Hydrology and
Earth System
Sciences

Discussions

Open Access

HESSD

Interactive
comment

Interactive comment on "Assimilation of citizen science data in snowpack modeling using a new snow dataset: Community Snow Observations" *by* Ryan L. Crumley et al.

Ryan L. Crumley et al.

ryanlcrumley@gmail.com

Received and published: 24 January 2021

DUE January 25th, 2021 Reviewer Comments and Responses

Reviewer #2:

The reviewer's comments are preceded by: Comment The authors responses are preceded by: Response

_____

Comment

The manuscript describes use of a new citizen science dataset (snow depth) to guide simulations of SWE via data assimilation (DA). The motivation is to include observations gathered from locations in the landscape that might not be monitored otherwise. The study is focused on a maritime snow climate of Alaska. The model used (Snow-Model) has a long history and is well established. A range of different observations (Snotel, field surveys with depth and SWE, and remotely-sensed snow depth) are used to gauge performance of the DA system, compared to simulations that do not include the depth observations.

The results presented are interesting and there appears to be considerable potential for use of the citizen science depth observations. However, major revisions could help make the manuscript more useful. The following issues should be addressed:

First, a more complete description of the DA approach is needed. In the introduction, a more detailed comparison of the approach used relative to other snow DA efforts should be provided – beyond what is currently included in the introduction (e.g., L80). How does the approach used compare to other methods, including direct insertion (e.g. Hedrick et al., 2018), particle-batch smoother (Margulis et al. 2019), particle filter (Smyth et al. 2019) and possibly EnKF.

Response:

The authors do not think this article needs to review all of the data assimilation methods used in snow science in great detail (particle filters, particle batch smoothers, Kalman filters, and ensemble Kalman filters). However, we understand the reviewer's desire to add context to the manuscript regarding other types of data assimilation methods. We propose the following paragraphs to replace the single paragraph in the methods section 3.2.4, since these new paragraphs more clearly describe the way SnowAssim works and they compare SnowAssim to other assimilation methods.

"To assimilate the CSO measurements, we used the sub-model SnowAssim developed in tandem with SnowModel (Liston and Hiemstra, 2008). The SnowAssim data

assimilation scheme is relatively simple when compared to other assimilation methods. Direct insertion methods often insert the observed state values into the modeled field in the locations and times where data is available (McGuire et al., 2006; Fletcher et al., 2012). Hedrick et al. (2018) outlines a 'modified' direct insertion method, where Airborne Snow Observatory LiDAR-based snow depth distributions are input into the iSnobal workflow in order to modify model state variables before a new initialization of the model begins. Liston and Heimstra (2008) describe a different type of modified direct insertion assimilation scheme (SnowAssim) used in the present study. Differences between the observed SWE depths and modeled SWE depths in time and location are calculated and interpolated to the entire model domain in the form of a correction surface. The final correction surface is spatially distributed (for each day of observations) using the Barnes interpolation scheme.

Note that CSO measurements are submitted as snow depths (m) and SnowAssim requires observational inputs to be SWE depths (m), so a conversion from depth to SWE was necessary. The snow depth to SWE conversion method for the current study will be discussed in the following section. Next, the model determines the dominant snow season phase (accumulation or ablation), and applies the correction factor surface to either a) the precipitation fluxes or b) the snowmelt factors during a second model simulation. Additionally, the Barnes interpolation scheme determines outliers within the observed dataset and determines the degree to which the assimilated values fit the modeled values. This determination creates a smoothed representation of the observed dataset in the assimilation results. For extensive details about the data assimilation scheme, see Liston and Heimstra (2008), their section 3, 4, and 5.

Other data assimilation methods include particle-batch smoother and particle filters. These methods are Bayesian data assimilation methods used to estimate system state variables using predicted estimates (modeled) and noisy measurement data (observed). These types of data assimilation methods rely heavily on characterizing and incorporating the predicted estimate uncertainties and measurement uncertainties into

the analysis using probability distribution functions (Magnusson et al. 2017; Margulis et al. 2015). In direct insertion or modified direct insertion methods like SnowAssim, modeled and observed state variable uncertainties are not explicitly characterized."

The authors also think that we should clarify a statement we made about assimilation in lines 108-110 of the submitted. As originally written, this sentence may be too strong of a claim to make based on the results of this paper, which was not our intent. We suggest replacing this sentence with the following, more accurate, sentence:

Previous sentence: "The CSO project adds to a growing body of research accomplished by citizen scientists in the natural sciences, and contributes to the connections between physics-based, process modeling and in-situ observations in data assimilation and snow science."

