

Snow depth from ICESat laser altimetry — A test study in southern Norway

D. Treichler^{a,*}, A. Kääb^a

^a*Department of Geosciences, University of Oslo, Sem Sælands vei 1, 0371 Oslo, Norway*

Accepted manuscript, ©2017. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Published version:

Remote Sensing of Environment, Volume 191, 15 March 2017, pp. 389–401. DOI: 10.1016/j.rse.2017.01.022

Abstract

Direct snow depth measurements are sparse, especially in remote areas. In this study, we assess the potential of ICESat laser altimetry for providing snow depths for its operational period 2003–2009 on the example of the Scandinavian Mountains in southern Norway. Snow cover during ICESat campaigns typically results in positive elevation differences (dh) between ICESat GLAH14 elevations and reference elevations from Digital Elevation Models (DEMs). Three DEMs are used: the Norwegian national DEM for the entire study area, and the SRTM DEM and a high-resolution airborne lidar DEM for a spatial subset on the Hardangervidda mountain plateau. To account for uncertainty in elevation data, ICESat samples are grouped into spatial subsets, elevation bands, and over time (e.g. all winter campaigns together). We find that ICESat has the potential to provide regional-scale snow depths for the years 2003–2009 for its winter (March) and late spring (June) campaigns. ICESat-derived snow depth time series for different elevation bands agree well with measured (RMSE 0.47 m) and modelled (RMSE 0.61 m) snow depths in the study area. Annual differences in snow amounts and the increase of snow depths with elevation and coastal proximity over the study area are correctly reproduced. Uncertainties in reference elevations exceed ICESat elevation uncertainty and good control over errors and biases in reference DEMs turns out essential. Spatially varying vertical offsets between ICESat and the reference DEMs make it necessary to bias-correct March/June snow depths with autumn dh per spatial unit or elevation band. Best results are achieved when samples are summarised per season over the entire observation period. After correction of local DEM biases, the spatial pattern of ICESat 2003–2009 March dh matches spatially distributed modelled snow depths in southern Norway with decimeter-scale accuracy. In the western part of Hardangervidda, ICESat-based March snow depths agree better with measurements (RMSE ≤ 0.15 m for all DEMs) than modelled snow depths do (RMSE 0.61 m). In eastern Hardangervidda, the coarse resolution SRTM DEM (RMSE 0.41 m) performs better than the 10 m Norwegian DEM (RMSE 0.64 m) which is based on a less consistent mosaic of elevation data. Using the high-resolution lidar DEM, even single footprints show good agreement (R^2 0.59, RMSE 0.94 m) with measured snow depths from the same year. Snow depth estimates could be further improved by using full waveform ICESat data or elevation measurements from ICESat-2 once this satellite is operational. Good quality reference DEMs may still be acquired in the future even in areas where no such data exists today.

Keywords:

ICESat, altimetry, snow depth

1. Introduction

Snow plays a key role for the hydrology, ecology and energy balance of large areas on Earth. In Norway, snow is

directly important for society e.g. through its role in hydropower production and winter recreation, but it also bears risks for floods or avalanches and affects transportation and accessibility (Engeset et al., 2004). Snow and its spatial distribution govern cryospheric processes such as glacier mass balance, permafrost distribution and surface energy balance, and also determine mountain ecology and habitats (Dietz et al., 2012). Consequently, snow is a driving

*Corresponding author

Email addresses: desiree.treichler@geo.uio.no (D. Treichler), kaeab@geo.uio.no (A. Kääb)

factor in models for many different aspects of our environment, worldwide. Yet, the amount of snow in remote areas is not well known due to a lack of measurements. Its estimation in mountainous terrain is possibly the most important unsolved problem in snow hydrology (Dozier et al., 2016; Lettenmaier et al., 2015). In rough terrain the snow cover varies highly both in space and with elevation due to orographic effects, and wind-driven redistribution in combination with small-scale topography. Spatially distributed information on snow over large areas is typically a product of modelling or interpolation based on sparse in-situ measurements, remotely sensed data, or climate reanalyses. The former method is how Norwegian national snow maps are produced: a snow model is forced by precipitation and temperature maps that are interpolated from spatially distributed, point-based station measurements (Saloranta, 2012).

Different metrics are used to express the amount of snow on the ground: the extent of the snow-covered area, snow depth and snow density, and snow water equivalent (SWE). The parameter of choice depends on the application (e.g. SWE for hydrological studies, snow-covered area for albedo, snow depth for heat transfer). On per-point basis, SWE can be measured in an automated way by snow gauges on meteorological stations, or snow pillows. Similarly, local snow depths are measured automatically with the use of e.g. a sonic ranger mounted on a meteorological station. On the scale of small catchments, SWE measurements are traditionally done manually, following a standardised protocol (snow course) where snow depth and density are measured at multiple, pre-defined locations several times throughout the winter. Direct measurements from meteorological stations or snow courses are mainly available for inhabited areas at lower elevations (stations) and/or catchments with hydropower generation (snow courses) and may be unrepresentative of conditions at higher altitude (e.g. Rasmussen, 2013; Dozier et al., 2016).

For larger areas, remote sensing techniques are used to derive snow metrics in a spatially distributed way. Terrestrial or airborne lidar (light detection and ranging) systems are currently the best providers of snow depths and distribution at catchment scale with decimeter-scale vertical accuracies (Deems et al., 2013) — but they require specifically planned campaigns and have considerable costs (airborne systems). Increasingly, camera systems using structure from motion (SfM) photogrammetry techniques are used to map snow depth and snow cover area (e.g. Nolan et al., 2015; Vander Jagt et al., 2015). Photogrammetric techniques require good image contrast which is not necessarily given on bright snow-covered areas, and advances in their application for snow depth retrieval with decimeter-scale accuracy have been made only recently (Bühler et al., 2015). A recent study by Marti et al. (2016) shows that also very-high-resolution optical satellite stereo imagery may be used for snow depth measurements at catchment scale — the swath width of such sensors is below 20 km. The method is limited to areas with significant snow accumulation due to decimet-

ric systematic and random errors.

Optical space-borne remote sensing data is commonly used to map snow cover extent at coarser spatial resolutions (Dietz et al., 2012; Lettenmaier et al., 2015). Numerous satellite-derived snow cover maps are available as ready-made products (e.g. the MODIS global snow cover products, Hall et al., 2002). However, optical satellite imagery does not provide information on snow depth and SWE of the snow-covered area. SWE has been successfully estimated from passive microwave sensors for dry snow packs <1 m (Dietz et al., 2012; Clifford, 2010). Unfortunately, products from passive microwave sensors are only available at a coarse spatial resolution of tens of kilometres which leaves large uncertainties at regional scales, in particular in mountainous regions (Rango, 1994; Vikhamar and Solberg, 2003; Lettenmaier et al., 2015). For applications where precipitation amounts are important, reanalysis data provides plausible values also at high altitudes (Rasmussen, 2013; Immerzeel et al., 2015) but only at even coarser spatial resolutions.

To the best of our knowledge, no method is available so far to measure snow depth on regional scales and from space. For certain applications or model assimilations, direct regional snow depth measurements would be useful. Additionally, the combination of regional snow depth measurements with above methods for retrieval of, for instance, snow cover extent or SWE could open up important synergies (Dozier et al., 2016; Lettenmaier et al., 2015).

In contrast to above methods or models, the NASA Ice, Cloud and land Elevation Satellite (ICESat) directly measured surface elevations during 18 campaigns (table 1) between 2003 and 2009 in northern autumn, winter and late spring, and thus has the theoretical potential to provide information on snow depths. Data from ICESat’s Geoscience Laser Altimeter System (GLAS) consists of point samples of surface elevation along near-repeat ground tracks. By using double-differencing techniques, i.e. comparing ICESat elevations with a reference Digital Elevation Model (DEM) and their subsequent evolution over time, Kääb et al. (2012) successfully detected glacier surface elevation change trends even in the rough Himalaya Mountains. The method relies on spatial grouping of samples to average out uncertainties, and has subsequently been applied successfully to other glacierised mountain areas and even globally (Kääb et al., 2015; Ke et al., 2015; Kropáček et al., 2014; Neckel et al., 2014; Treichler and Kääb, 2016; Gardner et al., 2013). The limits of this method were found to be the elevation accuracy of the reference DEMs rather than ICESat’s vertical accuracy (Treichler and Kääb, 2016). On Norwegian glaciers, Treichler and Kääb (2016) found clear indications of yearly varying snow depths that correspond well with measured winter mass balances in the area. On the Antarctic ice sheet, Bindenschadler et al. (2005) used elevation differences of intersecting ICESat orbits before and after a snow fall event to validate snow event detection from microwave data. To the best of our knowledge, explorations of seasonal ICESat-derived

elevation variations to analyse snow depths have not been published, even though a few authors mention this possibility (Jasinski and Stoll, 2012; Jasinski and Neumann, 2013; Stoll and Jasinski, 2012; Fassnacht et al., 2012).

In this study, we extend the glacier analysis of Treichler and Kääb (2016) to assess ICESat’s potential in providing snow depths also on the non-glacierised areas of the mountains in southern Norway. Most of the study area lies above the tree line and thus at elevations where snow depths are of special interest for e.g. hydropower generation (Saloranta, 2012). It is characterised by an east–west increase in precipitation due to predominantly western winds and the Scandinavian mountains acting as an orographic barrier. In the southern part of the area lies Hardangervidda, a high altitude plateau that covers approximately 4500 km² (fig. 1). Snow heights and distribution on Hardangervidda are comparatively well known, not least due to a recent study using high-resolution airborne lidar data (Melvold and Skaugen, 2013). The distinctive patterns of snow depths in southern Norway — increasing both with elevation and coastal proximity — make this region, and especially Hardangervidda, a suitable test area for ICESat-derived snow depths in mountainous terrain. In this study, we want to investigate the following questions in particular:

- Is the ICESat track density sufficient to provide representative results for the varied terrain in the study region, and even high enough to resolve spatial snow depth patterns?
- Is it possible to see snow cover variations over time and quantify snow depth with ICESat?
- How do the findings compare to snow depth measurements and model results?
- How does the reference DEM influence the quality of the results?

