The Cryosphere Discuss., 8, C1985–C1993, 2014
www.the-cryosphere-discuss.net/8/C1985/2014/

The Cryosphere
Discussions
Open Access

Interactive comment on "Independent evaluation of the SNODAS snow depth product using regional scale LiDAR-derived measurements" *by* A. Hedrick et al.

A. Hedrick et al.

andrewhedrick@boisestate.edu

Received and published: 7 October 2014

General comments

1. *A comparison of SNOTEL and SNODAS is lacking which will greatly improve the context of the results. Because SNOTEL roughly governs SNODAS (via data assimilation), this context will provide substance for extended discussion.*

   Response: This was an avenue that we considered early on, but abandoned it for two reasons. 1) SNOTEL is only a point measurement, and as such the assimilation of SNOTEL data into the model are only approximations for a larger

   area. SNODAS estimates are considered to be approximations of the average conditions within a 1 km grid cell. 2) We are simply comparing two snapshots of the snow conditions over a large area. Since the LiDAR flights only covered two SNOTEL sites (Columbine and Rabbit Ears), studying the SNOTEL vs. SNODAS time series would give context only at two wind-sheltered points within two pixels of the survey swath.

   The "run-time modifications", or MODS, are based in science but require human interpretations of data that is very difficult to replicate. Due to their relative sparseness, SNOTEL sites are often merely used to spatially align storm tracks and other regional-scale processes. There are interpolation components to the data assimilation methodology, but many other components would disrupt the "error increases with distance from a SNOTEL" assumption. Model operators may notice evolving discrepancies between SNOTEL observations and grid cell values and are often forced to make bulk corrections to repair multiple compounding model errors, without knowing the basis behind the errors. It is evident that the data assimilation process for SNODAS is closer to an "art form" based in science, which would be very difficult to quantify.

   However, that is not to say that an analysis of SNOTEL and SNODAS would be inconsequential. Further work examining a time series of SNOTEL and SNODAS could shed some light on how well the model is distributing the assimilated data over space. We did look at SNOTEL depths and swe between the LiDAR flights and compared them to the SNODAS output over that time period, and we will consider adding another figure depicting the time series of snow depths for the two overlapped SNOTEL sites and their corresponding SNODAS pixels. This will be at the editor's discretion.

2. *A discussion should more generally frame the results in terms of potential water yield. More importantly, what are the next steps for improving SNODAS using future LiDAR data sets similar to those in this paper? What aspect of the LiDAR*

acquisition will be key to get right next time? Are the LiDAR errors actually small enough for a comprehensive validation? This is important to consider because LiDAR snow depth is perhaps the best opportunity to understand and improve the SNODAS products or other, similar model estimates at large spatial scales. Could the LiDAR be assimilated? How much ground truth would be necessary to properly bias correct the LiDAR? Etc.

Response: A main goal of this paper is to foster more discussion about how to effectively use LiDAR snow depth campaigns for model calibration and validation. Even so, we need to be more vigilant about data quality when collecting the data in the first place, and that quality can change drastically depending on who/what is collecting the data. The question "Are the LiDAR errors actually small enough for a comprehensive validation?" is a very important one that we have struggled with over the entire course of this work. Though not collected to specifically validate the LiDAR, the HG in situ surveys were a boon to the study and really added another dimension to the LiDAR, and it was our judgment that the RMSD between the LiDAR and HG snow depth change was a reasonable uncertainty estimate. Additionally, if we had noticed no trends whatsoever in the comparison of Δ*LiDAR* vs. Δ*SNODAS* (Figure 8), we likely could not have continued. However, by quantifying the error for this particular survey and defining regions where Δ*SNODAS* disagreed with Δ*LiDAR*, we felt that this study could be a first step in really nailing down some of the issues facing large-scale energy balance models. All of this speaks to the importance of ground truthing remote sensing data upon acquisition.

Specific comments

1. *Given the analysis in the paper, my lingering question is "why is SNODAS wrong?" Or, why is SNODAS right?*

Response: After performing this comparison, we were pleasantly surprised at just how well SNODAS was able to predict snow depth change given that the primary goal of the model framework is to predict SWE. The estimation of depth heavily relies on physically based assumptions about the density of new falling snow, and how that snow evolves and compacts over time. With that being said, the three regions containing the largest model/observation discrepancies are definitely the traditional 'problem areas' for snow models. SNODAS is such a complex model framework that it is likely a combination of many small assumptions in the physics of the modeled processes that are causing particular physiographic locations to overestimate depth and others to underestimate. Ideally, higher temporally resolved LiDAR flights at this scale and extent would allow for more of a concise evaluation of the shortcomings of SNODAS.

