Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2018-505-RC1, 2018

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Comparison of approaches to interpolating climate observations in steep terrains with low-density gauging networks" *by* Juan Ossa-Moreno et al.

Anonymous Referee #1

Received and published: 3 December 2018

General comments:

The study from Ossa-Moreno, J., et al. aims to compare the performances of four different methods of increasing complexity, from Inverse Distance Weighting (IDW) and Lapse Rate (LR) to the Generalized Linear Mixed Model (GLMM), for spatially interpolating daily/monthly ground station observations of temperature and precipitation from a network with low spatial resolution. The study region is located in the area of the Aconcagua river basin, a mountain catchment located in the Central Andes of Chile. The comparison is performed using a leave-one-out cross-validation technique based on Root Mean Square Error (RMSE), integrated in the case of precipitation by two

other indices, the probability of detection and false alarm ratio. The authors also aim at evaluating the sensitivity of these methods to the number of available ground stations.

I think the topic of the manuscript quite interesting, potentially helping to provide valuable tools based on the integration of different sources of data (e.g., ground stations networks and remote sensed observations) especially in remote mountain areas where station networks are sparse, unevenly distributed, and difficult to maintain.

However, I think that in the present form this manuscript is not suitable for publication and I suggest that a major revision is necessary from the authors, advising them to provide effective responses to the issues here evidenced.

Among all, the present description of the methodology cannot help the reader to correctly interpret the showed results. Core methods are described but there is a general lack of clarity on the type of data in input to each interpolating method. Also, it is not clear which is the final product of these methods (e.g., daily/monthly gridded data on a regular grid of a specific resolution). This information is of particular interest given that the authors stated that the final aim consisted in providing inputs for hydrological and water resources models.

Another important aspect is associated with the sensitivity test with respect to the number of the gauges used for the interpolation. The authors compare the performances of the four interpolating methods, using data from unevenly distributed stations from a network with decreasing spatial resolution from about 1 gauge/400km2 to 1 gauge/4000km2. In this condition, it is not surprising that even a sophisticated (and high computationally demanding) method as the GLMM provides poor performance. Still, without more information on the methodology (see previous paragraph) it is not possible for the reader to interpret the results. This is particularly true since one of the main outcome of this study is that '…the WorldClim approach (ed., a combination of IDW with gauge data and WorldClim maps, monthly historical averages obtained by statistical analysis of worldwide weather observations between 1950 and 2000 and interpolated using latitude, longitude and elevation) may be recommended as being the more accurate, easy to apply and relatively more robust to tested reductions in the number of estimation gauges, particularly for temperature'. On the other hand, even with the lack of information on the methodology, I find very interesting that, besides the study region is a mountain area characterized by a very pronounced topographic gradient, based on the full spatial resolution of the gauges network (∼1 gauge/400km2) almost all methods seem to perform quite well for temperature at daily time resolution. Unsurprisingly, for precipitation, whose character is highly stochastic, the performances of all methods result so poor at daily time resolution that the comparison is performed at monthly time resolution. Concerning this specific aspect of the study, it is not clear why the authors did not apply to the CHIRPS gridded data (Climate Hazards Group InfraRed Precipitation with Station data) a similar approach to the one they used for the WorldClim maps, which would be undoubtedly very interesting especially at the daily time resolution.

Finally, the manuscript is not easy to read, the structure of many sections is confusing, mixing different aspects; more clearly structured sections would be advisable. Number of figures is unusually large, they are not optimized and with a poor layout. Also, most of them could be grouped in multi-panel figures to allow for an easier and faster results comparison.

Specific comments

1. Section 2.1: I would suggest the authors to separate the description of the geographical and climate settings. If the authors are interested in considering in their study the impact of climate variability on the model-parameter estimation, then the climate setting would deserve a more extensive description, including major and relevant literature and clearly provide known impacts on the variability of temperature and precipitation in the region under study. Fig. 1 is not easy to read, redundancy could be reduced by eliminating the actual large map, enlarging the small one and clearly drawing the divide of the portion of the catchment under study. I would suggest changing the color

of the divide being very similar to topography background, therefore difficult to be distinguished. The authors mentioned the importance of glacial/snow melt, a comment on the presence of glaciers in this catchment would be of interest. For example from the Randolph Glacier Inventory [RGI-Consortium, 2017] could be of help.