2. Data

2.1. ICESat elevation data

ICESat measured the Earth’s surface elevation in two to three campaigns per year from 2003 to 2009 (table 1). The campaigns were flown in autumn (ca. October–November, snow-free conditions over the study area except for very high elevations), winter (ca. March, snow-covered), and early summer (ca. June, only 2004–2006, mostly snow-free except for high elevations). Campaign start and end dates differ by up to several weeks for different years – we refer to ‘autumn’, ‘March’ and ‘June’ to distinguish the three seasons. The ICESat data consists of point samples that are separated by ca. 170 m along profile lines/orbits and ca. 40–47 km between orbits in southern Norway (fig. 1). Note that the shape of the study area is not square (coverage shown in red in the overview map), and that there was a transition between two different orbit patterns in the middle of campaign 2A (Schutz et al., 2005) which results in odd spacing between repeated tracks and orbits served during

the first two campaigns only (1AB/2A). One ICESat elevation sample corresponds to a ground footprint of ca. 70 m in diameter (Schutz et al., 2005). We used GLAS/ICESat L2 Global Land Surface Altimetry HDF5 data (GLAH14, release 33) which is optimised for land surfaces (Zwally et al., 2012). GLAH14 data was not affected by G-C offsets and elevation values have thus not changed between releases 33 and 34 (NSIDC, 2014; Borsa et al., 2014). In this data product, up to 6 Gaussian curves are fitted through the laser received waveform to find the mean surface elevation of the (potentially rough) terrain within the footprint area (NSIDC, 2014).

A direct comparison between ICESat elevations is difficult in mountainous terrain as ground orbits vary up to ± 150 m in space between overpasses in southern Norway. From comparison with reference DEMs, elevation uncertainty of GLAH14 data was found to be in the order of decimetres to metres in rough terrain (Treichler and Kääb, 2016). Dense vegetation cover contributes to the received laser waveform and may result in raised surface elevation values (Neuenchwander et al., 2008). Elevation biases and inconsistencies throughout ICESat’s lifetime, such as correction for waveform saturation and the optional range increment for land samples (attributes `d_satElevCorr` and `d_ldRngOff` in the GLAH14 dataset) or intercampaign bias (e.g. Hofton et al., 2013; Siegfried et al., 2011), are of centimetre to decimetre magnitude and found to be negligible compared to biases in the reference DEMs in southern Norway (Treichler and Kääb, 2016). Performance decline and differences between ICESat’s three laser systems result in lower sample numbers especially from December 2008 on (campaigns 2D–2F). Borsa et al. (2014) suggest the low transmit energy of the latter campaigns slightly increases the uncertainty of the elevations extracted from single peak, smooth return waveforms on ice sheets where different elevation extraction algorithms are used than for the GLAH14 product. No such studies exist for rough surfaces with complex return waveforms that likely prohibit the detection of potential performance-induced differences.

Before filtering, there are approximately 170’000 raw footprints in the study area, though sample numbers per campaign vary – they are larger in the beginning of the acquisition period and decrease over time (table 1) due to weakening of the laser before complete failure in 2009. The dataset used in this study has previously been used to assess the potential of ICESat data for glacier surface elevation change trends in southern Norway (Treichler and Kääb, 2016), and further details and data parameters are given there.

2.2. Reference surface elevations

Reference surface elevations from three DEMs were used: a) the Norwegian national DEM with 10 m spatial resolution, further referred to as Kartverket DEM, available for the entire study area from <http://data.kartverket.no>; b) the DEM from the Shuttle Radar Topography Mission (SRTM, Farr and Kobrick, 2000) for latitudes below 60.3° N, i.e. as far north as Hardangervidda. We used

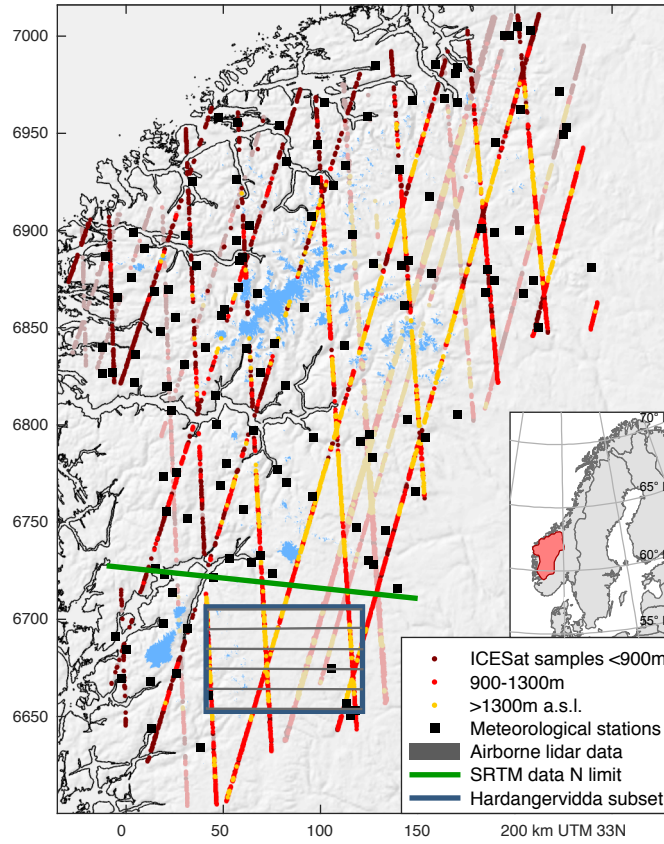


Figure 1: The study area in southern Norway (outlined in red in the overview map). Samples used in this study (after filtering) are coloured according to elevation above sea level. Greyed out orbits were only served in winter/autumn 2003 (no samples in 2004–2009). The Hardangervidda subset includes all samples within the area outlined in dark blue. Reference elevations from the SRTM DEM are available south of the green line, those from the lidar DEM only where the six east–west oriented stripes of data (grey) coincide with ICESat ground tracks. Glaciers are indicated in light blue.

Table 1: ICESat sample numbers and operational periods in southern Norway, and dates of lidar reference snow depth measurements. Modelled snow depths are available for every day of the ICESat operational periods. SN stands for Southern Norway, i.e. the entire study area, HV for the Hardangervidda subset. Sample numbers are after filtering and smaller for the lidar/SRTM reference DEM datasets due to the limited spatial extent of the lidar stripes and data gaps in the SRTM DEM respectively. The campaign lengths in days (# d) are slightly shorter for HV as not all SN orbits cross that spatial subset.

Laser	Dataset	Sample # SN	HV	SRTM	lidar	first day SN	last day SN	# d	HV
1AB	March	4901	512	512	28	24 Feb 2003	29 Mar 2003	34	32
2A	Autumn	7298	626	625	33	28 Sep 2003	17 Nov 2003	51	50
2B	March	3866	541	541	29	23 Feb 2004	20 Mar 2004	27	26
2C	June	2135	553	551	32	24 May 2004	19 Jun 2004	27	26
3A	Autumn	2548	225	225	10	09 Oct 2004	05 Nov 2004	28	27
3B	March	4684	652	651	39	17 Feb 2005	23 Mar 2005	35	34
3C	June	2036	373	373	17	26 May 2005	20 Jun 2005	26	9
3D	Autumn	1246	124	124	5	27 Oct 2005	22 Nov 2005	27	26
3E	March	4102	446	446	23	28 Feb 2006	26 Mar 2006	27	17
3F	June	2692	568	568	27	30 May 2006	25 Jun 2006	27	26
3G	Autumn	2354	401	401	28	31 Oct 2006	26 Nov 2006	27	16
3H	March	2761	613	611	31	17 Mar 2007	13 Apr 2007	28	27
3I	Autumn	3251	583	583	27	08 Oct 2007	03 Nov 2007	27	26
3J	March	1943	222	222	13	23 Feb 2008	20 Mar 2008	27	9
	April 2008 lidar DEM					3 Apr 2008	21 Apr 2008		19
	lidar reference DEM					21 Sep 2008			1
3K	Autumn	716	226	226	9	09 Oct 2008	18 Oct 2008	10	9
2D	Autumn	1781	173	173	11	28 Nov 2008	16 Dec 2008	19	8
2E	March	2005	362	362	14	14 Mar 2009	09 Apr 2009	27	17
	April 2009 lidar DEM					21 Apr 2009	23 Apr 2009		4
2F	Autumn	444	178	178	11	06 Oct 2009	07 Oct 2009	2	1
	Autumn, total	19638	2536	2535	134				
	March, total	24262	3348	3345	177				
	June, total	6863	1494	1492	76				

the SRTM non-filled DEM version at 3 arc-seconds resolution (corresponding to 93 m in y, and 45 m in x-direction at 60° N; the 1 arc version does not reach further than 60° N) which is accessible from the U.S. Geological Survey at <https://dds.cr.usgs.gov/srtm/>; and c) a lidar DEM at 2 m spatial resolution for six east-west oriented stripes crossing Hardangervidda. The lidar DEM is provided by the Norwegian Water Resources and Energy Directorate (NVE) and consists of 80 km long and 500 m wide stripes. The lidar data was acquired for a study on snow heights on Hardangervidda (Melvold and Skaugen, 2013), with campaigns on 21 September 2008 (minimum snow cover conditions), 3–21 April 2008, and 21–24 April 2009 (maximum snow accumulation). The dataset thus also includes gridded snow depth data at 2 m spatial resolution, resulting from elevation differences between the April datasets and the September reference.

DEM vertical accuracies (table 2) are on the order of centimeters for the lidar DEM (based on a kinematic GPS survey, Melvold and Skaugen, 2013), and on the order of metres for the SRTM and Kartverket DEMs (Rodriguez et al., 2006; Kartverket, 2016). Treichler and Kääb (2016) analysed elevation differences (further referred to as dh) between ICESat and DEM elevations of all three DEMs (after coregistration and bias correction, see section 3) of ICESat’s autumn samples in southern Norway. The results from that study are summarised in table 2 and confirm that the elevation errors in southern Norway are well within the accuracies stated by the data providers. Note that the empirical error estimates in table 2 comprise both DEM and ICESat elevation uncertainties (the normalised median absolute deviation/ dh (NMAD) is an estimate for the standard deviation of DEM differences that is resilient to outliers; Höhle and Höhle, 2009). The higher empirical elevation error (0.7 m) found for the lidar data (nominal accuracy: 0.12 m) can therefore be attributed to uncertainties from ICESat’s footprint size and the terrain therein. Systematic vertical biases in the coarse resolution reference DEMs cause uncertainties that are one magnitude larger than ICESat’s elevation uncertainties. The lower accuracy of the coarser resolution DEMs is caused by shifts between ICESat as a (globally) consistent elevation reference on the one hand, and unknown sub-units of the DEMs on the other hand. Such locally varying shifts are common in national/global DEMs that are a mosaic of multiple different elevation datasets. The horizontal and vertical shifts cause non-systematic vertical offsets between ICESat and the reference DEMs that increase uncertainty of derived parameters, such as snow depths. On Hardangervidda, the Kartverket DEM is likely based on old elevation data (ca. 1970, estimated from contour line timestamps). The older technology and methods behind the elevation source data may be the cause for the rather large dh between ICESat and the Kartverket DEM especially in eastern Hardangervidda. A detailed discussion of the elevation biases of the three reference DEMs and their effect on ICESat applications in southern Norway is given in Treichler and Kääb (2016).