New sentence: "The CSO project adds to a growing body of research accomplished by citizen scientists in the natural sciences, and demonstrates how CSO measurements can be assimilated into the process model workflow using SnowAssim to sometimes improve model results."

Hedrick, A.R., Marks, D., Havens, S., Robertson, M., Johnson, M., Sandusky, M., Marshall, H.P., Kormos, P.R., Bormann, K.J. and Painter, T.H., 2018. Direct insertion of NASA Airborne Snow ObservatoryâĂŘderived snow depth time series into the iSnobal energy balance snow model. Water Resources Research, 54(10), pp.8045-8063.

Liston, G.E. and Hiemstra, C.A., 2008. A simple data assimilation system for complex snow distributions (SnowAssim). Journal of Hydrometeorology, 9(5), pp.989-1004.

Magnusson, J., Winstral, A., Stordal, A.S., Essery, R. and Jonas, T., 2017. Improving physically based snow simulations by assimilating snow depths using the particle filter. Water Resources Research, 53(2), pp.1125-1143.

Margulis, S.A., Girotto, M., Cortés, G. and Durand, M., 2015. A particle batch smoother approach to snow water equivalent estimation. Journal of Hydrometeorology, 16(4),

pp.1752-1772.

————————————————————————————–

Comment:

The methods section provides only a limited description of how the model is adjusted for mismatch with observations ("SnowAssim aggregates all the assimilated observations by date and creates a spatially varying cor- rection surface that covers the entire model domain (Liston and Elder, 2008). These various correction surfaces are applied by adjusting the model precipitation fluxes and snowmelt factors between SWE observation dates during a second SnowModel simulation"). The 'adjustments' to the model are central to the effort, so the method should be described more completely in the manuscript.

Response:

The authors note that the literature on SnowAssim (Liston and Heimstra, 2008) is cited in the methods section 3.2.4 and we make efforts to not repeat information from that publication. First, see Figure RC1 (Figure 6a from Liston and Heimstra, 2008) included in our response below for the reviewer's ease, as an example of a correction surface referenced above. We do not include an example of the correction surface in the manuscript because it is explained in the original literature. However, see the previous answer for additional information that will be added to the manuscript's method section 3.2.4.

————————————————————————————–

Comment:

The results (or discussion) do not include any documentation of the 'adjustments' to the model, yet one of the benefits of DA is that the merging of data and models is one way to more completely understand the entire system (e.g., see Magnusson et al., 2014 and 2017 and their retrieved precipitation correction factor).

Response:

The authors agree that including some additional information from the correction factor adjustments during assimilation would benefit the arguments we make in the manuscript and elucidate the entire modeling/assimilation system. See Table RC1, which includes data from the best ranked assimilation runs using the time-series and spatial analysis. We plan to add the following paragraph to a new section (6.6) in the results.

" 6.6 Correction Factor Results SnowAssim generates a set of correction factors for each of the CSO ensemble member simulations. These factors correspond to the observed and measured differences in the SWE variable and are used to create a correction surface with the Barnes objective analysis. Table [RC1] reviews a subset of the correction factors, including data from the Best ranked CSO simulations according to the various temporal and spatial metrics previously reviewed in sections 6.1 and 6.2. The number of observations varies for the Best ranked simulation, as well as the precipitation correction factors, the use of a melt correction factor, and whether or not an interpolated correction surface was created. These correction factor results show that relatively few measurements are needed during assimilation and that there are multiple paths to improving model performance when assimilating CSO observations using SnowAssim."

_____

Comment:

Second, uncertainty of the observations and validation data should be described and incorporated into the analysis. One of the benefits of DA is that the magnitude of uncertainty can be explicitly included in the analysis (e.g., Magnusson et al., 2014 and 2017). It appears that uncertainty of the assimilated observations is not included in the analysis – is this the case? If not, why not?

The spatial representativeness of the depth measurements is mentioned in the discussion. One component of uncertainty is related to the conversion from depth to SWE, using the density estimation described in Hill (2019). In the region analyzed, SWE estimates based on density from Hill (2019) have an RMSE of 0.2-0.25 (normalized to snow season precipitation). Is this considered in the DA approach? Uncertainty (or biases) of the validation data is not described, thus it is implied that the data are 'perfect'. What is the error or uncertainty associated with the federal sampler data?