You are likely correct that SNODAS performs more poorly the further from an assimilation data point, and should be an entire publication in itself. In fact, Region 2 shows a stark contrast between Δ*LiDAR* and Δ*SNODAS* as the elevation decreases to the east of the Columbine SNOTEL. Nevertheless, as of yet we can merely show the locations of the discrepancy and do some cursory hand waiving as to the causes. We will include more discussion on this matter in the final manuscript.

(Comment cont.) *It will be impossible to answer this comprehensively because we don't know the exact assumptions in SNODAS. To address comment 1 of reviewer 1: yes, there are "MODS" in SNODAS (at least this is generally believed). This is (still) fairly standard practice for operational products (e.g. Seo et al, 2009). New validation products, such as presented in this paper, will hopefully lead to comparisons of MODS assimilations and automated assimilation procedures and advance the science. The upshot of the MODS is ambiguity in how to improve the results. This will make for challenging speculation in the discussion. However, efforts along these lines could be a significant benefit to the community* *and help push the science forward.*

My thoughts on the initial question:

(a) *SNODAS is going to be correct owing to SNOTEL observations*

(b) *It's going to be wrong moving away from SNOTEL in space as (MODS) assumptions about variability break down.*

To me this explains why SNODAS is not simply biased, but the line of best fit intersects the 1:1 line. I'd guess that the intersection is roughly near the magnitude SNOTEL observations. That's not going to be exactly true, but makes a reasonable story. I think this general idea is sketched on P3154 L3, but it deserves clarification and expansion along with the relationship of the SNOTEL observations to the results. There should be speculation about why the assumptions moving away from SNOTEL are likely wrong and how we might fix that. Wind is mentioned in passing. How about vegetation? Other differences in physiography with the assimilated SNOTEL observations?

Response: We would certainly like to examine the vegetation interaction with SNODAS, and elevation and terrain roughness are likely important factors as well. These physiographic influences would need to be studied on a smaller localized scale, such as only focusing on the Rabbit Ears region, or only the Medicine Bow foothills, for instance. From the regression analysis, the 1-km^2 bulk vegetation density, average elevation, IQR vegetation (basically canopy x-y density), and IQR elevation (basically surface roughness) all had a minimal effect on model-measurement discrepancies over the entire survey area. As a researcher, I am sure that these all of these variables affect SNODAS performance, but the effects were averaged out over the large spatial extent.

(Comment cont.) *The entire regression analysis centered on page 3154 strongly suggests to me that the SNODAS assimilation/MODS (which are the errors away from SNOTEL) are not based on any of these explanatory variables which we*

commonly expect to govern snow depth. This may or may not be true, but it appears that such predictors are not the basis of the MODS. I currently know of now snow depth/SWE assimilation technique that actually uses such variables. So, it's not really surprising.

I know there are already multiple analyses of different products in this paper, but a comparison is needed of the SNOTEL and SNODAS used in the study. This will help to illuminate the above points. For the two SNOTEL locations in the LiDAR footprint, I would suggest also including the LiDAR spread and mean information for the SNODAS pixel and also for a, say, 10-15m radius centered on the SNOTEL.

Response: The SNOTEL influence on the data assimilation was addressed in #1 above.

2. *Also hindering interpretation of the results is that SNOTEL information is somewhat hard to see where it does exist in the paper. While it's nice to see the SNOTEL positions overlaid, it makes it difficult to interpret the colored values in figs 2, 4, and 9 at the location of the SNOTEL, which is of acute interest. This is challenging to fix and make obvious, and is part of the reason I suggest treating these comparisons in a separate figure. Improving the readability of this figure is needed to help with interpretation in the spatial context. I don't have a great suggestion for how best to do this. Smaller symbols would help fix this, but be more difficult to see. Perhaps empty squares or diamonds centered on SNOTEL? Similarly the HG symbols can block the information that they overlay and make interpretation of the underlying values difficult. Figure 1 is very nice. My concern, again, is about promoting interpretability. Showing the ?Lidar in the figure detracts from our ability to use topography in that region as context for interpretation of results. I suggest removing it here and combining it with Figs 2 and 4, in this order 1, 4, 2 as panels of a single figure. I think being able to compare △LiDAR, △SNODAS, and △SNODAS−△LiDAR in the same figure is important*

to the interpretation. Flipping pages detracts from the interpretation. I'd also include SNODAS values in another panel in the same figure. This will help promote a coherent discussion where these things are easily compared.