2.3. Reference snow depth data

Reference snow depths for the ICESat era in southern Norway include time series from meteorological stations, modelled snow depths, and the lidar datasets on Hardangervidda for the winters 2008/2009. We used data from the 130 meteorological stations within the study area that directly measure snow depths. The data are provided by the Norwegian Meteorological Institute and accessible from www.eklima.no at daily temporal resolution. The stations are located mostly at lower elevations and up to 1140 m a.s.l and thus do not capture snow depths at higher elevations well.

Modelled snow depths for the entire study area are available from the seNorge snow model (Mohr, 2008; Saloranta, 2012). The seNorge snow depth dataset (v.1.1.1) is based on observed temperatures, precipitations and snow depths, and generated using a degree-day melt model (Saloranta, 2014). The maps are available daily (since 1957) for mainland Norway as interpolated, gridded datasets at 1 km spatial resolution, and are published at www.senorge.no. Due to the high spatial and temporal variability of snow depths in Norway, which is only partly captured by traditional, point-based measurements such as from meteorological data, the seNorge data is likely the best representation of the snow conditions in southern Norway currently available (Saloranta, 2012). Based on data from meteorological stations, snow pillows and manual SWE measurements, Stranden (2010) and Saloranta (2012) found seNorge snow depths to agree well with measurements at lower elevations east of the mountains but to overestimate snow depths on the Scandinavian mountains and at higher elevations. They attribute this bias to a rather serious overestimation of both SWE and snow pack density in the model – where the latter partly compensates the SWE surplus when recalculated to snow depths. The source of the overestimation lies likely in the model’s imperfect parameterisation of precipitation at high elevations due to too little model forcing data. Compared to the Hardangervidda lidar DEMs, also Melvold and Skaugen (2013) found seNorge to overestimate snow depths in particular in northwestern Hardangervidda, likely due to the lack of meteorological observations for interpolation/extrapolation of temperature and precipitation data in particular in this area.

The extensive Hardangervidda lidar survey is a representative measurement of 2008/2009 snow depths in this area (Melvold and Skaugen, 2013). In average, 2008/2009 snow depths represent long term snow depths at all elevations well (seNorge 08/09 average March snow depths exceed the 03–09 March average by ca. 0.1 and 0.2 m in western and eastern Hardangervidda respectively). The survey was not designed to fit the timing of ICESat winter campaigns (table 1) but to capture maximum winter snow depths (mid- late April; Melvold and Skaugen, 2013). This leaves a time difference of approximately one month to ICESat-derived snow depths. Based on the seNorge snow depth data archive, the average increase between March and April snow depths is only 0–0.13 m on Hardangervidda – except for areas west of

Table 2: Accuracies and specification of reference DEMs from data providers/literature and from comparison to ICESat elevations. Values from other sources are recomputed to 1 standard deviation (std). The normalised median absolute deviation (NMAD, Höhle and Höhle, 2009) for autumn samples after filtering is given for southern Norway (SN; 19638, 3829, and 134 samples for the SRTM/Kartverket/lidar DEMs) / the 134 autumn samples with data from all three reference DEMs on Hardangervidda (HV) respectively.

DEM (resolution)	Extent / data based on	elevation accuracy	corrected bias	NMAD	
				SN	HV
SRTM (45/93 m)	south of 6 720 000 N, >4 C-band radar scenes from Feb 2000	3.9 m (absolute error, 1 std)	20 m horizontal / 2.6 m vertical shift	2.6 m	1.8 m
Kartverket (10 m)	Norway; map data from ca. 1961–2011 at 1:50 000 scale	4–6 m (1 std)	coregistration, elevation-dependent bias	2.4 m	3.4 m
lidar (2 m)	Hardangervidda, six stripes of 80 km / 500 m. Lidar flights on 21 Sep 2008 (no snow), 3–21 Apr 2008, and 21–24 Apr 2009 (snow depth)	0.12 m (1 st. error), range: -0.95–0.51 m	—	0.7 m	0.7 m

60 km E where differences rapidly increase to up to 0.4 m. Spatial distribution of snow depths is unlikely to change in the course of a month: Melvold and Skaugen (2013) found interannual snow distribution patterns on Hardangervidda to be strongly consistent.

2.4. Map data

We used map data from the Norwegian Mapping Authority (Kartverket), the Norwegian Institute of Bioeconomy Research (NIBIO), and NVE’s glacier outlines (NVE, 2016) to classify ICESat footprints based on land cover. Kartverket’s N50 data (Norway’s topographical maps at 1:50 000 scale) are available from www.geonorge.no. The AR50 data from NIBIO can be accessed from www.nibio.no and contain information on land cover at ca. 1:50 000 scale. AR50 data are based on detailed land registry maps and, in mountain areas, on semi-automatic classification of satellite imagery (Aune-Lundberg and Strand, 2011).

3. Methods

ICESat–DEM elevation differences of single footprints in our mountainous study area contain larger uncertainties than is the case for footprints on flat surfaces such as ice sheets. The reason for the increased uncertainty lies in the within-footprint topography, in addition to uncertainties of all reference elevation datasets described above. As with ICESat studies on mountain glaciers, data pre-processing is required to remove biases and uncertainties as far as possible. The pre-processing steps for the datasets we use in this study are described in further detail in Treichler and Kääb (2016). We classified the ICESat samples based on N50 and AR50 map data and glacier outlines using footprint centre locations. The elevation datasets were co-registered using manual and automated co-registration techniques (Nuth and Kääb, 2011) to remove horizontal and vertical shifts between ICESat (snow-free, filtered samples only) and the three reference DEMs. Reference elevations were extracted from the DEMs by taking the median elevation of all grid cells within 35 m radius from each footprint centre from the Kartverket DEM, and bilinear interpolation of four adjacent

SRTM grid elevation values respectively. To reflect the distribution of terrain elevations and snow depths within single ICESat footprints, we extracted the median and 0.25/0.75 quantiles of elevations/snow depths for the footprints with lidar reference elevations. Two footprint representations were assumed: a) a 70 m circular footprint (the average shape), and b) the individually shaped elliptic footprint of each sample. Information on footprint transmit pulse azimuth, eccentricity, and major axis are metadata parameters of the GLAH14 data for all but the last three campaigns (2D–2F). For these, we used the campaign averages given in NSIDC (2016). The footprints of the 2008/2009 winter campaigns (3J, 2E) are 58–60 m in diameter and nearly circular.

The elevation differences between ICESat data and the reference DEM were used to correct a systematic, elevation-dependent bias of the Kartverket DEM in the order of a decimetre per 100 m elevation gain. dh are positive if the surface is covered by snow but should be zero for snow-free samples over stable ground. ICESat samples were filtered to remove footprints on non-stable surfaces (water, glaciers; ca. 11%), slopes $>10^\circ$ (ca. 43%), clouds (16%) and with gross measurement errors ($<0.1\%$, large dh for samples on e.g. vertical fjord walls). We used reference data from NVE and Kartverket, and a threshold maximum elevation difference based on the Kartverket DEM reference elevations ($|dh| < 100$ m) for filtering. The restriction to relatively flat terrain was done to reduce uncertainty from decreased ICESat elevation accuracy on steep slopes, as well as larger potential bias from unreal dh caused by horizontal shifts between the datasets (Nuth and Kääb, 2011). We found no indication of campaign-inherent biases such as from inter-campaign bias (Schutz et al., 2011) or the differently shaped footprints of individual campaigns (Abshire et al., 2005). After the above filtering, 50 763 samples remain (7 372 on Hardangervidda). Most of our study area is located above the tree line and the majority of samples fall on areas with sparse alpine vegetation cover (63%, 86% within the Hardangervidda area) and only few on forested ground (12%/5%). The remaining samples fall on non-vegetated terrain (13%/3%), mires and agriculturally used land. All vegetation types occur predominantly within certain elevation bands. Samples of

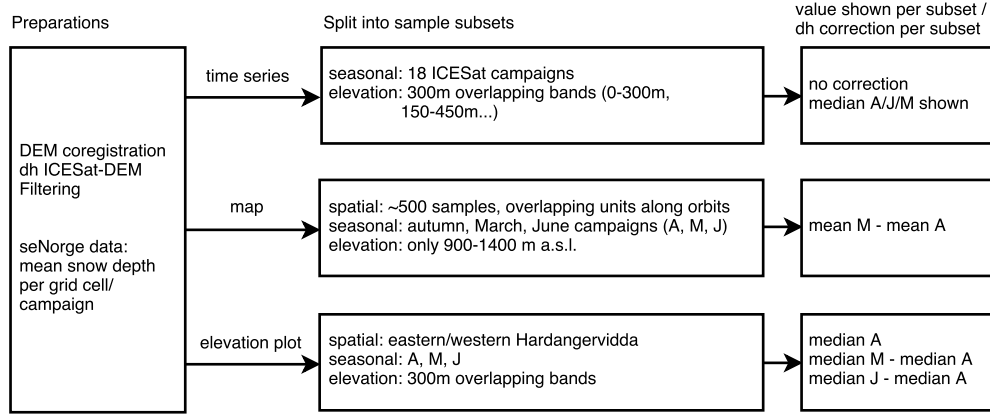


Figure 2: Workflow for ICESat sample subsetting and autumn dh offset correction per sample subset. The same workflow was applied to all three DEMs and also to the seNorge data where it is appropriate. A, J, M stand for autumn, June, and March respectively.

individual campaigns were found to be representative and consistent in how they sample terrain parameters such as elevation, slope, or aspect in the study area. Only two campaigns deviate from this: the 2003 winter campaign (laser 1AB, operated in a different orbit pattern; Schutz et al., 2005) and the 2009 autumn campaign (laser 2F, last campaign before complete instrument failure). Only three and two orbit tracks over our study region were covered by these two campaigns, respectively, while the spatial sampling distribution is consistent for all other campaigns. Campaign sample numbers are 450–7’300 for the entire study area, and 120–650 for Hardangervidda area (table 1).

The dh of snow-free samples in southern Norway are t-distributed and grouping of several samples was found to average out noise and errors. To reflect the main influencing parameters for snow depths in southern Norway (orographic effects and coastal proximity), we grouped samples in (i) elevation bands, (ii) spatial subsets, or (iii) over time to get average snow depths for ‘March’, ‘June’ and ‘autumn’ campaigns. Owing to the mosaicking from multiple source datasets for the production of spatially extensive DEMs such as the Kartverket and SRTM DEMs, spatial subsets smaller than a DEM tile are prone to be affected by remaining, spatially varying elevation biases of unknown spatial units. Such local shifts are not removed by global co-registration on e.g. DEM tile basis, as done here. In the glacier surface elevation change study described in Treichler and Kääb (2016), spatially varying elevation biases were effectively removed by the application of a per-glacier dh correction term. There, the correction is based on the median dh of snow-free autumn samples on each glacier. This corresponds to a local vertical co-registration of the spatial subsets (single glaciers) in the data. Here, we use an adaptation of this correction and correct ICESat March/June snow depths for vertical bias by removing average autumn dh of the same subset, where appropriate. In that case, the resulting snow depth estimates are bias-corrected to October-November surface elevation — or, in other words, refer to

the difference between June/March snow depths and autumn snow depths. The work flow in figure 2 shows how the samples were grouped and bias-corrected for the analyses presented in the results.