Response:

The authors think that the reviewer brings up an important point about the need for further uncertainty analysis in the manuscript. We note that SnowModel is a deterministic model and SnowAssim does not include an explicit characterization of uncertainties, so any perturbations need to be based on fieldwork measurements or other previously reported error values. We took this opportunity to characterize multiple sources of uncertainty mentioned by the reviewer in Table RC2.

One of the sources of uncertainty mentioned by the reviewer comes from the conversion of snow depth to SWE using the Hill et al. (2019) method. We can quantify this source of error using the reported values from Hill et al. (2019) or using the fieldwork measurements of co-located snow depth and SWE (see Table RC2 above). Another source of uncertainty is the spatial representativeness of the depth measurements across short distances. The average standard deviation of snow depth at all 70 fieldwork sites is 22 cm of snow depth, and after conversion to SWE, the average is 8.7 cm of SWE. We can assume that the spatial variability of snow depth plays a role in the conversion method uncertainty, so we decided to choose one of these values and perturb SWE measurements used in the data assimilation scheme. In an attempt to be conservative with our error estimates, we chose the highest reported/measured error value of 10.5 cm to create an envelope of uncertainty around our SWE values reported in the assimilation runs at the Upper Tsaina SNOTEL station. Figure RC2 contains the results of these additional model runs. This uncertainty analysis displays improved

simulation results, even after the error estimation from these sources of uncertainty have been taken into consideration. This figure further contextualizes the temporal results in section 6.1 and we suggest adding it to the manuscript with the following sentences.

"Using the snow depth to SWE conversion method during assimilation introduces uncertainty into the modeling process. Instead of using the global estimates of error reported in Hill et al. (2019; RMSE in SWE = 5.9 cm) we decided to calculate this source of error using our fieldwork site measurements. The RMSE in SWE due to the conversion method is 10.5 cm and we perturbed the CSO observations by this amount to depict the upper and lower boundaries of error associated with this source of uncertainty. Figure [RC2] displays the Best CSO simulation temporal results for each WY, along with the UTS station SWE record and the NoAssim case. These perturbations to the assimilated SWE show improved modeled SWE values at the UTS station when compared to the NoAssim case, even after this source of uncertainty has been accounted for."

Since there is additional uncertainty associated with the federal sampler data, we decided to add the following sentence to the methods section 3.2.3:

"Federal sampler data collection introduces uncertainty in the form of measurement error due to variable snow conditions and densities, hard impenetrable crusts, and loss during extraction. Dixon and Boon (2012) report the results of several studies showing that the Federal Sampler error, as a percentage of SWE depth, ranges from 4.6% to 11.2%. Our results presented in Section 6.4 include field measurements of SWE that use the higher 11.2% value for conservative SWE error estimation."

Because of this uncertainty we modified Table 4 in line 510 (Section 6.4) of the submitted manuscript to include a range of estimated densities and SWE (+/- 11.2%) instead of using only the measured values without the error estimation. See Table RC3 for the new values that will be added to the original Table 4 in red.

Lastly, we know there is uncertainty associated with the remote sensing acquisitions, and the sources of this error include flight trajectory and geometry, laser scan angle, density of canopy, steep gradients in the terrain, and more (Deems and Painter, 2006). We decided to report the estimated uncertainty for the RS datasets in the methods section 3.4.2 and also include them in the new version of Table 4. The following sentences will be added to the methods section 3.4.2:

"There is uncertainty associated with the RS dataset acquisitions, and the sources of error are related to flight trajectory and geometry, laser scan angle, density of vegetation and canopy, and steep gradients in the terrain (Deems and Painter, 2006). The mean error in snow depth for the photogrammetry and LiDAR datasets are estimated at 10.4 cm and 1.1 cm, respectively. The vertical RMSE in snow depth are estimated at 31.0 cm for the photogrammetry and 10.2 cm for the LiDAR dataset. While we acknowledge and report these error estimations, they are integrated into the results in Table 4 in Section 6.4 but not used in the spatial results reported in Section 6.2."

Dixon, D. and Boon, S., 2012. Comparison of the SnowHydro snow sampler with existing snow tube designs. Hydrological Processes, 26(17), pp.2555-2562.

Deems, J.S. and Painter, T.H., 2006, October. Lidar measurement of snow depth: accuracy and error sources. In Proceedings of the 2006 International Snow Science Workshop: Telluride, Colorado, USA, International Snow Science Workshop (pp. 330-338).

