Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-185-AC2, 2016

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Evaluation of snow data assimilation using the Ensemble Kalman Filter for seasonal streamflow prediction in the Western United States" *by* C. Huang et al.

C. Huang et al.

anewman@ucar.edu

Received and published: 6 October 2016

We greatly appreciate this anonymous referee for your thoughtful and positive comments on this manuscript. We have revised the manuscript accordingly. Below are detailed responses to the points raised. Our responses are in bold.

Anonymous Referee 2, Received and published: 4 September 2016

I found the topic relevant to HESS and a contribution to DA understanding for water resources in snow-dominated watersheds. While I found the paper well written, it was often difficult to follow because of the number of DA-model scenarios and corresponding acronyms (though I struggled to come up with good alternatives). I also

thought the results section lacked specifics and overly asked the reader to interpret the figures/tables. Finally, I found the major contribution of the paper to be its potential utility for improving streamflow prediction in watersheds with relatively low model skill. I would like to see the authors leverage their previous work to highlight the utility of the approach presented. It should be noted that I reviewed 'version 2' of the manuscript.

Response: Thank you for your overall summary comments of the paper. We agree with the general comment that additional specific analysis can be added to the results section. We will clarify the text throughout, with emphasis on the results section. Our replies to your specific comments give more detail to this general response.

Comment 1: Include more detail in the results. The reader is left to do most of the work in interpreting and quantifying many statements. Tell us how much and where things were improved and where they were not. Statements like this on line 211: "However, we also note that the ensemble observations of 7-day window can have a large variance, likely due to the more limited sample size for the regression, which can negatively impact DA performance (see Supplement Tables S1.1 and S1.2)." would strongly benefit from specific number. What is large variance? What is a negative impact to DA?

Response: The negative impact is truly a reduction in the positive impact of DA when comparing the 7-day window to the 3-month window. We will include specific numbers and clarify this text.

I became frustrated having to look at all the figures and table to understand what was meant by sentences like this. A number of examples are listed below, but I encourage the authors to re-read the manuscript to address this problem completely. Lines 221-225: Where by how much?

Response: We agree that the results section needs more analysis clearly stated in the text. We will tabulate key results for the metrics across example metrics for the entire basin set and add those results to the text.

Line 227-228: Which basins? By how much?

Response: Again, we will revise this discussion.

Lines 243-246: Improves runoff forecasts by how much

Response: We will quantify the improvement to daily flow for the example basins given.

Comment 2: Can you remove some of the acronyms or more clearly explain every acronym in the figure captions.

Response: Yes, we agree these need to be more clearly defined for each figure, or removed entirely in the captions. We will do that upon revision.

Comment 3: There should be more discussion of why the DA could make predictions worse and where that occurred. Should we be worried about this for future DA efforts? How might we screen sites to ensure that DA does not make predictions worse?

Response: We can see from the right two subplots in Figures 10-12 that the years when DA makes the simulated runoff worse is when runoff error is generally very small. Generally, those SWE increments are less than 10In terms of observational sites, the Merced River is the only basin to use state of California SWE observations, and these may be of lower quality as evidenced by the large amount of manual quality control we had to perform on the data and the quality control discussion of these data in Lundquist et al. (2015). This suggests that observed SWE data need to be of higher quality (or information content) than the calibrated model SWE to have a positive impact in the DA system. Conversely, there are years where the noDA runoff error is large, but the SWE increment is small in all three basins. This is not unexpected as spring SWE is not perfectly correlated with subsequent runoff. This may also hint at a level of data loss in the EnKF and modeling system, future work should compare streamflow hindcasts using this type of system with traditional statistical methods using SWE

as a primary input. We believe screening of observational sites is a difficult task. The above discussion and our results in California does suggest screening is needed. High quality sites with no information content would also need to be screened as well (also see discussion in reply to comment 4). It is possible that guidelines for this could be developed and then potentially automated, but this is likely a major undertaking. In this study, site selection was first taken using closest distances to the basin, then manual screening of suspect sites and sites that had little relationship with runoff were removed. It is possible some formalization of this methodology could be developed. That being said, the relationship between SWE and runoff will likely be basin dependent and the addition of an assimilation system and model forecast introduces information losses that are also likely basin dependent since the hydrologic modeling system is basin dependent, such that a screening methodology based solely on observations is likely to misidentify potential degradation or improvement when DA is applied.

Comment 4: It seems that one of the major contributions of the paper is pointing out that DA methods are likely only make improvements in snow dominated watersheds when model performance was <0.80 NSE. Given that Newman et al., 2015a has quantified the performance of SAC-SMA skill in >500 watersheds, I think a major contribution would be to discuss how many watersheds could benefit from DA and how they are spatially distributed. I think that this should be discussed in the context of where the DA methods did not perform well, i.e. comment 2.

Response: This idea you mention is an interesting topic. We will look back through the database and add some additional analysis examining spatial location of basins that may benefit from DA using the basic metrics of noDA NSE and contribution of SWE to runoff. That being said, a comprehensive description and analysis about how many watersheds could benefit from DA and how they are spatially distributed is a large topic and could be a separate paper. Preliminary screening of candidate basins would not only require the basic metrics of

being snow dominated, generally lower noDA skill, but also somehow assessing the quality of information from the nearby observation sites. Furthermore, we'd expect that implementation of the enKF DA would result in potential differences as there may be data loss in the observation transformation operator, etc.

Minor comments: 1. It seems odd to combine the discussion and conclusions section.

Response: We will revise the last two sections to be Results and Discussion and Summary. More discussion will be included in section 4, while the summary section will restate key discussion points and then findings of the study.

References: Lundquist, J. D., M. Hughes, B. Henn, E. D. Gutmann, B. Livneh, J. Dozier, and P. Neiman, 2015: High-elevation precipitation patterns: using snow measurements to assess daily gridded datasets across the Sierra Nevada, California. J. Hydrometeorology, 16, 1773-1792. doi: 10.1175/JHM-D-15-0019.1.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-185, 2016.