4. Results

Figure 3 (middle panel) shows time series of all samples after filtering, grouped into overlapping elevation bands of 300 m. Data from the entire study area are used. For each elevation band, median elevation differences between ICESat and the Kartverket DEM are shown. The change in surface elevation due to snow cover is clearly visible from the pattern of increasing and then decreasing dh between autumn and March campaigns. Snow depths increase with elevation to up to ca. 2 m at 1800 m a.s.l., but vary throughout the satellite’s operational period, with maximum snow depths in 2005 and 2007. For comparison, the lower and upper panel in fig. 3 show the medians (per elevation band) of modelled snow depth grid cells and snow depths from meteorological stations respectively. To summarise snow depth evolution during ICESat campaigns, seNorge/station data time series were averaged to one value per grid cell/station and ICESat campaign. Both ICESat and seNorge suggest consistently larger snow depths than data from meteorological stations above 500 m a.s.l. The latter are mainly located east of the main mountain ridge and experience less snow than coastal regions. Overall, the ICESat-derived snow depth pattern agrees well with observed (R^2 0.57) and modelled (R^2 0.65) snow depths. The root mean squared errors (RMSE, values based on the medians per elevation band/campaign) are 0.47 m and 0.61 m respectively. For comparison, the coefficient of determination between seNorge and station data (summarised as described above) is 0.74, with an RMS error of 0.38 m. Correlation between dh and seNorge snow depths is noticeably higher for March (R^2 0.71) and June campaigns (R^2 0.66) than for autumn campaigns (R^2 0.32). Also, dh of the middle elevation bands (600–1650 m) agree considerably better with

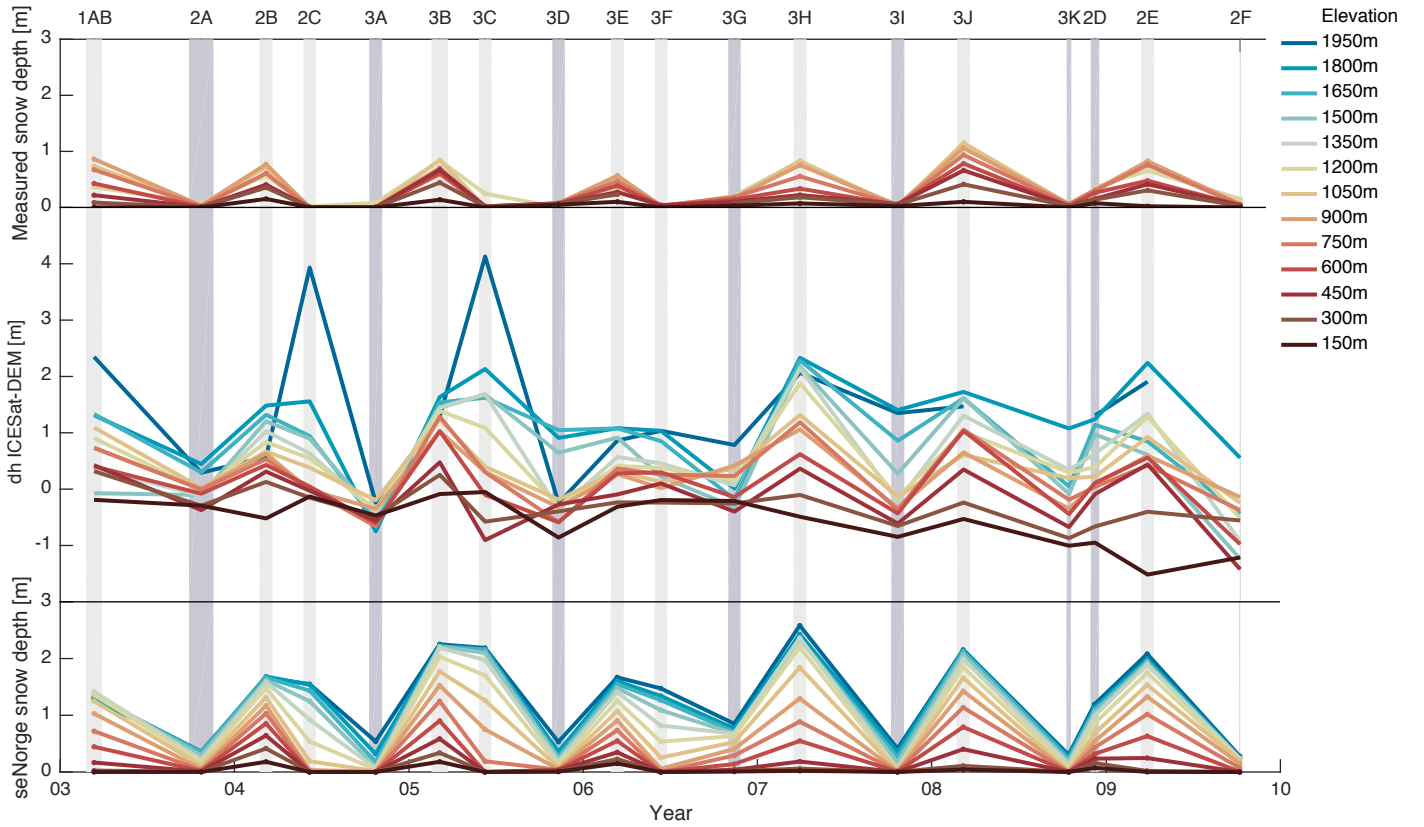


Figure 3: Snow depth time series from meteorological stations (top), ICESat (middle), and the seNorge model (bottom). Shown are the median values of observations/samples/grid cells within the entire study area, split into overlapping 300 m elevation bands (value ± 150 m). ICESat campaign times are indicated as dark (autumn) and light (March/June) grey stripes. Station/model daily time series were averaged over the ICESat campaigns for direct comparability (i.e. reference snow depth data between ICESat campaigns were excluded).

both reference datasets (R^2 between 0.62 and 0.87) than it is the case for the highest/lowest elevations. The latter have lower ICESat sample numbers (<3000 per band, all campaigns summarised; see table S1 in the supplement). F-statistics indicate a significant linear relationship with the reference datasets for all but the lowest/highest elevation band ($p < 0.001$ for 600–1800 m, $p < 0.05$ for 300/450 m).

ICESat’s June values match March snow depths above ca. 1700 m but rapidly decrease at lower elevations. Still, June dh tend to exceed autumn dh in the order of decimetres even for elevations where no snow is expected in late spring but where deciduous shrubs and trees grow. Between the November and December campaigns in 2008, the onset of snowfall is clearly visible at all elevations, with dh increasing in the range of decimetres. Potential systematic inconsistencies between ICESat’s 3K and 2D campaigns are in the order of centimetres only and thus can’t explain the December dh increase. Despite the large sample numbers per campaign, campaign means/medians per 300 m overlapping elevation bands vary considerably throughout time and include also negative values. This effect is also visible for autumn campaigns with no or minimal snow cover, i.e. where no elevation differences are expected between ICESat and the reference DEM. Due to cloud cover and weakening of the laser over time, the

pattern of acquired orbits as well as the sample distribution within orbits varies. Still, single campaigns were found to be representative in terms of spatial distribution and when compared to the slope, aspect, elevation and land cover of the terrain in southern Norway. We found no indications that any land cover type would result in larger absolute dh or larger dh uncertainty, including forested areas. Vegetation type/land cover correlate too strongly with elevation to analyse the influence of vegetation on snow depths or bias in retrieved dh in further detail in our study area. The unknown spatial distribution of remaining ICESat–DEM offsets, however, interferes with the varying sample spatial distribution. The bias pattern over time indicates that the DEM corrections do not capture these offsets well enough to treat entire southern Norway as one consistent spatial unit. Bias in median dh change over time due to the varying composition of differently offset sub-units of the reference DEM.

Figure 4 (left panel) shows the spatial pattern of March snow depths based on samples at elevations between 900 and 1300 m above sea level (bright red in fig. 1). The right panel of figure 4 shows the corresponding modelled/measured average March snow depths from 2003–2009. ICESat samples were grouped in time (all samples from ‘March’ campaigns together) and spatially – each circle in fig. 4 corresponds