_____

Comment:

Third, something seems strange about the calibration and validation methods and results. Are the NSE values in Table 1 correct? If the best simulation has NSE < 0, this would suggest that the calibration is not working very well. Additional details are required.

Response:

The values in Table 1 in the original manuscript are correct and they were the impetus for including the precipitation adjustment experiment in the manuscript. The initial calibration model runs and the final No Assimilation model runs were displaying time-series SWE values that were consistently high, throughout both water years, with both reanalysis products (see original Figure 5a and 5d). We know that biases in meteorological forcings are one of the most important factors in estimating snow depth and SWE magnitudes correctly (Liston and Heimstra, 2008; Margulis et al., 2015). So we decided to take a closer look at the precipitation totals with the CFSv2 product. See Figure RC3 that shows the total amount of precipitation over the calibration period when compared to the Upper Tsaina Snotel station and compared to when a precipitation adjustment factor is used. The authors would like to add this figure to the appendix to clarify the need for precipitation adjustments, whether via data assimilation or the precipitation adjustment experiment.

Because of this bias in the meteorological inputs, and after a conversation with the model developer about the calibration challenges in this region of Alaska, the authors were confident that making adjustments to the model parameters would only slightly improve our snow depth and SWE distributions and magnitudes (see Appendix A for a full list of the model parameter adjustments made during calibration). The improvements that could be made by adjusting model parameters were insignificant when compared to adjusting the precipitation fields.

Importantly, the reviewer's question speaks directly to why we think SnowAssim is the correct assimilation method for this research. SnowAssim adjusts the precipitation fluxes and/or snowmelt factors using only the additional observations provided by CSO participants. Nothing else is changed and no additional information is required for this type of data assimilation. Recall that we are not forcing the model with in-situ weather station data because the required meteorological variables are not available within the domain. Even with biased and coarse reanalysis forcing data, SnowModel

and SnowAssim are able to make snow depth and SWE magnitude improvements by the simple addition of several in-situ snow depth observations, strengthening our key claim that "that even modest measurement efforts by citizen scientists have the potential to improve efforts to model snowpack processes in high mountain environments."

Liston, G.E. and Hiemstra, C.A., 2008. A simple data assimilation system for complex snow distributions (SnowAssim). Journal of Hydrometeorology, 9(5), pp.989-1004.

Margulis, S.A., Girotto, M., Cortés, G. and Durand, M., 2015. A particle batch smoother approach to snow water equivalent estimation. Journal of Hydrometeorology, 16(4), pp.1752-1772.

––

Comment:

Is calibration for the entire year? The entire snow year? Why not at peak SWE?

Response:

As mentioned in lines 308 to 309 in the submitted manuscript, the calibration time period is for the entire water year, for 5 years. The calibration statistics cover the entire 5 year period. Two months (July and August) of data per water year, in which no snow was modeled, measured, or expected at the Upper Tsaina SNOTEL station in the domain, were removed from the calibration metrics. This was in an attempt to not bias the results of the RMSE and mean bias error metrics with months of corresponding zeros from the observed and modeled vectors.

––

Comment:

Results in Fig 5 also seem strange. Fig 5e: how can this be the 'best' simulation? There is a clear problem during the ablation period; is it really a "best" simulation if ablation is too rapid? If stats are calculated throughout the season, and ablation season is

short, it is easy to discount the errors during this time of year. But doesn't timing of snow disappearance matter? Perhaps a metric of snow disappearance date should be included? One could argue the result in 5f is much worse than 5d, so that assimilation is not improving the simulations, but actually making it worse.

Response:

Figure 5e represents the data from the Best ranked simulation according to the time-series data from water year 2018. The decision-making for characterizing the Best ranking is explained in lines 381 through 383 and also qualified in the discussion section, in lines 536 through 542 of the submitted manuscript. Characterizing and focusing on the Best results for some figures in the manuscript was a decision made by the authors to show a sampling of the many model runs we conducted during the analysis, instead of overwhelming the reader with too many figures. We acknowledge that if we focused on just the accumulation phase or just the ablation phase, the characterization of the Best results would indeed look different. The ablation phase of the snowpack in Figure 5e is not a perfect match to the 1:1 identity line, but the corresponding metrics (NSE, KGE, RMSE, Bias) all show improvements when compared to the No Assimilation case when averaging over the water year. The new SWE figure that includes error perturbations shows that the timing of the last day of SWE (snow disappearance date) in the Best CSO simulation in WY2018 is 6 days earlier than the SNOTEL snow disappearance date. The range of snow disappearance dates when accounting for some level of measurement and conversion model uncertainty is from 10 to 1 day(s) early. The NoAssim snow disappearance date is 7 days later than SNOTEL. The WY2017 snow disappearance dates are even better.

