Manuscript prepared for The Cryosphere
with version 5.0 of the LaTeX class copernicus.cls.
Date: 23 October 2014

Independent evaluation of the SNODAS snow depth product using regional scale LiDAR-derived measurements

Andrew Hedrick[1,2], Hans-Peter Marshall[1], Adam Winstral[2], Kelly Elder[3], Simon Yueh[4], and Donald Cline[5]

[1]Boise State University, Center for the Geophysical Investigation of the Shallow Subsurface, Boise, Idaho 83725, USA
[2]USDA-ARS Northwest Watershed Research Center, 800 Park Blvd., Suite 105, Boise, Idaho 83712, USA
[3]USDA Forest Service, Rocky Mountain Research Station, Fort Collins, Colorado 80526, USA
[4]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91125, USA
[5]NOAA-NWS, Hydrology Laboratory, Office of Hydrologic Development, Silver Spring, MD 20910, USA

Correspondence to: Andrew Hedrick
(andrewhedrick@boisestate.edu)

Abstract. Repeated Light Detection and Ranging (LiDAR) surveys are quickly becoming the de facto method for measuring spatial variability of montane snowpacks at high resolution. This study examines the potential of a 750 km^2 LiDAR-derived dataset of snow depths, collected during the 2007 northern Colorado Cold Lands Processes Experiment (CLPX-2), as a validation source for an operational hydrologic snow model. The SNOw Data Assimilation System (SNODAS) model framework, operated by the U.S. National Weather Service, combines a physically-based energy-and-mass-balance snow model with satellite, airborne and automated ground-based observations to provide daily estimates of snowpack properties at nominally 1-km resolution over the coterminous United States. Independent validation data is scarce due to the assimilating nature of SNODAS, compelling the need for an independent validation dataset with substantial geographic coverage.

Within twelve distinctive 500×500 m study areas located throughout the survey swath, ground crews performed approximately 600 manual snow depth measurements during each of the CLPX-2 LiDAR acquisitions. This supplied a dataset for constraining the uncertainty of upscaled LiDAR estimates of snow depth at the 1-km SNODAS resolution, resulting in a root-mean-square difference of 13 centimeters. Upscaled LiDAR snow depths were then compared to the SNODAS estimates over the entire study area for the dates of the LiDAR flights. The remotely-sensed snow depths provided a more spatially continuous comparison dataset and agreed more closely to the model estimates than that of the *in situ* measurements alone. Finally, the results revealed three distinct areas where the differences between LiDAR observations and SNODAS estimates were most drastic, providing insight into the causal influences of natural processes on model uncertainty.

1 Introduction

Meltwater from mountain snowpacks is an important component of Earth's water cycle. However, quantifying the amount of water stored in a snowpack from year to year remains difficult. Millions of people in the western United States rely on water that descends from the Rocky Mountains, where over 70% of the annual water supply is delivered from melting snow (Carroll et al., 2006). With the worldwide population growing exponentially, the importance of fine tuning our current hydrologic models is becoming more of a priority in order to mitigate flood disasters and water shortages.

The primary goal of most hydrologic snow models is to provide estimates of snow water equivalent, or SWE, over large mountain regions, but in addition most models include routines to estimate secondary snow properties. The methods used to estimate snowpack characteristics such as depth and density vary between models; some use empirical methods from available historical data, while others are more physics-based. Even so, SWE is but a function of depth and density, and if validation is achieved for either of these so-called secondary model components, then higher confidence can be placed into corresponding SWE estimates.

Since snow depth varies considerably more than bulk density over space (Sturm et al., 2010) and is also inherently easier to measure, this study purports to examine the snow depth prediction component of a gridded, spatially-distributed snow model. Specifically, we will demonstrate the value of repeated large-scale airborne LiDAR surveys, that have been subjected to ground validation error analysis, for assessing the ability of an operational physically-based snow model to estimate snow depths over a vast extent.

Due to the resolution capability and gridded nature of most distributed snow models, many small-scale features that affect spatial variability are either averaged or not considered, influencing the bulk estimates of total SWE and overall depth in each grid cell (Marchand and Killingtveit, 2005). Nevertheless, sub-grid spatial properties have been shown to have a significant effect on the accuracy of spatially distributed snow models (Luce et al., 1999; Liston, 2004; Skaugen and Randen, 2013), but the datasets required for parameter estimation and optimization are small and spatially sparse in high-elevation, tundra and shrubland environments (Elder et al., 1991; Sturm et al., 2001a,b; Hiemstra et al., 2002; Liston and Sturm, 2002; Schirmer and Lehning, 2011). Considerable variability in the spatial snow distribution can be introduced through the interaction between wind and snow with terrain and vegetation (Elder et al., 1991; Blöschl, 1999; Liston et al., 2007). In fact, wind has been shown to be the dominant influence on spatial variability of snow in complex terrain (Pomeroy et al., 1993; Winstral et al., 2002; Sturm and Wagner, 2010). Without prior knowledge of the spatial snow distribution in a given area, arbitrary manual snow measurements will not provide accurate estimates of snow depth over large alpine regions (Elder et al., 1991; Anderton et al., 2004; Erickson et al., 2005).

Various studies have shown that LiDAR (Light Detection and Ranging) surveys can provide spatial information on mountain snow depths at high-resolution over large areal extents that comprise various physiographic regimes (Hopkinson et al., 2004; Deems et al., 2006; McCreight et al., 2014). The first Cold Lands Processes Experiment (CLPX-1) of 2002-2003 in the Colorado Rocky Mountains, was the first large-scale coordinated study to use LiDAR acquisitions for the assessment of snow properties over a range of areas (Cline et al., 2009). Since then, numerous campaigns have used LiDAR to quantify spatial variability of snow depths in mountain terrain. Deems et al. (2006) used fractal analysis of the CLPX-1 LiDAR snow depths to determine scale-breaks, while Trujillo et al. (2007) found that spatial distributions of snow depth are strongly controlled by both wind redistribution and vegetation interception of snow over uneven surface topography in five of the CLPX-1 intensive study areas. More recently, LiDAR has been used with simple statistical models to determine scale invariance due to vegetation and wind direction (Trujillo et al., 2009) as well as to verify high-resolution dynamical snow models (Mott et al., 2011).

As the technology has become more widespread over the last decade, and as LiDAR for snow research has become increasingly relevant, more effort has been placed into increasing the measurement extent of LiDAR footprints. The advantages of LiDAR for spatially characterizing snow depths over large remote areas are finally being used to assess lower resolution operational hydro-meteorologic snow models. Melvold and Skaugen (2013) used six parallel 500-meter$\times$80-kilometer LiDAR surveys, each separated by 10 km, to investigate the Norwegian operational temperature index snow model, *seNorge*. After upscaling the LiDAR-derived 2-m resolution snow depths to the spatial resolution of the 1-km^2 gridded model output, the modeled results were found to accurately represent the remote sensing estimates despite the lack of sub-grid spatial information within the model structure. A similar approach for LiDAR upscaling is used in this study.

Even though depths can vary greatly over space in a snow pack, the overall distribution of snow has been found to exhibit spatial similarities from year to year (Hiemstra et al., 2006; Sturm and Wagner, 2010; Winstral and Marks, 2014). Repeated LiDAR surveys throughout single seasons (Schirmer and Lehning, 2011; Schirmer et al., 2011) and over multiple seasons (Deems et al., 2008) have found similar results through fractal analyses of the snow depth distributions. By comparing findings from large-scale LiDAR snow depth surveys to operational hydrologic models, we can pinpoint causes of any shortcomings and subsequently refine model results.

