Italicized text: Reviewer's comment

AR: Authors' response

Comments to reviewer #1

In this study, the authors explore the assimilation of discharge, SWE and SCA in a hydrologic model for the potential to improve streamflow forecasting in a mountainous basin in western Canada. Synthetic data sets are developed and used. The authors first determine which state variables are adequately predicted by the three data types that are candidate for assimilation. SCA was found to not be a good predictor. Then, the impact of assimilating SWE and discharge on hydrologic forecasts was tested.

Overall, this is an interesting study with good results. Forecasts were improved with SWE and Q assimilation both when assimilated individually and simultaneously. It is demonstrated that the data were useful for adjusting several model states (VOL, GWL, and SWE) in the CEQUEAU model. The methods in this paper show promise for applications in forecasting provided the results remain consistent for non-synthetic studies.

General comments:

1) There needs to be more detail provided in the methods section. As is, it appears as though the methods are valid, but I could not replicate this study with the information provided. I find myself having to assume I know what the authors did during some steps. Therefore, specific comments about where to add necessary detail are provided below.

AR: Information has been rearranged into new sections and additional details were added as recommended in the specific comments section. Section 3.1, dealing with the overall approach used during the synthetic experiment, now has three subsections explaining specifically how synthetic observations (3.1.1), synthetic meteorological input (3.1.2) and ESPs (3.1.3) were generated. Sections 3.2.2 and 3.2.3 now deal exclusively with meteorological ensemble and observation ensemble generation, respectively. Some information may appear to be redundant at first, but this is because similar approaches were used to generate perturbations required in the creation of single-valued synthetic observations/meteorological input and their ensemble version.

2) In the results section, the authors should make a stronger effort to link their findings to other studies. There are several papers referenced that explore assimilation of SWE and/or discharge in snow-dominated areas. There are also likely studies that have examined this type of data in other modeling and forecasting contexts. While there are a few comments about results from other studies, the authors should try to add more to the discussion.

AR: Comparisons with other studies have been added in multiple sections, including the state vector configuration when assimilating streamflow (4.1.1) and the impact of streamflow (4.2.1) and SWE (4.2.2) data assimilation on streamflow forecasts. Additional references will also be introduced for comparison purposes.

Specific comments:

Page 2, line 30. The last sentence might be better as a statement rather than a question. It seems out of character with the rest of the introduction.

AR: The sentence has been reformulated as a statement:

"The importance of state vector configuration when using multivariate DA for hydrological modeling has yet to be investigated."

Page 3, line 9: "such that the difference in elevation reaches about 1700m" is oddly stated. The difference in elevation between what? If 1700m is the total relief of the mountain, simply state it that way.

AR: The difference in elevation between the highest and lowest point in the watershed reaches about 1700m. This has been clarified in the revised manuscript.

Page 4, line 8: US Army Corps of Engineers should be capitalized. Also, is the inclusion of "1956" intended to be a citation? There is nothing listed in the references regarding this.

AR: The missing reference has been added and capitalized.

Page 4, line 17: SI is not in equation 1. How is it relevant to this discussion?

AR: SI is in the denominator of equation 1. It is one of three parameters used to convert snow water equivalent into snow cover area. It is included, along with other parameters, to be transparent in our approach and help readers understand the method used to produce the results shown in the manuscript.

Page 4, Line 18-19: As written it is implied that Hall et al. (2002) calibrated the three parameters in Equation 1. I do not think that is the case. Additional explanation of this calibration is needed. Who conducted the calibration? Was it conducted for this region? If not, is it considered to be universally applicable?

AR: The reference was initially meant to be a reference for the MODIS data, but this was not clear from the way the sentence was structured.

All the model parameters were not calibrated in the same way. CEQUEAU has numerous parameters that have been manually calibrated by engineers working for Rio Tinto, our industrial partner. These include snow, soil, evapotranspiration and transfer parameters. However, since snow cover area is not explicitly computed by CEQUEAU, a depletion curve had to be appended. The details pertaining to the depletion is not found in CEQUEAU's user manual or in any other study using or detailing how CEQUEAU works. To be transparent in our approach, we explain how the depletion curve is computed, requiring three parameters. The calibration of those parameters was conducted using the SCE-UA method (Duan et al, 1992) to minimize the root mean square difference between simulated snow cover area and MODIS data for each whole square within the Nechako watershed. The parameters are therefore not likely to be "universally applicable", but the approach may be.

The sentence has been reworked in the revised version of the manuscript to include more information about the calibration process.

Page 4, line 20: Tampered not tempered.

AR: This has been corrected in the revised manuscript for all occurrences of the term.

Page 6, line 6-8: this statement is becoming repetitive. It was mentioned several times in this section that it has been shown useful in hydrologic studies. I recommend removing earlier statements like this, or combine them into one or two sentences.

AR: The sentence has been modified to better transition into the following section.

Page 6, lines 14-24: It isn't clear what variables are referred to when using the term "observations". This may be stated earlier, but it would help the reader if they were explicitly stated here.