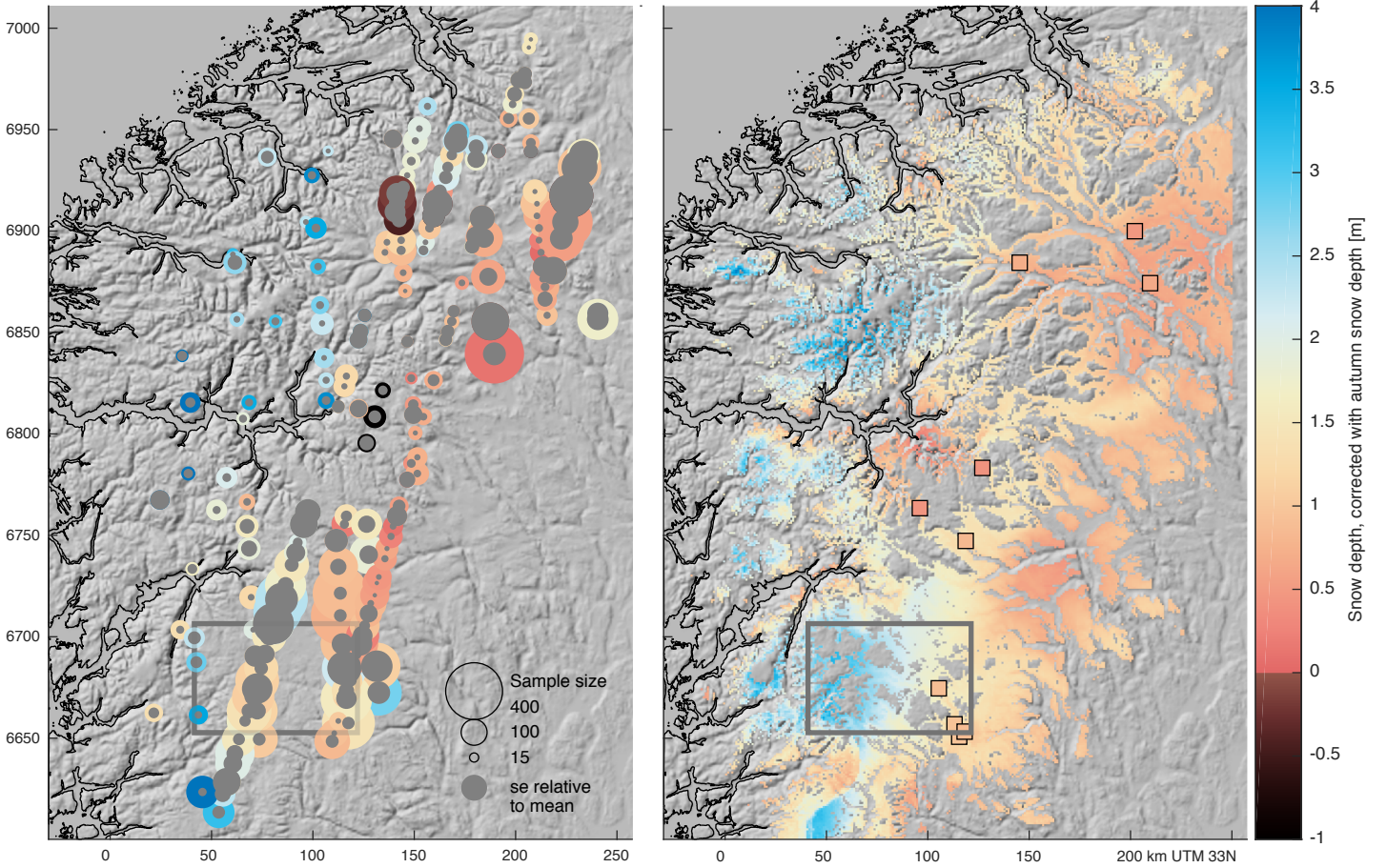


Figure 4: Spatial pattern of March snow depths between 900 and 1300 m a.s.l. in southern Norway. Left panel: snow depths derived from ICESat March campaign values compared to reference elevations from the Kartverket DEM (dh), corrected for local vertical bias with autumn (snow-free) dh. Each circle corresponds to the mean dh of a group of samples (circle size varies with sample numbers). The overlaid grey circles show the standard error (se) of the mean in relation to the mean value, i.e. if the grey circle fully covers the underlying colour, mean dh are not significantly different from zero (1σ). Right panel: seNorge modelled average March snow depths (background) and measured average March snow depths from meteorological stations (squares). Data is from 2003-2009 for grid cells/stations within the same elevation range. Average 2003-2009 autumn snow depths (mid October-mid November) are deducted for comparability with ICESat dh. The Hardangervidda subset is outlined in grey.

to 17–436 March samples (and also >15 autumn samples). Circles are scaled based on March sample numbers. Mean autumn dh per spatial subset were found to vary in the order of decimetres to metres. To coregister the elevation datasets locally, this spatially varying bias was removed from the March dh (see methods). Likewise, autumn snow depths were subtracted from seNorge model data (average values mid October–mid November, 2003–2009; see methods). In areas/elevations where snow is found in autumn already, the snow depths thus correspond to the difference between autumn dh (or snow depths) and March (or June) dh/snow depths, further referred to as difference in snow depths. We find a clear spatial pattern of increasing difference in March–autumn snow depths with coastal proximity, with up to 4 m snow in the westernmost parts. The values and their spatial pattern correspond well with the modelled/measured snow depths of the same period. The largest differences are found on Hardangervidda where the model data suggests considerably larger snow depths than ICESat. The standard error of the mean ICESat snow depths (overlaid grey circles,

standard error of mean autumn dh included in the value) is only in some cases larger than the mean value (where the coloured circle is not visible). For all other spatial subsets, ICESat-derived snow depths are significantly different from zero at 1σ confidence. This is also the case for a few circles with negative difference in snow depth/dh which is physically not plausible. The use of the median instead of mean per spatial unit results in fewer and less severely negative snow depth estimates. These occur spatially clumped (e.g. around 6920000N/1500000E), which could indicate poor DEM quality in this area, or they lie on orbits which were served in 2003 only, before ICESat’s orbit pattern changed. There, ICESat’s autumn elevation samples form a straight line (instead of a ca. 300m wide band for repeat orbits) which may have lead to poorer local vertical DEM coregistration in these areas. Sample numbers are smaller towards the coast where slopes generally exceed our slope threshold of 10° . The inclusion of samples with higher average footprint slope results in a comparable spatial pattern with more homogeneous size and spatial distribution of spatial

units (not shown). However, we find that increased uncertainty requires grouping of ca. twice as many samples per spatial unit to receive robust snow depth estimates when footprints with steeper slopes are included.

4.1. Snowdepths on Hardangervidda

Samples on Hardangervidda were also grouped per season, resulting in three datasets for autumn, March, and June dh. Per season and 300 m vertically overlapping elevation band, divided in eastern ($>85'000$ m easting in UTM 33N) and western Hardangervidda, sample numbers are typically 100–1000 (table S2 in the supplement) except above 1500 and below 900 m a.s.l. which mirrors the terrain elevation distribution. Only bands with at least 10 samples are shown. Sample numbers are one magnitude smaller for the lidar reference DEM (10–138). Vegetation cover is comparable for the entire Hardangervidda area (heathland with varying vegetation density) apart from samples above ca. 1500 m which are on bare ground. Forest samples occur only below the tree line ($< \text{ca. } 1000$ m). Figure 5 shows the median and 25/75% quantiles (shown as error bars) based on the three reference DEMs for all footprints per elevation bands/seasons/eastern and western Hardangervidda respectively. For comparison, modelled snow depths from seNorge’s 1 km grid cells as well as measured snow depths from the April 2008/2009 lidar campaign (March panel) are shown, too. For autumn samples (fig. 5, upper panel) with minimum snow cover, vertical offsets between ICESat and the reference DEMs are elevation-dependent and reach up to 1.5 m, in particular for the Kartverket DEM in eastern Hardangervidda. The offset is minimally positive for the lidar DEM. The difference between median dh from assuming constant, 70 m circular footprints or the individual, elliptic footprint shapes respectively was found to be negligible when several samples are grouped.

To remove the elevation-dependent biases in the reference DEMs, autumn median offsets per elevation bands were subtracted from March and June dh to render the datasets comparable (fig. 5, lower panels). The same was done for seNorge snow depths (autumn snow depths removed from June/March values). In the western part, autumn to June snow depth differences clearly increase with elevation to a maximum of ca. 1.5 m, independent of the reference DEM used. In eastern Hardangervidda, autumn to June snow depth differences are only slightly positive above 1200 m, but amount to about 0.5 m below 1200 m, if based on SRTM and Kartverket DEM. Using lidar reference elevations results in a constant offset of ca. -0.4 m dh difference, indicating June dh are lower than autumn dh, which is unlikely. This negative offset cannot be reproduced using Kartverket/SRTM DEM reference elevations on the lidar sample subset. seNorge June snow depths are ca. 0.5 m higher both for eastern and western Hardangervidda, at all elevations.

The same pattern as for June is visible for autumn to March snow depth differences. Both the east-west precipitation gradient and increase with elevation are well represented in March snow depths which amount to 1.5–2.5 m

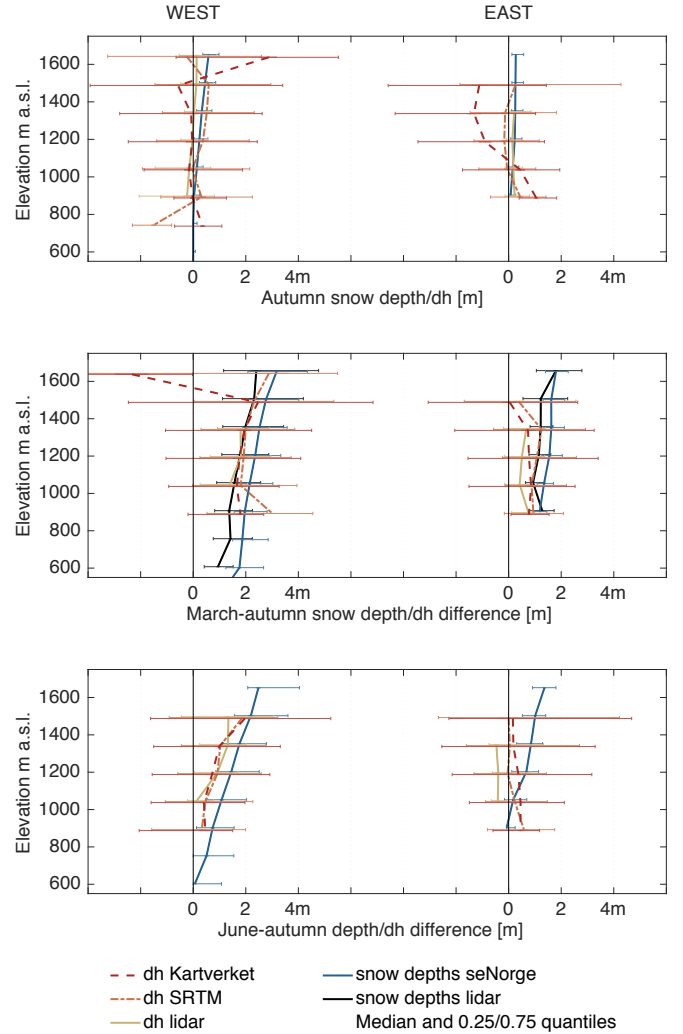


Figure 5: Snow depth vs elevation on Hardangervidda from ICESat and modelled/observed snow depths from 2003 to 2009. The plots show snow depths for three seasons: autumn (top), March (middle), and June (bottom). Data for each season is shown for two areal subsets, the western (left axis) and eastern (right axis) part of Hardangervidda area. Line types and colours distinguish the five different datasets used: dh ICESat-Kartverket/SRTM/lidar DEM (sample subset for lidar data), seNorge modelled snow depths from all grid cells within the Hardangervidda area, and measured winter snow depths (2008/2009) from lidar strips crossing the area. Elevation bands are 300 m wide and overlapping. June and March snow depths are bias-corrected with autumn snow depth/dh (i.e. autumn snow depth subtracted) of the same dataset and elevation band. The error bars indicate the 0.25/0.75 quantiles from the median of the samples/grid cells.