Developed by the National Operational Hydrologic Remote Sensing Center (NOHRSC) and first operationally implemented in 2004, the Snow Data Assimilation System (SNODAS) estimates various snow properties by merging satellite, airborne, and ground-based snow data with modeled approximations of snow cover (Barrett, 2003). Historical model output from SNODAS is stored and archived at the National Snow and Ice Data Center (NSIDC) in Boulder, Colorado for every day that the model has been executed since its inception. These eight snow properties are the primary estimates that are made available to the public:

1. Snow water equivalent (SWE)
2. Snow depth
3. Snow melt runoff from the base of the snowpack
4. Sublimation from the snowpack
5. Sublimation of blowing snow
6. Solid precipitation
7. Liquid precipitation
8. Snowpack average temperature

A large portion of model fidelity is directed towards SWE prediction rather than any of the other model outputs because the amount of total water storage within snowpacks is far more important for water managers. The physically-based energy- and mass-balance NOHRSC Snow Model (NSM), described by Carroll et al. (2006), is the primary component of SNODAS, while an assimilation step gives analysts the ability to decide every day whether to augment the model estimates with any available remote sensing or ground-based measurements. Ultimately, the final model product has a spatial resolution of approximately 1 km^2 over the coterminous United States.

Independent validation data for SNODAS is scarce as a consequence of the framework's data assimilating nature which ensures that all available data at the model scale (i.e. 1 km^2) is used to adjust estimates of the NSM (Barrett, 2003). An alternative validation method has been to perform comparisons of SNODAS with other hydrologic models and satellite remote sensing products. Rutter et al. (2008) compared various NSM properties with two energy-balance snow models, but found difficulty in constraining model uncertainties due primarily to the high sub-grid spatial variability exhibited in mountain snowpacks. Other studies have used SNODAS as the validation source for large-scale hydrologic models such as the Noah land surface model (Barlage et al., 2010), and SWE retrieval using satellite-based microwave radar remote-sensing platforms (Azar et al., 2008).

To our knowledge, only two validation studies of SNODAS' performance have been conducted using independent datasets and each of those studies relied on extensive *in situ* measurement campaigns. Clow et al. (2012) performed snow surveys of snow depth within 45 SNODAS pixels over a three month period in 2007. The results revealed that SNODAS performed satisfactorily for predicting snow depth in forested areas, but depth estimates in alpine areas were poor in comparison to manual measurements chiefly due to sub-grid scale variability from wind redistribution of snow. This discrepancy was addressed by applying a correction factor to account for wind redistribution of snow in the wind-affected alpine areas. In another study, Anderson (2011) intensively sampled three SNODAS pixels in the mountains just north of Boise, Idaho over the course of two winter seasons and found that SNODAS slightly under predicted snow depths in heavily-forested areas but maintained reasonable estimates of SWE overall. Each of the studies required an enormous amount of manpower and time to obtain the independent datasets for proper comparison, but came to somewhat different conclusions about the model performance most likely due to the individual locations of the collected data (Idaho and Colorado, USA). This study's goal was to increase the spatial continuity of the validation dataset in order to come closer to discovering individual biases with the SNODAS model framework.

2 Study Area

The second Cold Lands Processes Experiment campaign (CLPX-2, 2006-2008) was a multi-faceted mission designed to cover a much larger coincident extent than the previous campaign (CLPX-1, 2002-2003) three years prior. The primary objective of CLPX-2 was to acquire snow volume backscatter measurements from NASA's POLSCAT (POLarimetric SCATterometer) airborne Ku-band radar system and the necessary ground truth measurements (Yueh et al., 2009) for validation of the proposed NASA Snow and Cold Land Proceses (SCLP) and ESA Cold Regions Hydrology High-resolution Observatory (CoreH_2O) satellite missions (Rott et al., 2010). The airborne LiDAR portion of the campaign was intended to be an ancillary validation dataset for the radar measurements.

Three Intensive Observation Periods (IOPs) were organized over a 9×84 km rectangular swath to the south and east of the town of Steamboat Springs in northern Colorado, USA (Figure 8). During both IOP-1 (early December, 2006) and IOP-3 (late February, 2007), airborne LiDAR surveys were performed to provide high-resolution surface elevation change datasets to aid in the POLSCAT validation process. Covering approximately 750 km^2, the study area encompasses a wide range of elevations, terrain and vegetation types, and ecological classes. Maximum LiDAR-derived changes in snow depth varied from merely 30 cm in the central wind-swept prairies to over 4 meters in the drifts of the higher elevations.

The study area can be viewed as containing three main classification areas: 1) the grass-covered, low-elevation rolling farmland in the Yampa River Valley in the far west; 2) the coniferous forests of the Rabbit Ears Pass portion of the Park Range as well as the foothills of the Medicine Bow Mountains in the far east; and 3) the sagebrush-dominated high desert of the central North Park region. Six SNOw TELemetry (SNOTEL) sites, operated by the National Resources Conservation Service (NRCS), are located within 15 km of the study area and yield a relatively dense network of automated measurements of various snowpack properties. The data from these ground-based measurement stations are often assimilated by SNODAS in order to augment the NSM estimates.

3 Methods

3.1 LiDAR Acquisitions

Due to the supportive role of the LiDAR surveys, only two flights were planned and carried out concurrent with the POLSCAT radar acquisitions. On December 3rd, 2006 and February 22nd, 2007 LiDAR acquisitions were obtained by Fugro Horizons, Inc. using a Leica ALS50 laser range finder onboard a Cessna 310 aircraft flying at 3,000 meters above ground level. The 1064 nm laser wavelength is optimal for snow covered surfaces owing to the minimal penetration depth on the order of only 1 cm (Deems et al., 2013). The pulse rate of 32,500 Hz, combined with the aircraft's speed, altitude, and scan rate, resulted in raw point clouds with nominal point spacings of 2.0 – 2.5 meters, depending on surface roughness, canopy coverage and scan angle relative to the aircraft.

The LiDAR vendor filtered vegetation returns from ground returns using a minimum block mean algorithm and proprietary software to create vegetation-filtered point clouds for each flight with updated nominal point spacings of 2.5 – 3.0 meters, again depending on the terrain, canopy cover and scan angle. Various alternative filtering algorithms were explored during the course of this study, but the decision was ultimately made to utilize the vendor-filtered data in order to maintain consistency over the large variety of landscapes. Next, we applied the open-source Points2Grid interpolation tool, employing an inverse distance weighting scheme, to produce a 5-meter Digital Surface Model (DSM) for both of the vegetation-filtered point clouds. Because the CLPX-2 LiDAR scans were never acquired over an absolutely snow-free surface, as many of the higher elevations had already received snow by December 3rd, the interpolated surfaces were differenced to provide a raster of the estimated change in total snow height between December 3rd and February 22nd at 5-meter resolution (Figure 9a). This 5-meter gridded product of LiDAR-estimated changes in snow depth will hereafter be referred to as Δ*LiDAR*. Though less dense than the original CLPX-1 point clouds used by Deems et al. (2006), Trujillo et al. (2007), and McCreight et al. (2014), the CLPX-2 LiDAR footprint covered a greater variety of terrain, vegetation, and snowpack classes, thereby providing a useful comparison tool for hydrologic snow models over large spatial extents.