The claims that we make in the results and discussion section are specific to the entire water year and we are careful to not make any claims about improving the snow disappearance dates or the timing of the melt period. However, after doing additional uncertainty analysis as suggested by the reviewer, we are more confident that our overall snow disappearance dates are acceptable. The authors also note that the mag-

nitude of peak SWE is greatly improved in our best model runs when compared to the NoAssim case, and this may be more important for readers concerned with the water resources implications of our work. While we acknowledge that there are alternate ways to subset the data temporally, the authors stand by our decision to use water year averaged metrics to characterize the Best ranked simulation.

Lastly, since the reviewer requested, we plan to add the following sentences to the new paragraph accompanying the new SWE figure in section 6.1 in order to report the Best CSO simulation snow disappearance dates:

"Since the timing of snow disappearance is important for ecological systems in alpine environments and water resources managers, we calculated the range in snow disappearance dates from the best simulations from both water years. In WY2017 and WY2018, the snow disappearance date for the NoAssim case is 10 and 7 days later than the UTS station, respectively. In WY2017, the snow disappearance date in the Best CSO simulation, accounting for measurement uncertainty, ranges from 3 days early to 8 days later than the UTS station. In WY2018, the range is from 10 days to 1 day earlier than the UTS station. These ranges in snow disappearance date are acceptable and show improvements in model performance for some, but not all, of the Best CSO simulations after accounting for measurement uncertainty."

———————————————————

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-321, 2020.
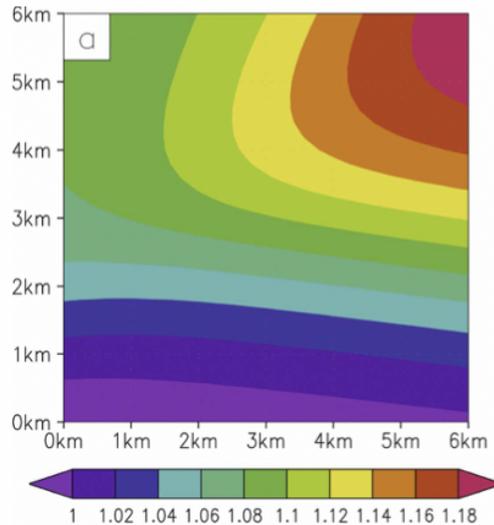
Figure RC1: This is Figure 6a from Liston and Heimstra (2008) showing an example of a spatially distributed correction factor surface created by the SnowAssim data assimilation scheme. These values modify the 1) precipitation inputs during the accumulation phase or 2) the melt rates during the ablation phase. We do not intend to add this figure to the manuscript.
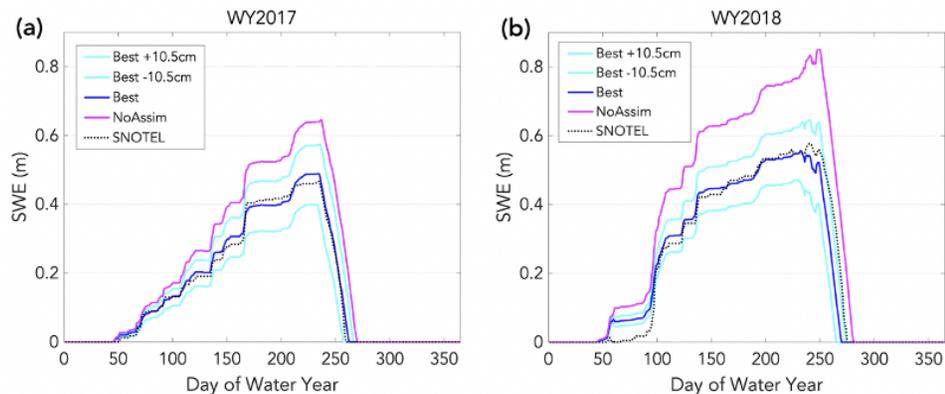
Fig. 1. Figure RC1

