Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2018-316-RC2, 2018

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Consistency of satellite precipitation estimates in space and over time compared with gauge observations and snow-hydrological modelling in the Lake Titicaca region" *by* Frédéric Satgé et al.

Anonymous Referee #2

Received and published: 22 November 2018

The manuscript presents an evaluation of several satellite-based precipitation products over a watershed in the Central Andes, using three different evaluation methods: (1) direct comparison with rain gauge data; (2) evaluation of the Nash-Sutcliffe efficiency of the calibration of a hydrological model forced with each precipitation data set on observed stream flow; and (3) evaluation of a hydrological model's capacity to predict snow cover when forced with each of the products. The study finds a quite consistent performance of the products over the three methods, with in general more recent and higher-resolution products performing best.

I agree that there is a need for better insights in the performance of satellite-based precipitation products, especially over complicated and data scarce terrains such as the central Andes. This makes the study relevant from both a scientific and operational perspective.

However, despite the manuscript's length, I do not think that the presented findings are significant enough to merit publication. The main issue is that the evaluation methods applied in the study are not designed to gain any insights in the underlying processes that differentiate the different products. Instead, the evaluations seem to be selected based on two specific applications: predicting streamflow in medium-sized watersheds, and predicting high-elevation snow cover. I can see how those applications may be very relevant in a local operational context, but they add very little to the scientific understanding of satellite precipitation and how satellite-gauge merging algorithms can be improved.

Indeed, despite the author's claim to present a new "protocol" for evaluation, the applied methods are very common (apart from perhaps the snow prediction) and some of the implementation decisions further reduce their capacity to gain process insights, in particular:

- The temporal aggregation to a 10-day period essentially reduces the test to an evaluation of the bias and seasonality of the satellite products, eliminating any insights in their capacity to capture individual events and higher intensities, and their propensity for false alarms.

- Focusing only on the calibration of the hydrological model provides little added value to the direct comparison with rainfall. I agree that a comparison with discharge data is warranted, as the provide a useful independent data set that makes it possible to evaluate potential biases of the precipitation product over a catchment area. This is potentially useful, because of the inherent weakness of comparing point rain gauge data with pixel-average SPEs. However this benefit reduces with increasing size of the rain gauge network. At the same time a hydrological model-based evaluation has other issues, such as errors in the model structure and ET estimates. None of this is discussed in detail.

- the decision to exclude elevation as a co-variable in the interpolation process nor to analyse the elevation-dependence of SPE performance explicitly, seems a wasted opportunity. Indeed, one of the main advantages of the study region is to understand the performance of the satellite products as a function of topographic characteristics. The cited study that shows that hydrological models are not very sensitive to elevation-dependent rainfall interpolation is in my opinion not a valid argument. In the case of SPEs there are good arguments to expect an elevation-dependent performance.

- I am afraid that I fail to see the purpose of the evaluation using remotely sensed snow cover data. I agree that solid precipitation is a major issue in SPEs that needs to be studied further, and that the study region would be an excellent opportunity to do so. I also agree that SPEs may be used to model river basins with snow cover, where such issues may propagate. But the implementation presented here does not generate any significant insights in either of these issues and I am really left wondering how can be learned from the presented results.

Because if these issues, I think that the current manuscript has only very limited scientific significance beyond the local scale, in the sense that none of the conclusions are sufficiently solid to gain significant insight in how the products may perform in other regions.

At the same time, the paper is very long and contains a lot of information that is readily available in the relevant literature and does not need to be repeated here. In fact, I think that the manuscript could easily be condensed into 1/2 or even 1/3 of the current length. This would make it much sharper and easier to assimilate the presented information.

However, in addition I think that it needs further analysis to increases the scientific significance and provides a more integrated and purposeful evaluation, as opposed to

the current combination of methods, which feel disjoint and ad-hoc. In my opinion this is would have to be a (very) major revision.

Some other issues I identify are:

- I think that it is a missed opportunity not to include IMERG. Of course IMERG does not go as far back in time as the other products. But given that the temporal analysis does not yield much of a trend, I don't think that that is a major issue. At the same time, the added value of the GPM data probably makes it currently one of the most relevant products from a water resources management perspective.

- The language is often rather imprecise, which may lead to misinterpretation. I did not have the time to make an exhaustive list, but a couple of examples include:

p12/8: "overloading": I don't think that you can "overload" research with "unnecessary" results of "useless" SPEs. Al this sounds rather unscientific. It is fair that a subset of product s is used to reduce the workload, but ideally on a scientifically sound and transparent basis.

p12/10: "Pref is influenced by the interpolation process": This is rather euphemistic as Pref is the direct result of the interpolation process.

p13/13: "regional" -> "spatially averaged"

P32/21: "globally": do you mean over the study region? This is surely not global?

Some other specific comments:

p8/6: no orographic effect: I don't follow the reasoning here. Why would accounting for the orographic effect not allow for an objective comparison?

p11/13: "useless SPEs": "useless is quite a strong word. Surely they are not useless for some applications. Perhaps a more scientific formulation can be found?

p12/13: "when more than 80% of daily values records were available": This is not very conservative and may lead to large errors. I suggest to take only 10-day periods that have complete daily records.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-316, 2018.