3.2 *In Situ* Measurements

All remote sensing methods are subject to an appreciable amount of measurement uncertainty which should be quantified, if possible, by ground truth validation. The CLPX-2 intensive manual measurement campaigns were arranged and completed by a team of 12–15 researchers during each IOP, and originally intended to be the primary ground truth dataset for the multi temporal POLSCAT radar acquisitions over the 2006-2007 winter season (Yueh et al., 2009). Twelve 500×500 m "hourglass" transects, (Figure 8, and henceforth referred to as *HG* sites) comprised of 47–50 evenly spaced snow depth measurements, were manually sampled during IOP-1 and IOP-3 within a day of each of the CLPX-2 LiDAR acquisitions. The *HG* sites were chosen to represent physiographically distinctive regions of the CLPX-2 survey swath. Ground crews made measurements at preprogrammed waypoints loaded onto mapping-grade handheld GPS units in order to maintain location consistency for each survey. We estimate the resulting relative point-to-point horizontal uncertainty between the *HG* surveys to be less than 2 meters while the *HG* transect locations themselves can be approximated to 7 meters in absolute space. The repeated *HG* surface elevation measurements were differenced to provide a similar comparison metric of snow depth change, or Δ*HG*, to the Δ*LiDAR* dataset.

3.3 SNODAS Snow Depths

SNODAS estimates of snow depth were downloaded from the NSIDC for the two dates of the CLPX-2 LiDAR acquisitions (December 3rd, 2006 and February 22nd, 2007), then spatially referenced to the UTM coordinate projection. The two rasters of snow depth were differenced to provide 1-km gridded model estimates of snow depth change, hereafter referred to as Δ*SNODAS*. Figure 9c depicts Δ*SNODAS* over the area surrounding the LiDAR swath, along with the locations of all nearby SNOTEL stations that can be used for model assimilation when necessary. In order to aid in uncoupling the causal influences of error within the model, we examined the SNODAS estimates of snow melt due to incoming solar radiation and ambient air temperature between the survey dates (Figure 10). Only in the North Park region did any appreciable melt occur (10–20% of the total snow precipitation), while everywhere else experienced negligible mass loss. Therefore, we can be more certain that Δ*SNODAS* discrepancies from Δ*LiDAR* were due to other factors such as sublimation and densification routines within the model or uncertainties in the LiDAR data.

3.4 SNODAS/*In Situ* Measurement Comparison

To provide a link to the previous ground-based SNODAS validation studies, we examined the ability of manual measurements from the twelve *HG* sites to represent Δ*SNODAS*. Mentioned previously, Clow et al. (2012) averaged depth measurements from snow surveys performed within 45 individual SNODAS pixels to perform model validation. We employed the same basic method to assess SNODAS-predicted snow depth changes using the CLPX-2 *in situ* Δ*HG* transects. The mean Δ*HG* over each *HG* site was calculated along with an associated interquartile range. Then, a new coincident 1 km^2 Δ*SNODAS* estimate was constructed around each *HG* transect site from the areal coverage fraction of the four overlapping SNODAS pixels, creating an area-weighted average of Δ*SNODAS* centered over each Δ*HG* measurement site. This spatial averaging was performed because the CLPX-2 campaign was not designed during the planning phase to be a validation source for SNODAS and the *HG* transects were therefore not aligned within individual model pixels.

3.5 Characterizing LiDAR uncertainty

The 750 km^2 CLPX-2 LiDAR dataset (Δ*LiDAR*) overlaps 980 individual SNODAS pixels completely[1], supplying a statistically robust validation dataset for determining contributing factors to SNODAS uncertainty. However, Δ*LiDAR* mea-

[1]Though previously stated as nominally 1 km^2, the actual resolution of SNODAS is 30 arcseconds because the model is implemented in the geographic coordinate system (Barrett, 2003). At the CLPX-2 latitude, 30 arcsecs $\approx$ 830 meters.

surements are fundamentally estimates themselves and require uncertainty assessments, which was available from the *in situ* Δ*HG* transects. To account for the horizontal position uncertainty in both the Δ*HG* and Δ*LiDAR* datasets, the 5-meter gridded Δ*LiDAR* estimates were averaged in a 10 meter radius around each reported *in situ* point measurement location and treated as a separate point measurement for comparison purposes.

To perform the model comparison, the 5-meter Δ*LiDAR* estimates were binned into the spatial extents of the 980 overlaid Δ*SNODAS* grid cells. Statistics were calculated within each 1-km pixel, resulting in a mean, standard deviation, and interquartile range of Δ*LiDAR* estimates over the CLPX-2 study area at the SNODAS model resolution. These mean Δ*LiDAR* estimates are portrayed in Figure 9b.

4 Results and Discussion

To link this study to previous SNODAS validation efforts that used independent manual measurements, we compared the twelve averaged *in situ* Δ*HG* transect pixels to the Δ*SNODAS* estimates to determine the feasibility of validating the model with in situ gathered data. The comparison is shown as the blue circles in Figure 11. The trend of this limited dataset of only 12 measurement points suggests that as the mean snow depth within a model pixel increases above approximately 40 cm, the ability of SNODAS to estimate the amount of total snow depth change decreases substantially. Also, it is not clear from the small sample size of the *in situ* data what physical factors could be influencing such discrepancies, and a much more spatially extensive dataset, such as the CLPX-2 LiDAR, is required for determining the underlying causes of model error.

The exhaustive CLPX-2 *in situ HG* measurement campaign provided an ideal dataset for limiting uncertainty in the large-scale LiDAR surveys of December 3rd, 2006 and February 22nd, 2007. The changes in snow depth as measured by the standard probing method and interpolated from the LiDAR surveys were compared throughout all twelve *HG* sites individually, and then averaged on a site by site basis similar to the averaging scheme used by Clow et al. (2012). The red crosses in Figure 11 indicate the correlation between the upscaled Δ*LiDAR* with the mean Δ*HG* measurements. As stated previously, the comparison dataset of LiDAR snow depth change was determined from the average 5 meter gridded Δ*LiDAR* estimate within a 10-meter radius surrounding each reported Δ*HG* measurement. The purpose of the Δ*LiDAR* areal averaging was to account for error in the handheld GPS units that were used to locate survey points. The resulting 12.9 cm RMS difference between mean Δ*HG* and mean Δ*LiDAR* point estimates for all twelve CLPX-2 *HG* sites is well within the bounds of conventional airborne LiDAR uncertainty estimates (Baltsavias, 1999; Hodgson and Bresnahan, 2004). The Δ*LiDAR* observations resulted in a higher r^2 value (0.942) than the Δ*SNODAS* depth estimates (0.655), but exhibited a small negative bias over all the *HG* sites.

The presence of a slight negative bias in the Δ*LiDAR* results with respect to the manual measurements could be due to a combination of contributing factors: 1) a number of December LiDAR returns may have not fully penetrated the low-lying brush and grass, resulting in lower estimates of snow depth change; 2) the snow depth probe tips may have penetrated the soil more easily during the February in situ measurement campaign, which would have produced higher estimates of snow depth change; and 3) the difference in measurement support that exists between the tip of a snow depth probe ($<$ 1 cm) and spatially averaged and interpolated 5 meter LiDAR may have had an effect on the 1-km^2 averaged snow depth change as the spatial variability of each site increased. Nevertheless, the negative bias of 5-15 centimeters shown in Figure 11 is on the order of the LiDAR uncertainty, and since the sample size of comparisons was so small relative to the total area of the survey footprint, no bias correction was performed on the Δ*LiDAR* data for the SNODAS validation.

The comparison between Δ*SNODAS* and mean Δ*LiDAR* within the model pixels (Figure 12) resulted in an $r^2 = 0.72$, signifying a reasonably strong correlation between the two estimate datasets. Since snow melt between the LiDAR flights was found to be an insignificant portion of the snowpack evolution (Figure 10), the actual changes in snow depth over the study area were primarily influenced by accumulation, densification, sublimation, and redistribution factors.