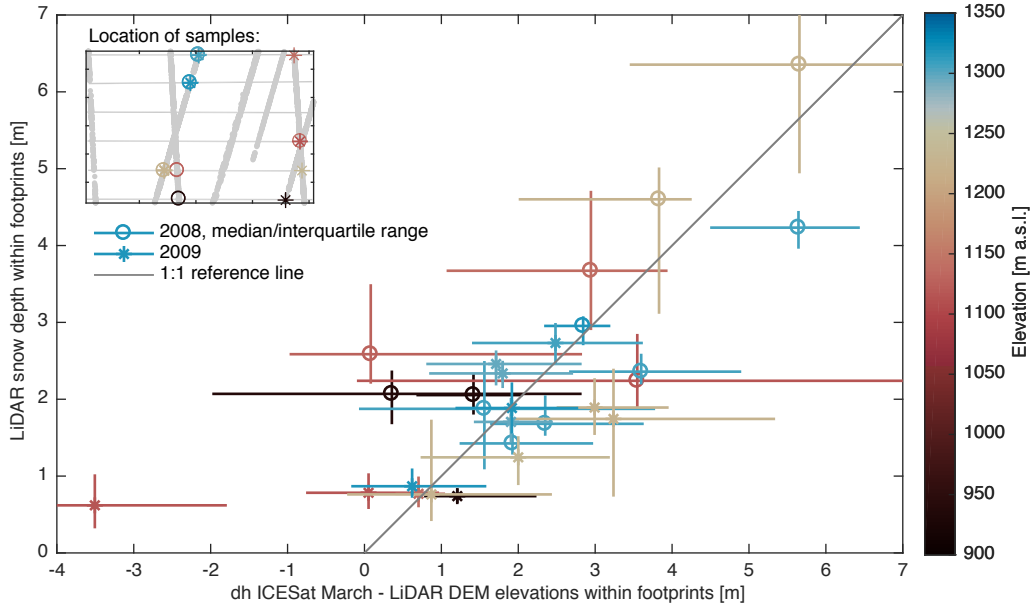


Figure 6: Within-footprint distribution of March 2008/2009 samples, showing elevation differences between ICESat and the snow-free lidar DEM (x-axis) vs lidar snow depths from April 2008/2009 (y-axis). Markers and error bars show the median dh/snow depth and 0.25/0.75 quantiles of the ca. 600 2x2 m reference elevation data points/snow depth measurements within each elliptic footprint. For comparison, the grey line shows the 1:1 reference line. The map inlet shows the location of the samples within the Hardangervidda area, distinguishing elevation (colour) and year (marker type).

in western and 0.5–1.5 m in eastern Hardangervidda respectively. In the western part, ICESat-estimated March snow depths agree for all reference DEMs used and also match snow depth measurements from April 2008/2009 (average of both years shown). For the middle four elevation bands which have >100 ICESat samples (>10 for lidar), the RMS errors between ICESat-derived and lidar snow depths are 0.09, 0.14 and 0.15 m for dh based on the Kartverket/SRTM/lidar DEMs respectively. Modelled differences in snow depths for the western Hardangervidda area are again ca. 0.5 m higher at all elevations. This is also the case in the eastern part, although estimates from the three reference DEMs differ considerably there. All shown elevation bands contain >100 ICESat samples (>10 for lidar). The differences in snow depth based on the low-resolution SRTM DEM agree substantially better with measurements (RMSE 0.4 m) than those based on the Kartverket DEM (RMSE 0.64 m). ICESat snow depth estimates based on the lidar DEM consistently underestimate measured 2008/2009 April snow depths by 0.5 m (RMSE 0.56 m) at all elevations on eastern Hardangervidda, an offset that is comparable to the -0.4 m bias visible in the June dataset. Interquartile ranges of all datasets increase with elevations, indicating larger variability in dh/snow depths at higher elevations.

4.2. Snow depth from single ICESat footprints

The dh ICESat–lidar DEM agree very well with measured snow depth even for single footprints for the 13 (2008) and 14 (2009) samples where ICESat March campaigns coincide with lidar snow depth measurements in April of the same years (fig. 6). seNorge modelled snow depths for the two

years suggest that changes in snow depths within the month-long time difference between ICESat and lidar data acquisitions are, at most, on the order of decimetres. 24 of the footprints fall entirely on areas with low/sparse vegetation (lichen, grass, heath), only the three footprints at around 930 m a.s.l. are just below the tree line and have sparse tree growth. The variability of dh and measured snow depths per sample is shown by interquartile ranges considering all single lidar reference DEM and snow depth grid cells within single footprints (elliptic shape, 450–700 valid cells each). The figure shows median dh against median snow depths with the interquartile ranges as error bars. Median dh correlate well with median measured snow depths. If the outlier with negative snow depth is excluded (footprint located on a slope between two narrow rock peaks), we receive an R^2 of 0.59 and RMSE of 0.94 m. Nash-Sutcliffe Model Efficiency is 0.46 (0.12 with outlier). If 70 m circular footprints are used instead, the error bars are 10% larger both in x and y direction (not shown) and correlation slightly poorer (R^2 0.52). Interquartile ranges within single footprints increase with increasing snow depth, and are larger for ICESat–DEM dh than for measured snow depths. The larger snow depths of 2008 (circles) as compared to 2009 (asterisk) are correctly represented. Within single years, an increase of snow depth with elevation can be seen from the 13 and 14 samples respectively.

5. Discussion

ICESat reproduces a realistic pattern of March and June snow depths in southern Norway. It shows snow amounts

to increase with elevation and coastal proximity and reproduces differences in snow depths for different years. Using ICESat data, we are able to estimate snow depths not least at elevations above the tree line where no or very few measurements from meteorological stations exist. However, in contrast to time series from meteorological stations or modelled snow depths, ICESat data can at most provide one single value each year around the months of March and June (only 2004 to 2006) respectively. High amount of noise due to uncertainties and biases in reference elevations, in combination with spatially varying sample distribution, can make it necessary to bin autumn/March/June dh rather than assessing time series. The resulting loss of temporal variability is compensated by a more robust estimate of average March/June snow depths and the possibility of finer spatial resolution. Yet, if snow depths vary considerably from year to year, one has to keep in mind the decrease of sample numbers over time due to weakening of the ICESat laser power. ICESat-derived March/June snow depths are thus biased towards snow depths at the beginning of the satellite’s acquisition period rather than later.

Our snow depth estimates agree on decimetre-scale with average snow depth values from the few meteorological stations at similar elevations in the study area, and they match the lidar-based snow depth/elevation relationship on Hardangervidda very closely. In contrast to that, modelled snow depths from the seNorge snow model are roughly 0.5 m larger at all elevations on Hardangervidda. This is in agreement with the findings of Stranden (2010), Dyrrdal (2010), and Melvold and Skaugen (2013) who compared seNorge data with time series from meteorological stations and the same lidar-based snow measurements that were also used in this study. Dyrrdal (2010) attribute the overestimation of seNorge snow depths to too steep precipitation gradients in the model — caused by the lack of meteorological stations in some areas to calibrate the model locally. In such a case, ICESat-derived elevation-dependent snow depth estimates as we present here for Hardangervidda could serve as reference to better fit the seNorge snow depth/elevation relationship to real snow depths — without the need of field measurements. Such an application of ICESat data would be especially valuable in other, more remote areas, where even less measurements are available to constrain lateral and altitudinal model gradients.

5.1. Vegetation

Care should be taken when interpreting June dh as snow depths. Our results show that June dh at elevations below 900 m exceed autumn dh although no snow is expected at these elevations in summer. On eastern Hardangervidda, autumn to June snow depth differences at lower elevations even exceed the values from higher elevations. Unfortunately, the elevation-dependency of the vegetation types in our study area does not allow a more detailed statistical analysis of the influence of vegetation on dh. The reason for the June offsets could be the influence of dense, leaf-shedding (shrub) vegetation (Neuenschwander et al., 2008)

on the ICESat return waveform in June. For the elevations in concern, not only trees but also heathland plants with decimetre heights could possibly contribute to the signal. Stronger penetration of the ICESat signal into the vegetation in autumn, when most leaves will have fallen off, results in small differences of the ICESat waveform centroid elevation to reference (ground) elevations. Consequently, estimation of snow depths from ICESat during the vegetation period (June) works best in non-vegetated areas where the overlapping signals from snow and vegetation cover cannot be confused. On the other hand, we find no systematic bias in March or autumn dh from forested footprints compared to footprints with low/no vegetation cover in our study area. This indicates that ICESat may be capable to provide snow depth estimates also in forested areas — which is not the case for photogrammetric methods which only provide snow depths in open terrain (Marti et al., 2016). The AR50 land cover data used in this study does not distinguish coniferous and deciduous forests in mountainous areas (it does so in inhabited areas in the lowlands with sample distribution across entire southern Norway which prohibits a sound statistical analysis, given the potential for DEM bias influence). To determine the bias potential from vegetation cover, further work on ICESat-derived snow depths should focus on areas where the influence of elevation and vegetation cover on snow depths can be clearly separated.

5.2. Consequences of DEM biases

Local vertical offset correction of March/June dh with autumn dh considerably reduces elevation-dependent bias in reference elevations. However, such a correction might bias snow depth estimates negatively at highest elevations where snowfall is not unlikely during ICESat’s (late) autumn campaigns. In other areas, for example regions with summer monsoon snowfall (e.g. parts of High Mountain Asia) or on the southern hemisphere, the vertical bias correction might have to be based on reference seasons other than autumn, or might not work at all when no snow-free season is covered by ICESat campaigns. In latter cases snow depth differences between seasons may be derived, but not total snow depths. Further, our results show that even the correction of vertical offsets of small spatial units and/or elevation bands is not everywhere able to remove all bias. The influence of reference DEM quality becomes obvious from the samples on the Hardangervidda high-altitude plateau which is covered by all three reference DEMs. If all reference DEMs showed true terrain elevations, we would receive the same pattern from all DEMs — which is not the case. Treichler and Kääb (2016) showed that the uncertainties caused by within-footprint topography average out when a larger number of samples are summarised, in contrast to spatially consistent vertical bias in the reference DEMs. The example of Hardangervidda snow depths confirms the findings of that study: If enough samples are available, ICESat applications are limited by the quality of reference DEMs rather than ICESat’s elevation accuracy on rough surfaces. Thereby, spatial resolution of the reference DEM is of less importance than

the absence of (spatially varying) shifts or other biases in the data: in eastern Hardangervidda area, the 90 m SRTM DEM clearly performs better in reproducing actual snow depth compared to the 10 m Kartverket DEM that seems to be of poorer quality in this area. In highly alpine terrain with large elevation variation within a single footprint, the spatial resolution of the DEM would likely play a more important role than in more gentle mountain terrain like Hardangervidda.

The uncertainty in the reference DEMs is in complex interaction with the changing spatial distribution of the ICESat samples and the non-constant snow depths within the study area. This mainly has implications on the sample size needed in order to distinguish the snow depth signal from noise. If the spatial area is too large, the varying sample distribution can cause artificial patterns. An example: Our snow depth time series of the entire study region correspond well to modelled/measured snow depths, but biases from changing spatial distribution are clearly visible — they are a product of the changing composition of associated offsets and errors. The results are more robust if the temporal component is removed, and samples are grouped by yearly repeating campaign season — as this decreases the spatial area over which samples are averaged. Elevation errors in the reference DEMs are more likely to be constant over smaller spatial units. Our sample numbers per spatial unit or elevation band are similar to sample numbers for spatial units that other studies used for glacier surface elevation trends in mountainous terrain (e.g. Kääb et al., 2015; Treichler and Kääb, 2016; Gardner et al., 2013). However, there are many more samples on stable ground than on glaciers. As shown in this study, this should allow resolving average snow depths at higher spatial resolution compared to glacier studies in the same mountain area. Application of ICESat data for snow depths in other mountainous areas might be limited by the lack of sufficiently flat areas, or by lower snow depths that are distinctly smaller than elevation uncertainty. In this study, we excluded footprints on slopes $>10^\circ$ to remove slope-induced elevation uncertainty — which still leaves enough samples in the rounded Norwegian mountains, smoothed by the Scandinavian ice sheet, but potentially not in younger mountains such as the Alps or the Himalayas. In southern Norway, the inclusion of footprints on steeper slopes does not result in worse snow depth estimates but sample numbers have to be ca. doubled to compensate for the larger dh uncertainties.

