Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2020-42-AC5, 2020

Hydrology and Earth System Sciences Discussions
Open Access

Interactive comment on "The use of personal weather station observation for improving precipitation estimation and interpolation" *by* András Bárdossy et al.

András Bárdossy et al.

jochen.seidel@iws.uni-stuttgart.de

Received and published: 2 April 2020

We thank Marc Schleiss for taking his time to carefully read our paper and for his interesting discussion on the methodology. Here are our responses to the major comments

a) The authors should provide more details about the kriging part. - How did you estimate the variograms? (with/without zeros?) - How do the variograms look like? - Which variogram model did you use and how well does it fit the empirical variogram? - How do you deal with cases in which there are not enough data to reliably fit a variogram? - How do you deal with spatial anisotropy and intermit-

Printer-friendly version

Discussion paper

tency during interpolation?

The variograms used for this paper were calculated using the observations of the primary network only. The variograms were calculated on in the rank space which leads more robust results (Lebrenz and Bárdossy 2017). Further as the kriging weights do not change if the variogram is multiplied by a constant in this study the estimation of the range of the variogram was the major task. We assumed that there is no nugget (precipitation amounts are spatially continuous). The possible measurement error was included in the kriging with uncertainty. Anisotropy was not considered, the main reason for this was that the primary network did not give robust results. In the future we intend to estimate anisotropy from the corresponding radar images. The kriging weights are not very sensitive to the choice of the range and the variogram type as it was investigated in the paper (Bárdossy 1988). The variograms used for the second filter are the rescaled (adjusted to the variance of the observed event) variograms calculated from the percentiles. A discussion on the variogram calculation and fitting including the corresponding references will be added to the paper.

b) Ordinary kriging makes rather strong assumptions about the data (such as second-order stationarity). The latter might not be very realistic in heavy localized rain events. Kriging is also relatively slow compared with other deterministic interpolation methods and its accuracy strongly depends on the density and number of primary observations. For example, the estimation and fitting of a variogram model (from a small number of samples) might introduce additional errors into your predictions that are due to modeling choices rather than the quality of the data. So my question is: why did you choose ordinary kriging? Please motivate this choice by some form of cost/benefit analysis, for example by comparing it to simpler, faster alternatives such as inverse weighted distance interpolation or bilinear interpolation (which make different modeling assumptions).

c) Related to the previous comment. Please note that during cross-validation, one part of the error is due to the spatial interpolation method that you use (i.e., kriging). If you had taken a different interpolation method (say IDW or Bilinear), perhaps the usefulness of the PWS data would have been different. I think it is important that you assess this part of the error by using at least one alternative non-parametric interpolation method other than kriging (e.g., bilinear interpolation). My point here is that in some cases, you might see improvement for one particular interpolation method but not for another.

Regarding both comments above, we assume local second order stationarity - this means kriging is carried out using a few neighbouring stations only. The assumption partly accounts for the non-stationarity. There are several studies which compared different interpolation methods for precipitation which in most cases showed that kriging is superior to other techniques. We compared the interpolation with inverse distance and nearest neighbour for the selected events. For all three interpolation methods the usage of the filtered and corrected PWS lead to an improvement of the cross validation. The selected OK approach was superior to the others. We did not want to overload the paper with the other interpolation results. We also tested different Co-Kriging approaches which also lead to improvements, compared to the inverse distance and nearest neighbour interpolations, but remains slightly inferior to the simplest OK approach. Therefore not to overload the paper these results are not included.

d) The cross-validation part lacks crucial details about parameter estimation. For example, did you use the same variograms or recalculate them based on the selected subset of observations? Theoretically, you should recalculate the variograms on the smaller subset.

Variograms were recalculated for each subset. Due to the relatively large number of primary stations and the fact that we used percentiles the change in the ranges was minor.

e) The second step (i.e., amount estimation) involves a quantile mapping. Accord- ing to your Figure A1, this mapping is different for each PWS. However, this would mean that you need to estimate and fit a separate variogram model (with different nugget/range/sill) for each PWS location at which you want to interpolate. Is that correct? This would be computationally heavy. Please add more details to help me understand this.

Variograms of the quantiles are estimated from the primary stations only. Thus there is no need to recalculate the variograms for each PWS. The appropriate quantiles are also estimated from the primary stations for each PWS locations. For each event this requires one additional OK. The example in the Appendix shows the procedure.

f) Wind is known to cause localized biases in rain gauge measurements in the order of 10-30%. The latter are not stationary over time and space and can significantly affect the ordering of your data, therefore violating your model assumptions (i.e., monotonic link between quantiles of primary and secondary variables). This is not catastrophic but will occasionally affect the accuracy of your rainfall estimates and lower the reliability of your method. I think this issue should be clearly mentioned and discussed in the paper, along with the other limitations in the methodology mentioned by the other reviewers.

You are right - wind has a strong effect on precipitation bias. However this applies for both networks. Our methodology is presently focussing on adjusting the PWS to the primary network. We intend to consider wind dependent corrections in the future. Several PWS measure local wind speed this could be used for further investigations.

Printer-friendly version

Discussion paper

g) Tables 3 and 4: Your evaluation of the improvement in terms of a binary response (yes/no) is not very informative. Improved by how much? Some conditional error distributions (for both cases) might help shed some more light on best/worst case scenarios and what to expect in practice.

We'll add one or two figures on showing error distributions. The main reason for this is to provide a transparent evaluation showing that for the majority of the stations and events there is an improvement, but not for all.

h) I agree with Lotte de Vos (referee 1) when she says that more details about the limitations of the method need to provided. I would go one step further and say that right now, the paper is heavily focused (biased?) towards demonstrating potential and improvements over the status quo. However, the numbers suggest there are also a lot of cases in which the PWS data deteriorate the accuracy of the predictions. Perhaps you could show a few of these cases and comment on them. By explicitly showing what can go wrong, you may be able to provide concrete recommendations for future developments.

We do not agree that the paper would be optimistically biased. In Tables 3 and 4 (which you previously criticized) we show the frequencies of cases when the method was better and when it was worse than the standard. This information is usually not provided and shows that there are cases and locations where there are no improvements. Summary statistics as in Table 5 are usually shown and do not provide this information. The locations with no improvements can easily be identified as those where the density of PWS is small. The reason why the PWS bring no improvements for some events is not clear. As the these cases are rare (< 10 % for short durations) we do not consider this as a major drawback. Of course further research is needed to improve the interpolation, but we believe that the current results are encouraging.

Printer-friendly version

Discussion paper

The minor comments will be considered while preparing the revised manuscript.

Bárdossy,A., Notes on the robustness of the kriging system, *Mathematical Geology,* Vol. 20, No.3, pp 189-203, 1988

Lebrenz, H. and A. Bárdossy, Estimation of the variogram using Kendall's tau for a robust geostatistical interpolation, *Journal of Hydrologic Engineering*, **22**, 2017

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-42, 2020.