To investigate the primary cause of disagreement between Δ*SNODAS* and Δ*LiDAR*, seven potential explanatory physiographic variables were culled from the LiDAR data to perform a regression analysis. In addition to the Δ*LiDAR* estimates and the vegetation-filtered elevations, vegetation height and canopy coverage across the survey swath was calculated at 5-m resolution using both the raw and vegetation-filtered December LiDAR point returns. Vegetation heights and elevations were each upscaled to the 1-km SNODAS resolution in a similar fashion to the LiDAR snow depth change, while the vegetation density was calculated by finding the number of 5-meter pixels within each 1-km SNODAS grid cell that contained LiDAR first returns greater than 50 cm above the filtered surface. Lastly, the interquartile range of the 1-km averaged (upscaled) variables was determined to result in the following group of 7 individual predictor variables for regression analysis:

1. Vegetation density [%]
2. Mean vegetation height [cm]
3. Inter-quartile range of vegetation height [cm]
4. Mean snow depth change [cm] (Dec. 3rd – Feb. 22nd)
5. Inter-quartile range of snow depth change [cm]
6. Mean elevation [m]
7. Inter-quartile range of elevation [m]

The upscaled snow depth changes (#4) were overwhelmingly found to best predict the discrepancy between $\Delta SNODAS$ and $\Delta LiDAR$, indicating that none of the other six explanatory variables singularly influenced SNODAS performance over the entire study area. However, as terrain and canopy coverage certainly have an influence on model performance and LiDAR uncertainty, it should be noted here that smaller subsets of the survey area might have yielded differing results from our analysis of the entire 750 km^2 study area. Figure 13 shows the plot of mean $\Delta LiDAR$ estimates against the $\Delta SNODAS - \Delta LiDAR$ differences within each model pixel. Within that plot, the pink vertical and blue horizontal shaded areas between -13 and +13 cm on each axis represent the minimum attainable resolution of the $\Delta LiDAR$ estimates determined from the ΔHG ground-based measurement data. Three regions have been circled in the figure, each corresponding to portions of the difference dataset that were found to be outside the uncertainty levels of the LiDAR-derived changes in snow depth.

Contrasting the images of $\Delta SNODAS - \Delta LiDAR$ (Figure 14) and mean $\Delta LiDAR$ (Figure 9b) reveals the geographic locations of the three regions within the survey swath containing the greatest SNODAS and LiDAR disagreements. Within these three regions specific physiographic factors are likely the causes of greater relative discrepancies.

Region #1: North Park

The region within the survey area exhibiting the lowest annual snow totals is approximately delineated within Figure 14, and is comprised of pixels that SNODAS has estimated to have had a larger positive change in snow depth than that of the LiDAR acquisitions (Figure 13). However, the LiDAR snow depth changes within these pixels are well below the trusted LiDAR uncertainty level (the pink vertical shading). These pixels are located in the North Park region of the survey area, where the flat landscape is densely populated by low sagebrush ($\approx$ less than 30 cm) and high winds frequently scour the snow above and near the height of the sage throughout the winter. The snow that remains is subsequently packed between the low vegetation and the snow height changes very little throughout the year once it has reached a height similar to the sagebrush. SNODAS does incorporate a sublimation factor due to wind into the accumulation model, but requires an accurate representation of wind speed and direction as input to the assimilation step. In the case of the prairie-like North Park area the nearest meteorological station used by the model assimilation step is a sheltered SNOTEL site located nearly 15 kilometers to the southwest in very different terrain and at a higher altitude, affecting not only the wind forcings, but also solar radiation as well. Further study is required to quantify the effect of the distance from assimilation measurement sites on SNODAS performance for remote areas such as North Park.

Region #2: East Slope of Rabbit Ears Pass

Pixels that comprise region #2 in Figure 13 are where snow depths are similarly estimated by SNODAS to have accumulated more snow than observed by the LiDAR. However, the pixels are in a region with higher snow accumulation totals, which are above the lower LiDAR uncertainty level of 13 cm. Again delineated in Figure 14, these pixels are nestled directly to the east of Rabbit Ears Pass where the Columbine SNOTEL station provides assimilation data for SNODAS. Since the relative error of the LiDAR observations is small and a large altitudinal effect can be seen in the $\Delta LiDAR$ (Figure 9b), this discrepancy can possibly be attributed to SNODAS distributing the SNOTEL information to areas of lower elevations and vegetation types. Future study on regions such as this would be important for determining optimal precipitation forcings by the SNODAS data assimilation process.

Region #3: Rabbit Ears Pass

Finally, the region #3 pixels represent an area where the upscaled LiDAR changes in snow depth are significantly larger than the SNODAS estimates. These pixels occur primarily in topographically complex areas with exceptionally high snow totals and dense coniferous forests, once again outlined in Figure 14. The probable controlling factor of underestimation by SNODAS in this region is the sub-kilometer scale heterogeneity of snow distribution caused by both vegetation and topography. Furthermore, canopy interception remains an important aspect of the mountain snow energy balance that is still not well understood (Marks et al., 2008; Pohl et al., 2014), adding uncertainty to the assimilation model framework. SNODAS has been found to underestimate snow depths in similar forested alpine terrain (Anderson, 2011), so this result is not unexpected. Areas such as Rabbit Ears Pass are of primary interest to water managers due to the amount of water stored in the snowpack, so more analysis is required to effectively constrain SNODAS uncertainty in complex, deep snowpacks.

5 Conclusions

Over the past decade, high resolution snow depth information has become a highly sought-after data product by snow researchers and many scientific questions have been addressed using the spatial continuity and extent provided by LiDAR surveys. This study first examined the ability of ground-based measurements to constrain remote sensing uncertainty, and in turn compared the remote sensing estimates to an operational hydrologic model for validation purposes. In this case, the CLPX-2 ground truth campaign was vitally important for quantifying uncertainty in the LiDAR snow depth estimates, emphasizing the necessity of similar in situ campaigns to complement future LiDAR remote sensing missions.

From the comparison study, three distinct regions were extracted from the survey footprint that exhibited greater disagreement than could be explained by LiDAR estimate uncertainty alone. It is our opinion that the distinct physiographic characteristics within these three regions ultimately affected the accuracy of the SNODAS predictions of snow height change between the two LiDAR acquisitions.

To further investigate model performance, more studies are needed from subsequent large extent LiDAR surveys to focus on the accuracy of SNODAS as a function of distance from SNOTEL stations. Additionally, micro-scale wind redistribution effects could be applied within the model structure to assist in areas where blowing snow transport is a major cause of spatial variability. Finally, large-scale coincident density surveys would allow model validation with LiDAR-derived snow depths as well as *in situ* estimates of SWE, for which SNODAS is likely to be more accurate than compared with depth alone.

Author Contributions. Adam Winstral supplied modeling expertise, Kelly Elder produced the archived LiDAR datasets and orchestrated the *in situ* CLPX-2 measurement campaign, Simon Yueh and Donald Cline managed and coordinated the CLPX-2 campaign of 2006-2007, and Andrew Hedrick performed the validation comparison and manuscript preparation with significant support from Hans-Peter Marshall and contributions from all co-authors.

Acknowledgements. The authors would like to express their gratitude to all the researchers involved in the intensive ground-based measurement campaign during CLPX-2. The CLPX-2 LiDAR datasets were archived and maintained by Fugro Horizons, Inc. Daily SNODAS model runs from 2003 to the present day are archived at the National Snow and Ice Data Center in Boulder, Colorado. This research was funded in part by NASA grant #NNX10AO02G (NASA New Investigator Program), NASA grant #NNX10AN30A (NASA EPSCoR Program), and the USDA-ARS CRIS Project 5362-13610-008-00D: "Understanding Snow and Hydrologic Processes in Mountainous Terrain with a Changing Climate".