5.3. Uncertainties and potential of ICESat-derived snow depths

From the comparison to measured and modelled snow depths we find uncertainties and errors of up to several decimetres in ICESat-derived, regional-scale snow depths in southern Norway. This is comparable to the decimetre-scale uncertainties of satellite-based stereo photogrammetry (Marti et al., 2016) but rather less accurate than air-

borne/terrestrial lidar and photogrammetry (Bühler et al., 2015; Deems et al., 2013). The latter methods provide snow depth maps at a much better spatial resolution of decimetres or metres — but they cannot cover more than a few square kilometres. On the scale of an entire mountain range (the Sierra Nevada), Dozier et al. (2016) find that different methods to estimate SWE differ by up to several hundred mm water equivalent. This corresponds to metre-scale differences in snow depths. On regional scale, ICESat-based snow depths could thus be a valuable data source for SWE estimation, in particular in combination with other (remote sensing) data sets and models (Dozier et al., 2016).

On the other hand, we find surprisingly good agreement of per-footprint lidar snow depth measurements with ICESat-derived snow heights even for single footprints. The few measurements available correlate well — despite the one month time difference in acquisition dates between ICESat and lidar snow measurements, and in defiance of the simple statistical representation of within-footprint snow depth distribution. The agreement might be further improved by considering the terrain influence on the laser return waveform as it is done for vegetation studies (e.g. Harding and Carabajal, 2005; Carabajal and Harding, 2006; Neuenschwander, 2008). The promising results in combination with high-resolution lidar data hints that the potential of ICESat in mountainous terrain is still not yet fully exploited. Compared to the already well established glacier trends from ICESat data, the snow analyses presented in this study have one distinct advantage: On stable ground, better reference data can still be acquired in the future, and new upcoming DEMs could improve ICESat analysis in retrospect (e.g. TanDEM-X or new local and national DEMs from lidar or high-resolution digital photogrammetry). The current increase of SfM-based 3D representations of the terrain, not least from unmanned aerial vehicles/systems (UAV, UAS), might also contribute to more and better reference elevation data over large areas useful for ICESat processing. This opens up the potential for retrospect snow depth estimates for the years 2003–2009 also in other remote areas where actual snow depths are not well known, and even for areas where we currently don't yet have sufficiently accurate reference DEMs.

While ICESat's short life and large footprints limit our snow-depth estimates to the last decade and a spatially coarse resolution, the specifications of its successor satellite promise much better data. ICESat-2 will have a denser cross- and along-track sampling as well as smaller ground footprints (Kramer, 2015) which should increase its accuracy over rough surfaces considerably. Given accurate reference DEMs, there is a potential that ICESat-2 will be able to provide near-realtime snow depth measurements (Jasinski and Neumann, 2013).

6. Conclusions

Given an accurate reference DEM, ICESat has the potential to provide snow depth data for the years 2003–2009 in mountainous terrain on regional scale and with decimetre-scale accuracy. ICESat data has several limitations: its low

temporal resolution of only two to three campaigns per year, the sparse spatial sampling with ca. 45 km between ground orbits in southern Norway, and the coarse spatial resolution due to the ca. 70 m large ground footprints. It also suffered from elevation inconsistencies throughout its lifetime — although these are of negligible concern over rough mountain terrain, where within-footprint topography is a priori the main source of surface elevation uncertainty. The terrain-driven uncertainties have random character with respect to ICESat sampling and average out when samples are grouped. However, the retrieval of snow depths requires that ICESat elevations are compared to reference surface elevations. On snow-covered ground, the difference between ICESat elevations and a reference surface correspond to snow depths. The reference DEMs typically have metre-scale elevation accuracy and contain spatially varying elevation biases. We find that the vertical biases in the reference DEMs are the main limitation of ICESat-derived snow depths. Other sources of elevation bias are steep slopes (in particular in combination with lateral DEM shifts) or vegetation cover. The larger the potential bias, the more samples have to be grouped to distinguish the snow signal from noise: for example into elevation bands, spatial units, or in time.

Time series of the elevation differences capture yearly differences in snow depth in southern Norway. On smaller spatial scale, ICESat’s campaigns around the month of March (or June) of all five years have to be combined to provide robust multi-annual average snow depths. The loss of the temporal dimension is compensated for by finer distinction of spatial patterns. Increasing snow depths with elevation or coastal proximity are correctly represented with decimetre-scale accuracy in our study. When grouped in elevation bands or spatial sub-units, the ICESat-derived snow depths agree well (RMSE 0.15–0.6 m) with campaign-averages from meteorological stations, airborne lidar snow depth maps on Hardangervidda, as well as modelled data from the spatially distributed seNorge model.

Our results bring out the importance of good control over errors and biases in reference DEM elevations. The need for grouping of several samples has the disadvantage that small-scale spatial variability of snow depths may be averaged out, too. Where errors are not random but systematic, they have the potential to severely bias snow depth estimates. This becomes evident from comparison of snow depths on Hardangervidda, derived from three different reference DEMs but using the same ICESat elevations. The different offsets for the three DEMs indicate that uncertainties in reference DEM elevations exceed ICESat uncertainties and, thus, limit ICESat snow applications. To remove spatially varying vertical DEM errors, we bias-correct March/June dh with median autumn (assumed snow free) dh of the same spatial subset/elevation band.

Comparison of ICESat-derived snow depths using a high-resolution lidar DEM and measured snow depths from lidar data of the same winter reveals good agreement (R^2 0.59, RMSE 1 m) even for single footprints. This indicates that

the potential of ICESat for snow depth retrieval is not yet fully exploited. On stable ground, it is not too late to acquire more accurate reference DEMs in the future. ICESat represents thus a source of regional-scale March snow depths and patterns at decimetre-scale accuracy for the years 2003–2009 in areas where no measurements are available. Even greater potential has its successor ICESat-2 which will provide more samples and smaller footprints once this data becomes available.

Acknowledgements

The study was funded by the European Research Council under the European Union’s Seventh Framework Programme (FP/2007-2013)/ERC grant agreement no. 320816, the ESA project *Glaciers_cci* (4000109873/14/I-NB) and the Department of Geosciences, University of Oslo. We thank the two anonymous reviewers whose comments and suggestions helped improve and clarify this manuscript. We are very grateful to NASA and USGS for free provision of the ICESat data and the SRTM DEM version we used, respectively, to NIBIO and the Norwegian mapping agency for their land cover and topographic data including the national DEM, the Norwegian Water Resources and Energy Directorate for providing the lidar data and modelled snow depths as well as glacier maps, and to the Norwegian Meteorological Institute for weather station data.

Author Contributions

D. Treichler designed the study, performed data analyses and wrote the paper. A. Kääb designed the study and edited the paper.

References

- Abshire, J.B., Sun, X., Riris, H., Sirota, J.M., McGarry, J.F., Palm, S., Yi, D., Liiva, P., 2005. Geoscience laser altimeter system (GLAS) on the ICESat mission: On-orbit measurement performance. *Geophysical Research Letters* 32. doi:10.1029/2005GL024028.
- Aune-Lundberg, L., Strand, G.H., 2011. Land resource classification in mountain areas. Examination of the classification system used in land resource mapping of Norwegian mountain areas. Report 01/11: 42 p. Skog og landskap, Norway.
- Bindschadler, R., Choi, H., Shuman, C., Markus, T., 2005. Detecting and measuring new snow accumulation on ice sheets by satellite remote sensing. *Remote Sensing of Environment* 98, 388–402. doi:10.1016/j.rse.2005.07.014.
- Borsa, A.A., Moholdt, G., Fricker, H.A., Brunt, K.M., 2014. A range correction for ICESat and its potential impact on ice-sheet mass balance studies. *The Cryosphere* 8, 345–357. doi:10.5194/tc-8-345-2014.

