöntal catchment), whereas in the larger, rain dominated catchments this effect will have less influence on the results. Second, the climatological reference period (1994-2014) and the period of the meteorological observations of the ESP predictions (1980-2014) do not exactly coincide with each other. The meteorological observations used for the ESP covers a longer period and may exhibit a trend in temperature that, in addition, is stronger pronounced between 1980 and 1990. This could influence our analysis and would be an argument to repeat the ESP prediction with using the same period as for the climatology. But on the other hand, to ensure an ensemble size that is large enough, we decided to use the full available period (1980-2014) to generate the ESP ensembles. However, we plan to discuss this in the revised manuscript and put it in relation to existing literature.

L15 As mentioned above, we will specifically include the range of the scores.

Page 12 L15

We will rephrase the sentence accordingly to make a clearer statement that the reliability in the corrected forecasts is a result of the combination of the NWP model and QM, and thus cannot be realized by QM alone.

P14 L9-10 We will rephrase this sentence to account as well for the uncertainty induced by the hydrological forecast model. Furthermore, this will be accounted for when the hydrological forecasts are in addition post-processed, which is intended in future work.

Your input to discuss the reliability of the ensembles verified against observations will be included in the revised manuscript. As you expected, if the ensembles of the reforecasts are verified against the observed streamflow instead of the pseudo observations from the reference simulation, we do see a more pronounced overconfidence of the forecasts as well, especially at short lead times. We plan to mention this in the revised manuscript.

In addition, we will give more detailed answers to all your comments during the upcoming revision process and include your suggestions which we think will clearly improve the manuscript.

References: Monhart, S., Spirig, C., Bhend, J., Bogner, K., Schär, C. and Liniger, M. A.: Skill of Sub-seasonal Forecasts in Europe: Effect of Bias Correction and Downscaling using Surface Observations, J. Geophys. Res. Atmos., 1–18, doi:10.1029/2017JD027923, 2018.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-458, 2018.