References

Anderson, B. T.: Spatial Distribution and Evolution of a Seasonal Snowpack in Complex Terrain: An Evaluation of the SNODAS Modeling Product, Masters thesis, Boise State University, Boise, Idaho, United States, 2011.

Anderton, S. P., White, S. M., and Alvera, B.: Evaluation of spatial variability in snow water equivalent for a high mountain catchment, Hydrological Processes, 18, 435–453, 2004.

Azar, A. E., Ghedira, H., Romanov, P., Mahani, S., Tedesco, M., and Khanbilvardi, R.: Application of satellite microwave images in estimating snow water equivalent, Journal of the American Water Resources Association, 44, 1347–1363, 2008.

Baltsavias, E.: Airborne laser scanning: basic relations and formulas, ISPRS Journal of Photogrammetry and Remote Sensing, 54, 199–214, 1999.

Barlage, M., Chen, F., Tewari, M., Ikeda, K., Gochis, D., Dudhia, J., Rasmussen, R., Livneh, B., Ek, M., and Mitchell, K.: Noah land surface model modifications to improve snowpack prediction in the Colorado Rocky Mountains, Journal of Geophysical Research, 115, 2010.

Barrett, A. P.: National Operational Hydrologic Remote Sensing Center SNOw Data Assimilation System (SNODAS) Products at NSIDC, NSIDC Special Report 11, p. 19, 2003.

Blöschl, G.: Scaling issues in snow hydrology, Hydrological Processes, 13, 2149–2175, 1999.

Carroll, T., Cline, D., Olheiser, C., Rost, A., Nilsson, A., Fall, G., Bovitz, C., and Li, L.: NOAA's national snow analyses, Proceedings, 74th Annual Meeting of the Western Snow Conference, pp. 1–14, 2006.

Cline, D., Yueh, S., Chapman, B., Stankov, B., Gasiewski, A., Masters, D., Elder, K., Kelly, R., Painter, T. H., Miller, S., Katzberg, S., and Mahrt, L.: NASA Cold Land Processes Experiment (CLPX 2002/03): Airborne Remote Sensing, Journal of Hydrometeorology, 10, 338–346, 2009.

Clow, D. W., Nanus, L., Verdin, K. L., and Schmidt, J.: Evaluation of SNODAS snow depth and snow water equivalent estimates for the Colorado Rocky Mountains, USA, Hydrological Processes, 26, 2583–2591, 2012.

Deems, J. S., Fassnacht, S. R., and Elder, K. J.: Fractal Distribution of Snow Depth from Lidar Data, Journal of Hydrometeorology, 7, 285–297, 2006.

Deems, J. S., Fassnacht, S. R., and Elder, K. J.: Interannual Consistency in Fractal Snow Depth Patterns at Two Colorado Mountain Sites, Journal of Hydrometeorology, 9, 977–988, 2008.

Deems, J. S., Painter, T. H., and Finnegan, D. C.: Lidar measurement of snow depth: a review, Journal of Glaciology, 59, 467–479, 2013.

Elder, K., Dozier, J., and Michaelsen, J.: Snow accumulation and distribution in an alpine watershed, Water Resources Research, 27, 1541–1552, 1991.

Erickson, T. A., Williams, M. W., and Winstral, A.: Persistence of topographic controls on the spatial distribution of snow in rugged mountain terrain, Colorado, United States, Water Resources Research, 41, 1–17, http://www.agu.org/pubs/crossref/2005/2003WR002973.shtml, 2005.

Hiemstra, C. A., Liston, G. E., and Reiners, W. A.: Snow Redistribution by Wind and Interactions with Vegetation at Upper Treeline in the Medicine Bow Mountains, Wyoming, U.S.A., Arctic, Antarctic, and Alpine Research, 34, 262–273, 2002.

Hiemstra, C. A., Liston, G. E., and Reiners, W. A.: Observing, modelling, and validating snow redistribution by wind in a Wyoming upper treeline landscape, Ecological Modelling, 197, 35–51, 2006.

Hodgson, M. E. and Bresnahan, P.: Accuracy of Airborne Lidar-Derived Elevation : Empirical Assessment and Error Budget, Photogrammetric Engineering & Remote Sensing, 70, 331–339, 2004.

Hopkinson, C., Sitar, M., Chasmer, L., and Treitz, P.: Mapping Snowpack Depth beneath Forest Canopies Using Airborne Lidar, Photogrammetric Engineering & Remote Sensing, 70, 323–330, 2004.

Liston, G. E.: Representing subgrid snow cover heterogeneities in regional and global models, Journal of Climate, 17, 1381–1398, 2004.

Liston, G. E. and Sturm, M.: Winter Precipitation Patterns in Arctic Alaska Determined from a Blowing-Snow Model and Snow-Depth Observations, Journal of hydrometeorology, 3, 646–660, 2002.

Liston, G. E., Haehnel, R. B., Sturm, M., Hiemstra, C. a., Berezovskaya, S., and Tabler, R. D.: Simulating complex snow distributions in windy environments using SnowTran-3D, Journal of Glaciology, 53, 241–256, 2007.

Luce, C. H., Tarboton, D. G., and Cooley, K. R.: Sub-grid parameterization of snow distribution for an energy and mass balance snow cover model, Hydrological Processes, 13, 1921–1933, 1999.

Marchand, W. D. and Killingtveit, A.: Statistical probability distribution of snow depth at the model sub-grid cell spatial scale, Hydrological Processes, 19, 355–369, 2005.

Marks, D., Winstral, A., Flerchinger, G., Reba, M., Pomeroy, J., Link, T., and Elder, K.: Comparing Simulated and Measured Sensible and Latent Heat Fluxes over Snow under a Pine Canopy to Improve an Energy Balance Snowmelt Model, Journal of Hydrometeorology, 9, 1506–1522, 2008.

McCreight, J. L., Slater, A. G., Marshall, H. P., and Rajagopalan, B.: Inference and uncertainty of snow depth spatial distribution at the kilometre scale in the Colorado Rocky Mountains: the effects of sample size, random sampling, predictor quality, and validation procedures, Hydrological Processes, 28, 933–957, 2014.

Melvold, K. and Skaugen, T.: Multiscale spatial variability of lidar-derived and modeled snow depth on Hardangervidda, Norway, Annals of Glaciology, 54, 273–281, 2013.

Mott, R., Schirmer, M., and Lehning, M.: Scaling properties of wind and snow depth distribution in an Alpine catchment, Journal of Geophysical Research, 116, 1–8, 2011.

Pohl, S., Garvelmann, J., Wawerla, J., and Weiler, M.: Potential of a low-cost sensor network to understand the spatial and temporal dynamics of a mountain snow cover, Water Resources Research, 50, 2533–2550, 2014.

Pomeroy, J., Gray, D., and Landine, P.: The Prairie Blowing Snow Model: characteristics, validation, operation, Journal of Hydrology, 144, 165–192, 1993.

Rott, H., Yueh, S. H., Cline, D. W., Duguay, C., Essery, R., Haas, C., Heliere, F., Kern, M., Macelloni, G., Malnes, E., Nagler, T., Pulliainen, J., Rebhan, H., and Thompson, A.: Cold Regions Hydrology High-Resolution Observatory for Snow and Cold Land Processes, Proceedings of the IEEE, 98, 752–765, 2010.

Rutter, N., Cline, D., and Li, L.: Evaluation of the NOHRSC Snow Model (NSM) in a One-Dimensional Mode, Journal of Hydrometeorology, 9, 695–711, 2008.

