Italicized text: Reviewer's comment

AR: Authors' response

Comments to reviewer #2

The manuscript titled Combined assimilation of stream flow and snow water equivalent for mid-term ensemble stream flow forecasts in snow-dominated regions is an attempt to apply the Ensemble Kalman Filter to the application of stream flow prediction in regions where the majority of the water involved comes from snow.

The manuscript undertakes a set of studies, the first is to ascertain which of the 7 possible state variables they consider are sensitive to the assimilation of the 3 observation types they consider. They indicate that snow cover appears to not be an important factor in the forecasts that they seek.

The manuscript is well written and upon a second reading easy to follow. However, i do have a couple of points that need to be addressed before i can sign off on publication

Major Comments:

1) My first concern relates to the generation of the ensemble perturbations and the perturbations to the observations. You indicate that you use Gamma, lognormal of beta distributions yet the EnKF is highly reliant on these perturbations, and hence the errors being Gaussian distributed. My query, and i am requesting graphs of these, is to see the distributions plots for the distributions that you mentions with the parameters in the manuscript. My hunch is that these will look quite close to a Gaussian distribution of some form and as such is why you obtain the results, which are great results, but it could be misleading to have these distributions when really they are close to a Gaussian.

AR: Many studies skip the controlled experiment and go straight to the assimilation of real data, but we feel this is not the ideal approach. As a first step to test DA potential on streamflow predictions, a near-ideal framework should be constructed in order to reduce the number “outside variables” that can influence results and mislead the analysis. Since the EnKF is used, all perturbations should have a normal distribution in order to obtain optimal results. However, observations like the ones used in our study have physical limits that cannot be breached (eg: SWE must be >= 0). Alternatives approaches must be used. In our study, we decided to use different distributions that 1) resemble normal ones when the mean is away from the limits, 2) prevent any violation of the physical limits of the variable and 3) are unbiased. The distributions introduced in the manuscript fit that description.

Examples of beta distributions obtained using the parameters presented in the manuscript, along with their analogous normal distributions obtained using the same mean and variance, are shown below (Fig. 1a of this response, has been added as Fig.4a in the revised manuscript). The variance is defined as a function of the mean in such a way that it is largest at 0.5 and smallest at the extremes in an attempt to reflect MODIS SCA retrieval’s greater uncertainty during the transition periods when patchy snow is

prominent. The beta distribution also prevents perturbations from violating the physical limits of the variable, which the normal distribution cannot guarantee. A visual glance at the graphs shows that the beta distributions resemble normal distributions most near 0.5 and deviate more as the mean gets closer to the extremes.

In a similar fashion, examples of gamma and lognormal distributions using the parameters (variance values) described in the manuscript are show in Fig. 1b of this response (has been added as Fig.4b in the revised manuscript). A mean of 1 has been used for all examples for an easier comparison between examples, but the distributions themselves are visually independent of the mean since the variance is defined as directly proportional to the mean. The distributions resemble most a normal one for small variances relative to the mean, but important deviations can be noted as the relative variance increases due to the lower limit (zero) imposed on gamma and lognormal distributions. The variances shown are the ones used to perturb SWE and streamflow observations, as well as meteorological input.

While those distributions may or may not reflect entirely the real error of the observations and meteorological input, therefore yielding “optimistic” results, we feel they are a good compromise between the “normality” required by the EnKF and physical limits of the variables. As future works, one could investigate other distributions, introduce a bias, etc. and analyse their impact, but this falls outside the scope of the current manuscript.

The assumptions and limitations of the method are stated in the conclusion. In order to make it clearer that the experiment is near ideal, an additional sentences has been added in the revised manuscript.