AR: The term "observations" used in the manuscript refers to observations to be assimilated using the EnKF, namely streamflow, SWE and SCA. The expression "meteorological input" or "weather input" is used to refer to precipitation and mean air temperature. There is one occasion in the manuscript where the expression "meteorological observation" is used and has been corrected to avoid confusion. Additional clarifications have also been added to explicitly state what are "observations" and "meteorological input".

In general, this section lacks detail. In what way and by how much were the data perturbed? How do you get synthetic observations by perturbing "true states"? More specific terminology and combining or pulling in information from Section 3.2 would be helpful in understanding the procedure of creating synthetic data.

AR: Initially, sections 3.2.2 and 3.2.3 described how both the individually perturbed and ensemble version of meteorological input and observations were generated since they both used a similar approach using the same perturbation factors. The information was added in the “Hyper-parameter tuning” section since it partly dealt with specifying errors, which is required when generating ensembles.

However, this could lead to confusion between the two sets and appeared to create a void in the description of the synthetic experiment. To avoid confusion, the information from sections 3.2.2 and 3.2.3 strictly dealing with initial perturbations of observations and meteorological input have been used to create additional subsections (3.1.1 and 3.1.2) following the overview of the synthetic experiment.

The methods section includes very little description of how ESP forecasts are generated. A more thorough explanation should be provided for readers unfamiliar with the process. Please clarify whether only 20 years of meteorological data (1990-2000) were used to generate the ESP ensembles. Also, was only the mean value of the state variable predicted by the EnKF used to generate each ensemble in the ESP forecast, or were multiple state values from the state variable ensemble used?

AR: ESPs were generated everyday over the entire study period (10 years) using a forecast horizon spanning 50 days. The multiple state values resulting from the EnKF were used to generate ESPs and is the only factor differentiating ensemble members during the forecast phase. The same true meteorological input and model parameters were during the forecast phase. No meteorological ensemble was used to generate ESPs. Although this generates forecasts which are not “realistic” since they get better over the forecast horizon, this is a way to evaluate the potential impact of data assimilation on streamflow forecasts, which is one of the two goals of the study. Using a meteorological ensemble forecast would simply add unnecessary noise. An additional section (3.1.3) has been added on ESP generation.

Page 10, line 24 onward: What is the timestep of the data evaluated? Hourly, daily, etc? I cannot find where this is explicitly stated, but it is important to understanding the results of the study. If the streamflow is evaluated at a daily timestep, it makes sense that the SWE is not beneficial for predicting streamflow; however, results might be quite different if output is evaluated at a monthly or seasonal timestep. In addition, I could not find the interval between assimilation, is it done at each model timestep, daily, weekly?

AR: The time step is daily. This has been clarified in the revised manuscript by mentioning a daily time step for the model and daily availability for the observations.

Page 14, lines 1-6 and Figure 8: It is not quite clear what VOL is in the model. As presented on page 4, it appears to be water that is being routed to the outlet (i.e. runoff). If that is the case, the quick decline in adjusting the VOL state on the CRPSS makes sense not because of a linear relationship with discharge, *but because of the likely short residence time of the water represented by VOL within the watershed. The authors discuss the residence time issue in the next paragraph with respect to GWL and SML, I would like to see similar insight regarding the VOL as the linear relationship explanation is not obvious.*

AR: The initial impact strength is due to the close link between streamflow and VOL. Since the correlation between the streamflow observation and VOL is relatively high, VOL globally experiences a positive update (as seen in figure 4). In return, since simulated streamflow depends strongly on VOL, simulated streamflow initially experiences an important positive impact. However, the duration of this impact is indeed caused by the water residence time. A short-lived impact on CRPSS as a function of forecast horizon would be caused by a relatively short residence time. This was not obvious from the way it was stated in the initial manuscript and has been clarified. Additional clarifications have also been added to section 2.2 concerning the role of VOL and streamflow.

Page 16, lines 14-16 and Figure 10: It would be helpful to add a sentence putting results from Figure 10 in context of results from assimilating only Q or only SWE.

AR: A few sentences have been added to address the comparison between the combined and individual data assimilation scenarios:

"CRPSS values for combined assimilation of both streamflow and SWE observations were superior to CRPSS values for individual assimilation of streamflow or SWE over all forecast horizons, with the exception of forecast horizons higher than 45 days, where CRPSS values for SWE DA are slightly higher. This reveals that while the updated VOL and GWL by streamflow data assimilation may be very beneficial for short-term forecasts, they do not further improve the mid-term forecasts when combined with SWE data assimilation in comparison with the scenario where only SWE data is assimilated."

Page 17, lines 28-29: The SCA was not tested on the forecasts due to the lack of improvements in state variables (page 13, lines 19-20). The authors should not make any conclusions regarding the impact on forecast skill from this study, and restate this with respect to state variable improvements.

AR: This has been corrected in the revised manuscript.

Page 18, line 4: The statements made in this paragraph are not hypotheses, they are assumptions and limitations

AR: This has been corrected in the revised manuscript.

References :

Duan, Q. Y., Sorooshian, S. and Gupta, V.: Effective and Efficient Global Optimization for Conceptual Rainfall-Runoff Models, Water Resour. Res., 28(4), 1015–1031, doi:10.1029/91wr02985, 1992.