Schirmer, M. and Lehning, M.: Persistence in intra-annual snow depth distribution: 2. Fractal analysis of snow depth development, Water Resources Research, 47, 1–14, 2011.

Schirmer, M., Wirz, V., Clifton, A., and Lehning, M.: Persistence in intra-annual snow depth distribution: 1. Measurements and topographic control, Water Resources Research, 47, 1–16, 2011.

Skaugen, T. and Randen, F.: Modeling the spatial distribution of snow water equivalent, taking into account changes in snow-covered area, Annals of Glaciology, 54, 305–313, 2013.

Sturm, M. and Wagner, A. M.: Using repeated patterns in snow distribution modeling: An Arctic example, Water Resources Research, 46, 1–15, 2010.

Sturm, M., Liston, G. E., Benson, C. S., and Holmgren, J.: Characteristics and Growth of a Snowdrift in Arctic Alaska, U.S.A., Arctic, Antarctic, and Alpine Research, 33, 319–329, 2001a.

Sturm, M., McFadden, J. P., Liston, G. E., Chapin III, F. S., Racine, C. H., and Holmgren, J.: Snow–Shrub Interactions in Arctic Tundra : A Hypothesis with Climatic Implications, Journal of Climate, 14, 336–345, 2001b.

Sturm, M., Taras, B., Liston, G. E., Derksen, C., Jonas, T., and Lea, J.: Estimating Snow Water Equivalent Using Snow Depth Data and Climate Classes, Journal of Hydrometeorology, 11, 1380–1394, 2010.

Trujillo, E., Ramírez, J. A., and Elder, K. J.: Topographic, meteorologic, and canopy controls on the scaling characteristics of the spatial distribution of snow depth fields, Water Resources Research, 43, 1–17, 2007.

Trujillo, E., Ramirez, J. A., and Elder, K. J.: Scaling properties and spatial organization of snow depth fields in sub-alpine forest and alpine tundra, Hydrological Processes, 23, 1575– 1590, 2009.

Winstral, A. and Marks, D.: Long-term snow distribution observations in a mountain catchment: Assessing variability, time stability, and the representativeness of an index site, Water Resources Research, 50, 293–305, 2014.

Winstral, A., Elder, K., and Davis, R. E.: Spatial Snow Modeling of Wind-Redistributed Snow Using Terrain-Based Parameters, Journal of Hydrometeorology, 3, 524–539, 2002.

Yueh, S. H., Dinardo, S. J., Akgiray, A., West, R., Cline, D. W., and Elder, K.: Airborne Ku-Band Polarimetric Radar Remote Sensing of Terrestrial Snow Cover, IEEE Transactions on Geoscience and Remote Sensing, 47, 3347–3364, 2009.

Fig. 1. Location of the CLPX-2 LiDAR footprint in Colorado, USA with nearby towns, SNOTEL sites, and IOP *in situ* hourglass (HG) measurement transect locations indicated.

Fig. 2. Estimates of snow depth change between December 3rd, 2006 and February 22nd, 2007 along with the six nearby SNOTEL sites used by SNODAS for data assimilation. **(a)** represents the 5-meter resolution LiDAR-derived snow depth change, $\Delta LiDAR$, **(b)** shows the upscaled LiDAR estimates of snow depth change at the 1-km SNODAS resolution, and **(c)** is the difference in SNODAS estimates of snow depth, $\Delta SNODAS$, on the dates of the LiDAR acquisitions, with the LiDAR footprint outlined for reference.

Fig. 3. SNODAS estimates of snow melt as a percentage of the estimated mass lost from the estimated mass gained due to accumulation between December 3rd, 2006 and February 22nd, 2007. Mass loss due to ambient air temperature and solar radiation between the survey dates can be effectively eliminated as a cause of model error over much of the CLPX2 survey footprint.

Fig. 4. $\Delta SNODAS$ (blue circles) and $\Delta LiDAR$ (red crosses) snow depths evaluated over the centers of the twelve ΔHG measurement transects. The $\Delta LiDAR$ points were determined by averaging each reported 5 meter resolution $\Delta LiDAR$ snow depth within a 10 meter radius of each reported HG measurement, then averaging again over each HG transect site. The $\Delta SNODAS$ estimates were the areal-weighted averages of the four nearest SNODAS pixels to the center of each HG transect site.

Fig. 5. SNODAS model estimates plotted against mean LiDAR-derived snow depth change within all 1-km^2 SNODAS pixels (n=980).

Fig. 6. Pixel by pixel $\Delta SNODAS - \Delta LiDAR$ differences of snow depth change plotted against the mean $\Delta LiDAR$ within each SNODAS pixel. The pink and blue shaded areas represent the $\pm$13cm error threshold for the upscaled LiDAR estimates determined from the CLPX-2 *in situ* ΔHG measurements. Three distinct regions are circled that fall outside the 13cm error threshold, signifying a particular physiographic forcing factor present in the three specific areas. Also plotted is a histogram of differences showing a bias toward higher SNODAS estimates across the CLPX-2 study area.

Fig. 7. Image of the difference ($\Delta SNODAS - \Delta LiDAR$) between model and remote sensing estimates of snow depth change between December 3rd, 2006 and February 22nd, 2007 over the CLPX-2 study area. The three outlined regions correspond to the regions highlighted in Figure 13.

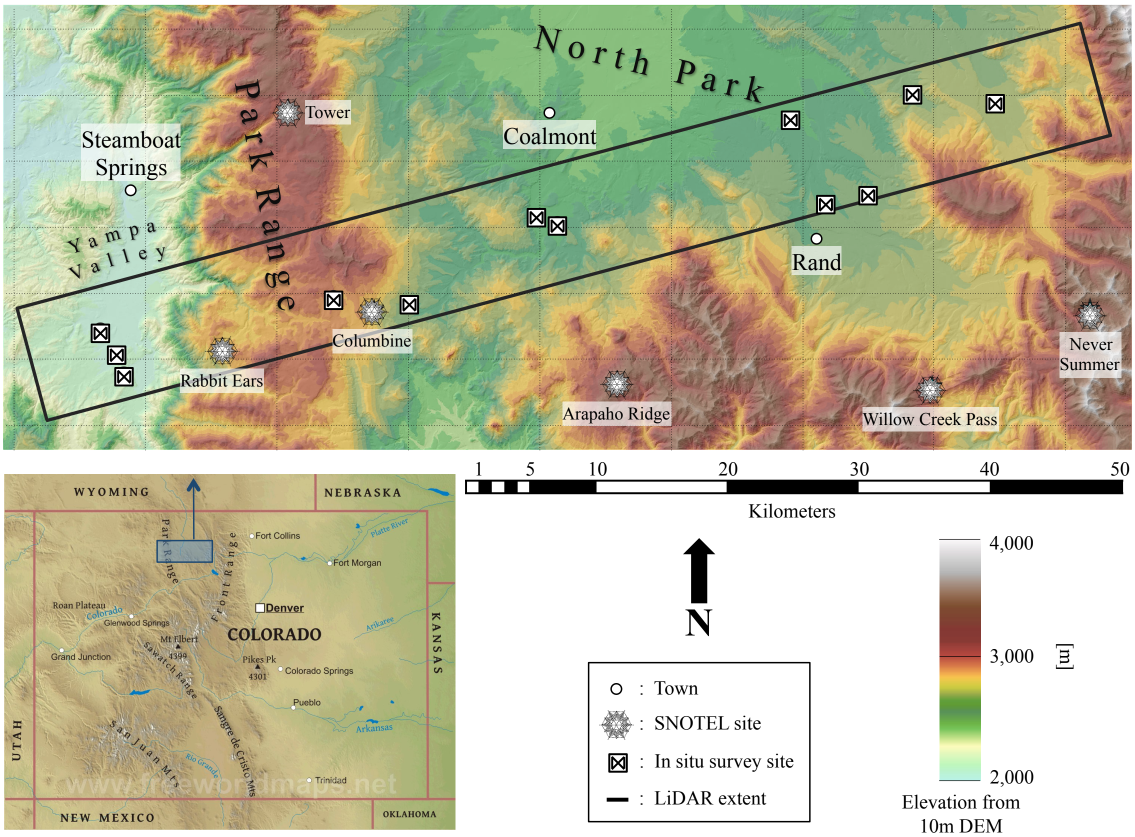

Fig. 8. Location of the CLPX-2 LiDAR footprint in Colorado, USA with nearby towns, SNOTEL sites, and IOP *in situ* hourglass (HG) measurement transect locations indicated.