Figure RC2: Snow water equivalent (SWE) time series results with measurement uncertainty included. The simulations with ±10.5 cm of SWE represent the upper and lower boundaries of error introduced when converting snow depth measurements to SWE using the Hill et al. (2019) method.
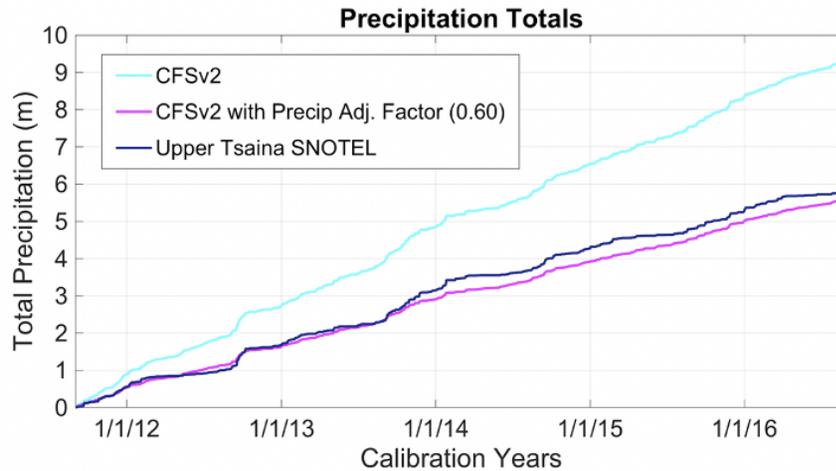
Fig. 2. Figure RC2

Figure RC3: Precipitation totals at the Upper Tsaina SNOTEL station compared to the CFSv2-forced model totals and the CFSv2-forced model totals with a precipitation adjustment factor. This overestimation of precipitation by the reanalysis product is a major factor in the quality of the calibration results.

Fig. 3. Figure RC3

Table RC1: Correction factors from the assimilation scheme for the best ranked simulations from both water years. The model determination for precipitation vs melt correction factors is included and whether or not the Barnes objective analysis created a spatially distributed correction surface.

Type	Ranking	Year	# of Obs	Precip Correction Factors	Melt Correction Factors (-)	Interpolated Surface?	Dates
Temporal	Best	2017	2	0.45, 1.04	n/a	Yes	4/29/17
Temporals	Best	2018	2	0.68, 0.76	n/a	Yes	5/15/18
Spatial	Best	2017	8	0.30, 0.50, 0.73, 0.86, 1.36	6.32, 2.29, 22.6	Yes	4/29/17; 5/8/17
Spatial	Best	2018	1	0.32	n/a	No	5/22/18

Fig. 4. Table RC1

Table RC2: Sources of uncertainty, calculated or reported error, and metric used for each dataset. We do not intend to add this table to the manuscript, it is included for the reviewer only. All of the data will be included in various locations within the final, edited manuscript.

Source of Uncertainty	Error	Metric
Conversion Method (reported)	5.9 cm	RMSE
Conversion Method (measurements)	10.5 cm	RMSE
Variability of Snow Depths (measurements)	8.7 cm	Std. Dev. Ave.
Federal Sampler Measurement Error (reported)	11.2%	% Error
2017 Photogrammetry RS Dataset (measurements)	10.4 cm, 31.0 cm	Mean Error, RMSE
2018 LiDAR RS Dataset (measurements)	1.1 cm, 10.2 cm	Mean Error, RMSE

Fig. 5. Table RC2

Table RC3: Spatially Averaged Variables in the RS Region
The spatially averaged results were calculated using the RS region in WY2018, the RS dataset (±1cm error), the spatially averaged density, and the modeled results. The spatially averaged SWE depth for the RS survey was estimated using the average density (± 11.2%) measured during April 2018 fieldwork.

Dataset	Spatially Averaged Snow Depth (cm)	Spatially Averaged Density (kg/m^3)	Spatially Averaged SWE Depth (cm)	Total RS Region Water Volume (km^3)
RS Survey 2018	130 ±1 (RS survey)	331 ± 37 (fieldwork)	38 - 48 (estimated)	0.06 - 0.07 (estimated)
Best CSO Simulation 2018	130 (modeled)	400 (modeled)	52 (modeled)	0.08 (modeled)
NoAssim 2018	267 (modeled)	430 (modeled)	115 (modeled)	0.17 (modeled)

Fig. 6. Table RC3