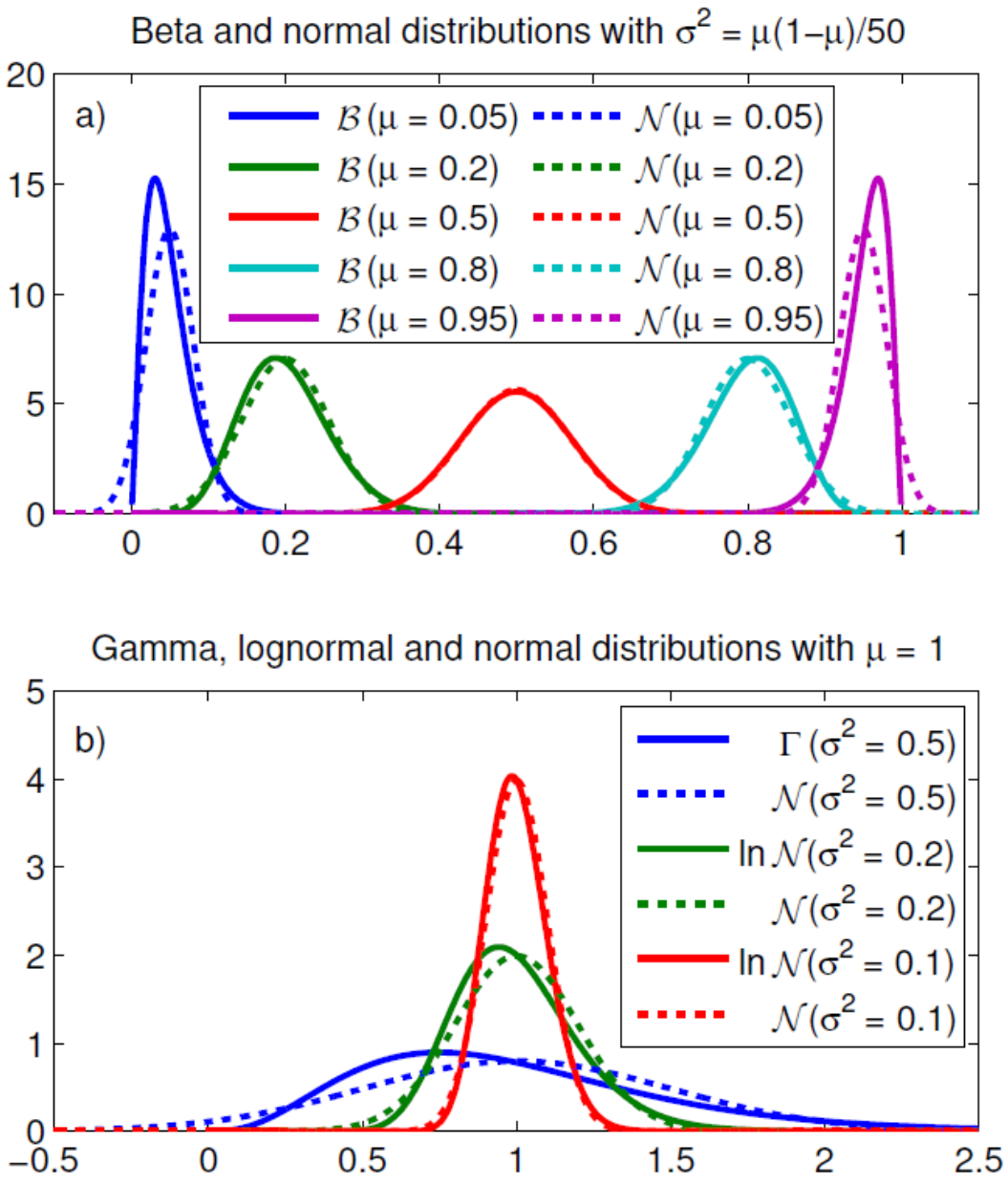

Figure 1. Examples of a) beta and b) gamma and lognormal distribution compared with their analogous normal distributions using the same mean and variance.

2) You need to provide a better justification to the use of these distributions on page 7.

AR: A paragraph has been added in the revised manuscript to justify the use of non-normal distributions. Essentially, it is to satisfy hard boundaries on observations and meteorological input. Since all observations and meteorological input in the study have limits (ex: range between [0 infinity] or [0 1]), adding a normally distributed perturbation can mean those limits are sometimes exceeded. A simple way

to get around this problem might be to set all exceeding values at the boundary (ex : all negative values set to zero). However, this introduces a bias, which is another and likely bigger problem. Other distributions were therefore used to satisfy the physical limits, while keeping some visual similarities with a normal distribution if possible.

3) You need to rewrite the paragraph starting on page 6 at line 14 as it is confusing as it would appear that it looks like you are referring to equations.

AR: As suggested by another reviewer, the word “step” has been added before each number in parentheses to avoid confusion.

4) The statement on page 17, line 30 does not make sense and is confusing about the need for linear relationships which you really should have with the EnKF.

AR: The sentence is a relic of a previous formulation and has been removed.

5) On page 9 you are finishing the details about the localization but i am concerned that because you achieve this wrt the true state that this may not be the case in the real data situation and you need some sort of disclaimer here as you are kind of using the true localization which would not be the case in reality.

AR: This was meant to be discussed in the conclusion, in the paragraph explaining the assumptions and limitations of the results. This seems like a more appropriate section in the manuscript than the experimental design since 1) the results are not known yet in the experimental design section and 2) it allows us to generalize all results as valid within the synthetic limitations of the experiment. This avoids repetition since it applies to all steps, not only the covariance localization. An additional sentence has been added in the conclusion to explicitly state that dependency.

Minor comments:

1) Page 3, line 27 remove to

AR: This has been corrected in the revised manuscript.

2) Page 5, line 17, you mention the gain yet you have not defined it.

AR: The whole paragraph has been moved to the end of the section, as well as partly merged with the (previously) last paragraph and reworked to avoid repetition.

3) Page 11, line 33, remove the first that

AR: This has been corrected in the revised manuscript.

4) Page 15, line 15, sits is not a very scientific way to describe where the site is.

AR: The expression “sits at 682 mm” does not refer to the physical location of the site, but mean annual maximum SWE as described earlier in the sentence. Nonetheless, this has been rephrased to “contains a mean annual maximum SWE of 682 mm”

5) General comment. you use both Gaussian and normal please be consistent and only use one of them

AR: This has been corrected in the revised manuscript by using “normal” throughout the manuscript.