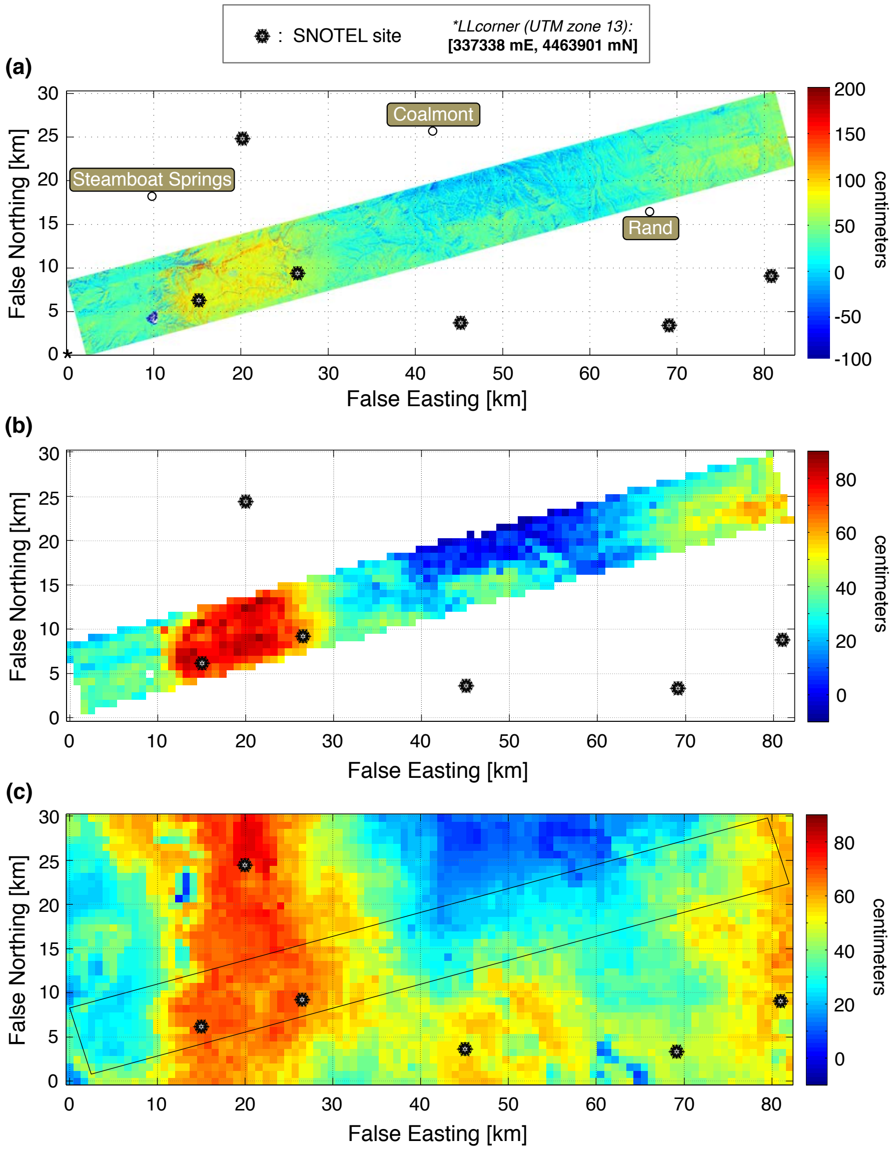

Fig. 9. Estimates of snow depth change between December 3rd, 2006 and February 22nd, 2007 along with the six nearby SNOTEL sites used by SNODAS for data assimilation. **(a)** represents the 5-meter resolution LiDAR-derived snow depth change, $\Delta LiDAR$, **(b)** shows the upscaled LiDAR estimates of snow depth change at the 1-km SNODAS resolution, and **(c)** is the difference in SNODAS estimates of snow depth, $\Delta SNODAS$, on the dates of the LiDAR acquisitions, with the LiDAR footprint outlined for reference.

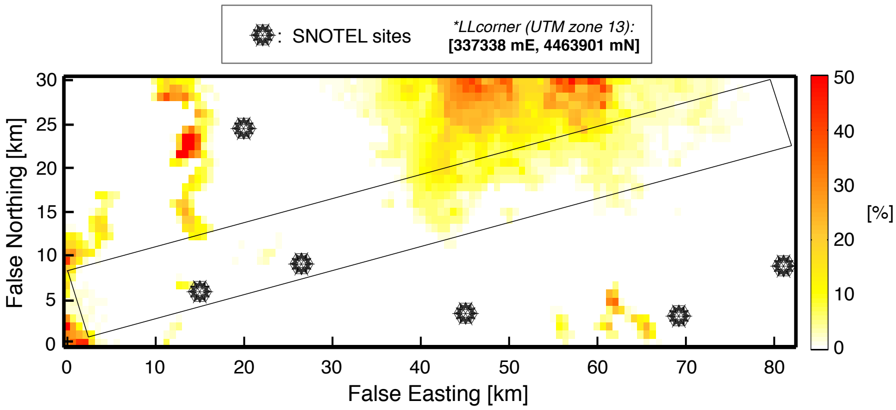

Fig. 10. SNODAS estimates of snow melt as a percentage of the estimated mass lost from the estimated mass gained due to accumulation between December 3rd, 2006 and February 22nd, 2007. Mass loss due to ambient air temperature and solar radiation between the survey dates can be effectively eliminated as a cause of model error over much of the CLPX2 survey footprint.

