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.

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.

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? 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? Line 227-228: Which basins? By how much? Lines 243-246: Improves runoff forecasts by how much

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

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?

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.

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