2. Section 2.2: I would suggest to clearly state at the very beginning the total number of time series and the maximum time span covered by the considered time series. Fig. 2 and 3 could be merged in a 2-panel figure.

3. Section 2.3: I would include references of studies that have evidenced decreased skill of remote sensed products in the mountain environment (lines 27-29). Fig. 4 and 5 could be merged in a 2-panel figure. What is the DEM dataset for? Only for the regressions? Also, the authors could consider including a plot of the MEI index discussing the occurrence of El Nino or La Nina years during the period for which data are available. Given the short length of the used time series, it could be difficult to have enough ENSO cycles to get a significant correlation between the observations and the MEI index.

4. Section 3: in the first paragraph (lines 2-11) it is not clear if the authors refer to literature or to the methods that were used in the present study. I would suggest being clearer and more direct. In fact, it would be useful for the reader to have in this introductory paragraph of the applied methods a structured list of the methods, possibly referring to the literature for advantages/drawbacks. Also, It would be also useful to discuss why the GLMM method (which provides '. . . larger flexibility to analyze random effects. . .' than GLMs) is potentially a good tool for interpolating daily temperature and precipitation observation in a complex mountain region. Also, it is not clear which type of data (ground station data or spatial data) are used as inputs for each method, which is quite important information that should be integrated. Finally, it is not clear at which resolution the final interpolated variables (temperature and precipitation) are provided (the same for each method, i.e. the WCA resolution?). The authors are strongly suggested to provide this information.

5. Section 3.1: Within the GLMM method description the reader finds that monthly data and not daily observations (as they were initially defined) are used by the authors for precipitation. This generates confusion, I would suggest to state clearly in the abstract and data description (section 2.2) which is the time resolution for both datasets. If precipitation data have monthly time resolution, it is not clear which added information would bring the test indices POD and FAR.

6. Section 3.2: Is the WCA method based on IDW using both station data and WC map data? Table 1: please consider adding a column for the data used as inputs (station data, spatial data, DEM,. . .). Fig. 6 and 7 could be merged in a 6-panel figure.

7. Section 3.3: I suggest the authors to clearly divide the two comparison tests, LOOCV and sensitivity and avoid mixing the two tests. The authors also add that the RMSE estimation for the GLMM method was performed using the expected values of each variable for each time step (line 18). How it is calculated for the other methods? Should not be the same? That is, for all the methods, do not the authors generate time series of gridded data that are compared then with the station values with the LOOCV method? This aspect should be clarified; otherwise it is very difficult to correctly interpret the results in Table 3 and 4. In fact, it is not clear how the authors could obtain a so small RMSE for the raw WorldClim maps values (monthly worldwide estimated averages compared with daily station data?).

8. Section 4: In the results section, the authors provide correlation values but they do not explain the purpose of this part of analysis. It seems quite an important pre-processing step that aims at testing which variates are to be used for parametrizing the relevant interpolating schemes. I suggest the authors adding a paragraph in the methods section that explains this aspect. Fig. 8 to 13 could be merged in a 6-panel figure.

9. Section 4.1: How did the authors calculate the daily temperature averaged over the 5-y period? Is i-th daily value the simple mean of all days i (i.e., average of 5 values)?

Or did the authors consider a moving window k-days long (i.e., average of 5*k values)? I would suggest clarifying this and add a sentence describing how the data in Fig. 14 to 16 have been estimated including smoothing method description. Furthermore, it would be also interesting to provide a comment on why we look at the daily temperature averaged over the 5-y period, i.e., which added information this comparison provides. Finally, Fig. 14 to 16 could be merged in a 3-panel figure.

10. Section 4.2: Why the CHIRP data have not been used in association with the station data? RMSE comparison with WCA indicates a better performance of CHIRPS data in the raw configuration, therefore it is expected that they would much better perform in combination with station data than the rest of the methods. Consider merging Fig. 17 and Fig. 23 (maybe scatter plots are easier to be looked at), Fig. 19 to 21, Fig. 18 and Fig. 22.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-505, 2018.