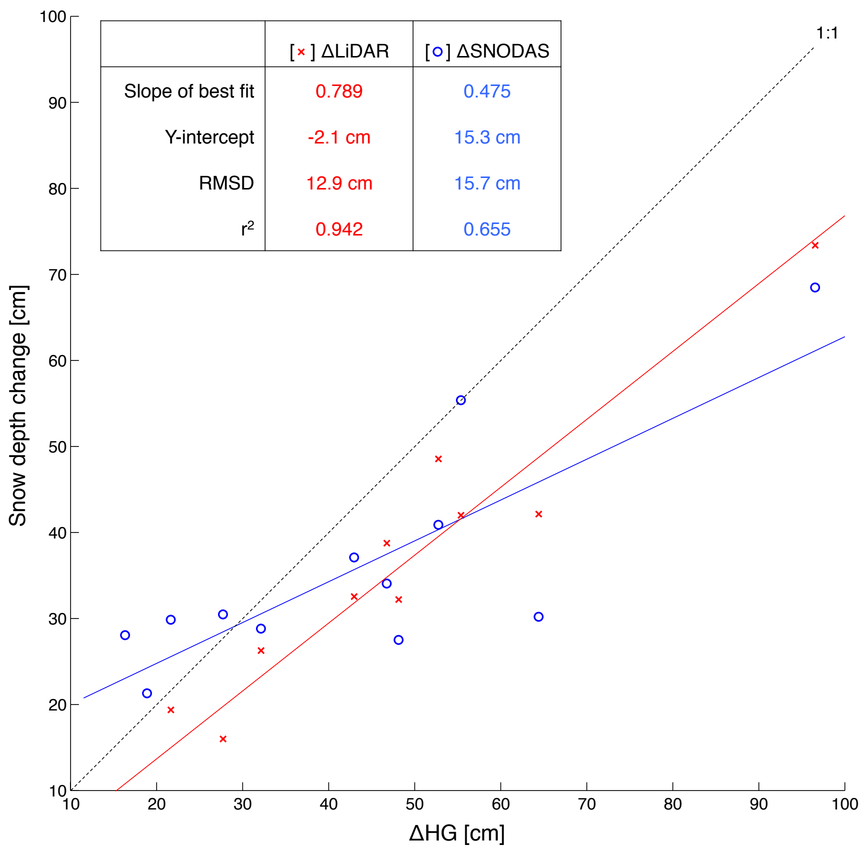

Fig. 11. Δ*SNODAS* (blue circles) and Δ*LiDAR* (red crosses) snow depths evaluated over the centers of the twelve Δ*HG* measurement transects. The Δ*LiDAR* points were determined by averaging each reported 5 meter resolution Δ*LiDAR* snow depth within a 10 meter radius of each reported HG measurement, then averaging again over each HG transect site. The Δ*SNODAS* estimates were the areal-weighted averages of the four nearest SNODAS pixels to the center of each HG transect site.

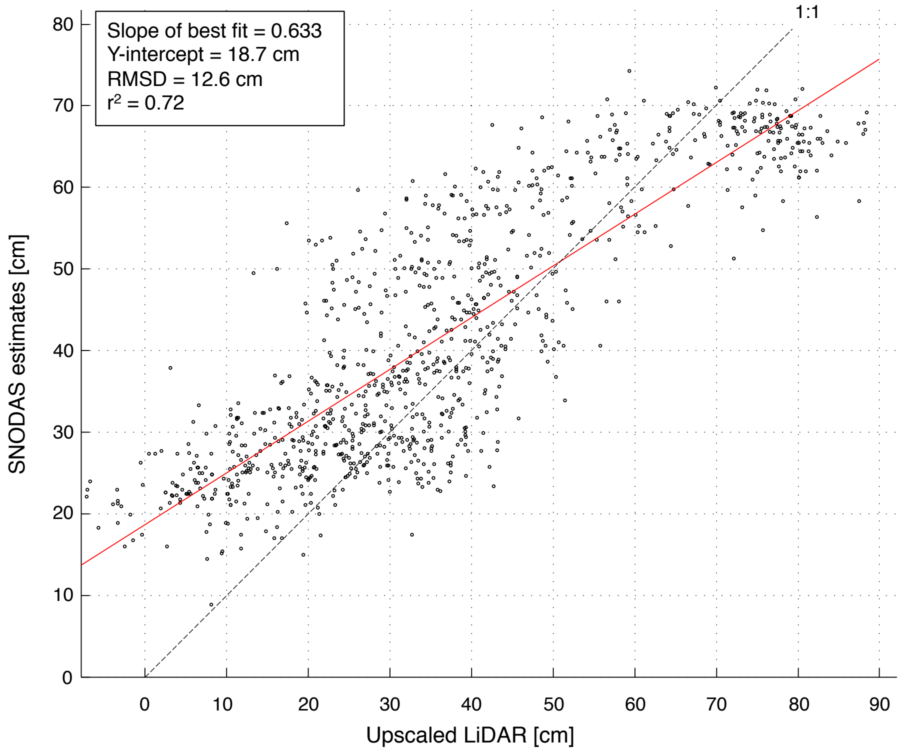

Fig. 12. SNODAS model estimates plotted against mean LiDAR-derived snow depth change within all 1-km^2 SNODAS pixels (n=980).

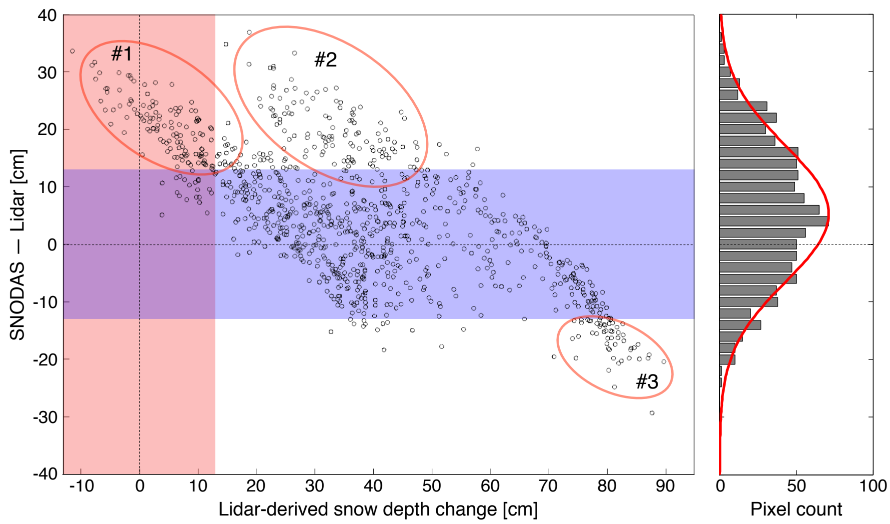

Fig. 13. Pixel by pixel $\Delta SNODAS - \Delta LiDAR$ differences of snow depth change plotted against the mean $\Delta LiDAR$ within each SNODAS pixel. The pink and blue shaded areas represent the ±13cm error threshold for the upscaled LiDAR estimates determined from the CLPX-2 *in situ* ΔHG measurements. Three distinct regions are circled that fall outside the 13cm error threshold, signifying a particular physiographic forcing factor present in the three specific areas. Also plotted is a histogram of differences showing a bias toward higher SNODAS estimates across the CLPX-2 study area.

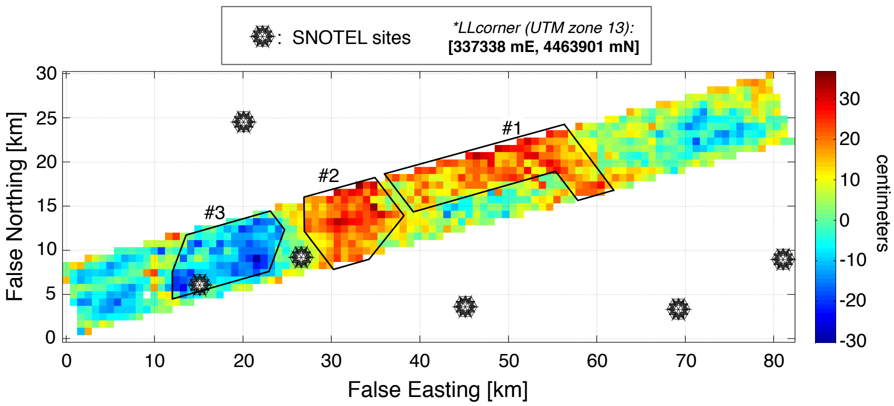

Fig. 14. Image of the difference ($\Delta SNODAS - \Delta LiDAR$) between model and remote sensing estimates of snow depth change between December 3rd, 2006 and February 22nd, 2007 over the CLPX-2 study area. The three outlined regions correspond to the regions highlighted in Figure 13.