- Bühler, Y., Marty, M., Egli, L., Veitinger, J., Jonas, T., Thee, P., Ginzler, C., 2015. Snow depth mapping in high-alpine catchments using digital photogrammetry. *The Cryosphere* 9, 229–243. doi:10.5194/tc-9-229-2015.
- Carabajal, C.C., Harding, D.J., 2006. SRTM C-band and ICESat laser altimetry elevation comparisons as a function of tree cover and relief. *Photogrammetric Engineering and Remote Sensing* 72, 287–298. doi:10.14358/PERS.72.3.287.
- Clifford, D., 2010. Global estimates of snow water equivalent from passive microwave instruments: history, challenges and future developments. *International Journal of Remote Sensing* 31, 3707–3726. doi:10.1080/01431161.2010.483482.
- Deems, J.S., Painter, T.H., Finnegan, D.C., 2013. Lidar measurement of snow depth: a review. *Journal of Glaciology* 59, 467–479. doi:10.3189/2013JoG12J154.
- Dietz, A.J., Kuenzer, C., Gessner, U., Dech, S., 2012. Remote sensing of snow — a review of available methods. *International Journal of Remote Sensing* 33, 4094–4134. doi:10.1080/01431161.2011.640964.
- Dozier, J., Bair, E.H., Davis, R.E., 2016. Estimating the spatial distribution of snow water equivalent in the world's mountains. *Wiley Interdisciplinary Reviews: Water* 3, 461–474. doi:10.1002/wat2.1140.
- Dyrddal, A.V., 2010. An evaluation of Norwegian snow maps: simulation results versus observations. *Hydrology Research* 41, 27–37. doi:10.2166/nh.2010.019.
- Engeset, R., Tveito, O.E., Udnæs, H.C., Alfnes, E., Mengistu, Z., Isaksen, K., Førland, E.J., 2004. Snow map validation for norway, in: *Proceedings XXIII Nordic Hydrological Conference 2004*, Tallinn, Estonia. pp. 8–12.
- Farr, T.G., Kobrick, M., 2000. Shuttle radar topography mission produces a wealth of data. *Eos, Transactions American Geophysical Union* 81, 583–585. doi:10.1029/E00811048p00583.
- Fassnacht, S.R., Brogan, D., Sextstone, G.A., Jasinski, M., Lopz-Moreno, J.I., Skordahl, M., 2012. Spatio-temporal variability of snowpack properties: Comparing operational, field, and ICESat remote sensing data over Northern Colorado, United States, in: *2012 IEEE International Geoscience and Remote Sensing Symposium*, pp. 1576–1577. doi:10.1109/IGARSS.2012.6351094.
- Gardner, A.S., Moholdt, G., Cogley, J.G., Wouters, B., Arendt, A.A., Wahr, J., Berthier, E., Hock, R., Pfeffer, W.T., Kaser, G., Ligtenberg, S.R.M., Bolch, T., Sharp, M.J., Hagen, J.O., van den Broeke, M.R., Paul, F., 2013. A reconciled estimate of glacier contributions to sea level rise: 2003 to 2009. *Science* 340, 852–857. doi:10.1126/science.1234532.
- Hall, D.K., Riggs, G.A., Salomonson, V.V., DiGirolamo, N.E., Bayr, K.J., 2002. MODIS snow-cover products. *Remote Sensing of Environment* 83, 181–194. doi:10.1016/S0034-4257(02)00095-0.
- Harding, D.J., Carabajal, C.C., 2005. ICESat waveform measurements of within-footprint topographic relief and vegetation vertical structure. *Geophysical Research Letters* 32. doi:10.1029/2005GL023471.
- Hofton, M.A., Luthcke, S.B., Blair, J.B., 2013. Estimation of ICESat intercampaign elevation biases from comparison of lidar data in East Antarctica. *Geophysical Research Letters* 40, 5698–5703. doi:10.1002/2013GL057652.
- Höhle, J., Höhle, M., 2009. Accuracy assessment of digital elevation models by means of robust statistical methods. *ISPRS Journal of Photogrammetry and Remote Sensing* 64, 398–406. doi:10.1016/j.isprsjprs.2009.02.003.
- Immerzeel, W.W., Wanders, N., Lutz, A.F., Shea, J.M., Bierkens, M.F.P., 2015. Reconciling high-altitude precipitation in the upper Indus basin with glacier mass balances and runoff. *Hydrology and Earth System Sciences* 19, 4673–4687. doi:10.5194/hess-19-4673-2015.
- Jasinski, M.F., Neumann, T., 2013. The ICESat2 Mission: Objectives, concept, and opportunities for snow. *NASA Snow Remote Sensing Workshop CIRES*, August 14; Boulder, Colorado, USA. URL: <http://nasasnowremotesensing.gi.alaska.edu/sites/default/files/Jasinski.pdf>.
- Jasinski, M.F., Stoll, J., 2012. Feasibility of estimating snow depth in complex terrain using satellite lidar altimetry. *Eastern Snow Conference Meeting*; 5–7 June 2012; Claryville, NY, USA. URL: <http://ntrs.nasa.gov/search.jsp?R=20120013477>.
- Kääb, A., Berthier, E., Nuth, C., Gardelle, J., Arnaud, Y., 2012. Contrasting patterns of early twenty-first-century glacier mass change in the Himalayas. *Nature* 488, 495–498. doi:10.1038/nature11324.
- Kääb, A., Treichler, D., Nuth, C., Berthier, E., 2015. Brief Communication: Contending estimates of 2003–2008 glacier mass balance over the Pamir-Karakoram-Himalaya. *The Cryosphere* 9, 557–564. doi:10.5194/tc-9-557-2015.
- Kartverket, 2016. Terrenghogmodeller — land. Data description. The Norwegian Mapping Authority, Norway. Online, accessed 6 June 2016. URL: <http://www.kartverket.no/data/kartdata/terrenghogmodeller/Terrenghogmodell-10-meters-grid/>.
- Ke, L., Ding, X., Song, C., 2015. Heterogeneous changes of glaciers over the western Kunlun Mountains based on ICESat and Landsat-8 derived glacier inventory. *Remote Sensing of Environment* 168, 13–23. doi:10.1016/j.rse.2015.06.019.

- Kramer, H.J., 2015. ICESat-2 (ice, cloud and land elevation satellite-2). Online, accessed June 2015. URL: <https://directory.eoportal.org/web/eoportal/satellite-missions/i/icesat-2>.
- Kropáček, J., Neckel, N., Bauder, A., 2014. Estimation of Mass Balance of the Grosser Aletschgletscher, Swiss Alps, from ICESat Laser Altimetry Data and Digital Elevation Models. *Remote Sensing* 6, 5614. doi:10.3390/rs6065614.
- Lettenmaier, D.P., Alsdorf, D., Dozier, J., Huffman, G.J., Pan, M., Wood, E.F., 2015. Inroads of remote sensing into hydrologic science during the WRR era. *Water Resources Research* 51, 7309–7342. doi:10.1002/2015WR017616.
- Marti, R., Gascoin, S., Berthier, E., de Pinel, M., Houet, T., Laffly, D., 2016. Mapping snow depth in open alpine terrain from stereo satellite imagery. *The Cryosphere* 10, 1361–1380. doi:10.5194/tc-10-1361-2016.
- Melvold, K., Skaugen, T., 2013. Multiscale spatial variability of lidar-derived and modeled snow depth on Hardangervidda, Norway. *Annals of Glaciology* 54, 273–281. doi:10.3189/2013AoG62A161.
- Mohr, M., 2008. New routines for gridding of temperature and precipitation observations for seNorge.no. met.no note 08/2008, The Norwegian Meteorological Institute, Oslo, Norway.
- Neckel, N., Kropáček, J., Bolch, T., Hochschild, V., 2014. Glacier mass changes on the Tibetan Plateau 2003–2009 derived from ICESat laser altimetry measurements. *Environmental Research Letters* 9. doi:10.1088/1748-9326/9/1/014009.
- Neuenschwander, A.L., 2008. Evaluation of waveform deconvolution and decomposition retrieval algorithms for ICESat/GLAS data. *Canadian Journal of Remote Sensing* 34, S240–S246. doi:10.5589/m08-044.
- Neuenschwander, A.L., Urban, T.J., Gutierrez, R., Schutz, B.E., 2008. Characterization of ICESat/GLAS waveforms over terrestrial ecosystems: Implications for vegetation mapping. *Journal of Geophysical Research* 113. doi:10.1029/2007JG000557.
- Nolan, M., Larsen, C., Sturm, M., 2015. Mapping snow depth from manned aircraft on landscape scales at centimeter resolution using structure-from-motion photogrammetry. *The Cryosphere* 9, 1445–1463. doi:10.5194/tc-9-1445-2015.
- NSIDC, 2014. GLAS/ICESat L1 and L2 Global Altimetry Data, Version 34. Data description. National Snow and Ice Data Center. University of Colorado, Boulder, Colorado, USA. URL: http://nsidc.org/data/docs/daac/glas_icesat_11_12_global_altimetry.gd.html.
- NSIDC, 2016. Attributes for ICESat Laser Operations Periods: Latest Release. Data description. National Snow and Ice Data Center. University of Colorado, Boulder, Colorado, USA. Online, accessed 30 October 2016. URL: ftp://sidacs.colorado.edu/pub/DATASETS/nsidc0587_ICESAT_GLAS_LTA/Section3_Product_Documentation/glas_laser_ops_attrib.pdf.
- Nuth, C., Kääb, A., 2011. Co-registration and bias corrections of satellite elevation data sets for quantifying glacier thickness change. *The Cryosphere* 5, 271–290. doi:10.5194/tc-5-271-2011.
- NVE, 2016. Climate indicator products. Online glacier database. Norwegian Water Resources and Energy Directorate. Online, accessed 31 May 2016. URL: <http://glacier.nve.no/viewer/CI/>.
- Rango, A., 1994. Application of remote sensing methods to hydrology and water resources. *Hydrological Sciences Journal* 39, 309–320. doi:10.1080/02626669409492752.
- Rasmussen, L.A., 2013. Meteorological controls on glacier mass balance in High Asia. *Annals of Glaciology* 54, 352–359. doi:10.3189/2013AoG63A353.
- Rodriguez, E., Morris, C.S., Belz, J.E., 2006. A global assessment of the SRTM performance. *Photogrammetric Engineering & Remote Sensing* 72, 249–260. doi:10.14358/PERS.72.3.249.
- Saloranta, T., 2014. New version (v.1.1.1) of the seNorge snow model and snow maps for Norway. NVE report 6, 2014. Norwegian Water Resources and Energy Directorate, Oslo, Norway.
- Saloranta, T.M., 2012. Simulating snow maps for Norway: description and statistical evaluation of the seNorge snow model. *The Cryosphere* 6, 1323–1337. doi:10.5194/tc-6-1323-2012.
- Schutz, B., DiMarzio, J., Luthcke, S., Hancock, D., Urban, T., 2011. Notice Concerning Detection of ICESat/GLAS Inter-Campaign Elevation Biases. Technical Report. The National Snow and Ice Data Center. URL: https://nsidc.org/sites/nsidc.org/files/files/inter-campaign_bias_notice_v1.pdf.
- Schutz, B.E., Zwally, H.J., Shuman, C.A., Hancock, D., DiMarzio, J.P., 2005. Overview of the ICESat Mission. *Geophysical Research Letters* 32. doi:10.1029/2005GL024009.
- Siegfried, M.R., Hawley, R.L., Burkhart, J.F., 2011. High-resolution ground-based GPS measurements show inter-campaign bias in ICESat elevation data near Summit, Greenland. *IEEE Transactions on Geoscience and Remote Sensing* 49, 3393–3400. doi:10.1109/TGRS.2011.2127483.

- Stoll, J., Jasinski, M.F., 2012. Estimation of Snow Depth in the Uinta Mountains using ICESat/GLAS Observations. AGU Fall Meeting Abstracts 2012 .
- Stranden, H.B., 2010. Evaluering av seNorge: data versjon 1.1. NVE dokument 4, 2010. Norwegian Water Resources and Energy Directorate, Oslo, Norway.
- Treichler, D., Kääb, A., 2016. ICESat laser altimetry over small mountain glaciers. *The Cryosphere* 10, 2129–2146. doi:10.5194/tc-10-2129-2016.
- Vander Jagt, B., Lucieer, A., Wallace, L., Turner, D., Durand, M., 2015. Snow depth retrieval with UAS using photogrammetric techniques. *Geosciences* 5, 264. doi:10.3390/geosciences5030264.
- Vikhamar, D., Solberg, R., 2003. Subpixel mapping of snow cover in forests by optical remote sensing. *Remote Sensing of Environment* 84, 69–82. doi:10.1016/j.rse.2003.06.004.
- Zwally, H., Schutz, R., Hancock, D., Dimarzio, J., 2012. GLAS/ICESat L2 Global Land Surface Altimetry Data, Versions 33/34, GLAH14. Dataset. NASA DAAC at the National Snow and Ice Data Center, Boulder, Colorado USA. doi:10.5067/ICESAT/GLAS/DATA207.