Response: Yes, the figures will all be adjusted to ease the interpretation, based on these recommendations.

3. *Also, would a vegetation figure (e.g. NLCD or MODIS) contribute to interpretation?*

   Response: The regression analysis over the entire survey area showed the LiDAR measured vegetation to have a negligible effect on the SNODAS-LiDAR difference. However, vegetation likely has a strong effect on the difference at smaller localized regions within the study area. A new analysis on vegetation effects on SNODAS could be its own paper, but gets back to the issue of localized causes of variability. Also, this particular LiDAR survey had a relatively low point density, which would not allow for fine-scale analyses of vegetation interactions. The data included in our regression analysis was merely a first cut at trying to tease out correlations from our limited dataset, and a more thorough analysis would require the higher resolution information that future LiDAR acquisitions should be able to provide.

4. *As mentioned above, the results (the difference between LiDAR and SNODAS changes) should be put into the context of difference in potential water yield or potential energy balance effects. I'm talking about back of the envelope calculations with simple assumptions. Other suggestions for discussion questions were also offered above.*

   Horizontal error bars on figures 5 and 6 could place SNODAS more broadly within the measurement context.

   Response: Agreed. We added error bars to the plots for both the range and 1st standard deviation of measurements within the pixels. However, since the spatial

variability was so high in the deeper sites, those error bars made interpretation more difficult so we decided not to include them.

5. *(Same as reviewer 1, comment 7; P3153 L18) It nags at my conscience that you're using* ± 13*cm from RMSD as the error range for the LiDAR. I think a simple discussion (1 sentence?) justifying why this is appropriate would be helpful. My concern is that the errors are biased so that they are not symmetric about zero. If you plot the distribution of these errors, the mean is not zero. The assumption of* ± 13*cm is similar to assuming 1 standard deviation of a mean-zero distribution? Also why is 1 standard deviation, or whatever exactly RMSD represents, appropriate? It's not the same as 1 standard deviation if there's a bias. The conclusions are somewhat dependent on this assumption, so it should be clearly argued.*

   Response: We don't believe that the sample size of just 12 in situ points where the LiDAR was analyzed provides enough spatial representation of the 750-km^2 study area to apply a bias correction. With a bias on the order of the LiDAR uncertainty, we did not feel that the bias correction would largely affect the Δ*SNODAS* vs. Δ*LiDAR* comparison results. This was also addressed within the Author Comments to Reviewer #1.

6. *(Same as reviewer 1, comment 3; P3150 L20) The argument about why melt being insignificant is lost on me. Why is this important? Clarification needed. Related to this, a time series "spaghetti-plot" of all the SNODAS pixels along with their mean would illuminate SNODAS behavior during the* Δ*time.*

   Response: This analysis was intended to simply narrow down the processes that could be causing uncertainties in SNODAS. By eliminating melt as a contributing factor, the focus turns to the densification, compaction, and new snow density routines within the NSM portion of SNODAS.

7. *(Same as reviewer 1, comment 5; P3150 L22) Do you mean just change in depth (what I'm calling* Δ*) instead of melt? Sublimation doesn't really cause melt; it*

probably has the opposite effect like sweat cools the body.

Response: Yes these are two completely separate processes, and we removed any mention of sublimation from the text, since we did not perform an analysis on the sublimation products. SNODAS treats processes affecting mass loss individually within the energy balance model, and melt and sublimation are modeled as individual products: sublimation due to wind, sublimation within the snowpack, and melt due to solar radiation.

8. *(P3152 L26) "...mean HG" wasn't defined as "mean HG difference" previously. I assume that's what you mean. Generally I'd suggest revising the notation to use deltas, it would be clearer: $\triangle$LiDAR, $\triangle$SNODAS, $\triangle$HG.*

   Response: This suggestion is great. We will go back and change all the surface difference variables to include deltas.

9. *(P3153 L21) Seems like bias and RMSE should be mentioned in this paragraph. It's on the figure and important.*

   Response: We will add some discussion to address the bias issue to this paragraph.

Interactive comment on The Cryosphere Discuss., 8, 3141, 2014.