

UNIVERSITY
OF OSLO

Master thesis

Snow depth Retrieval and Downscaling using Satellite Laser Altimetry, Machine Learning, and Climate Reanalysis

A Case Study in Mainland Norway

Zhihao LIU

Department of Geosciences
The Faculty of Mathematics and Natural Sciences

2023



Snow Depth Retrieval and Downscaling using Satellite Laser Altimetry, Machine Learning, and Climate Reanalysis

A Case Study in Mainland Norway

Zhihao LIU



2023-05-13

Acknowledgements

I am deeply grateful to the individuals who have supported and contributed to the completion of my thesis.

First and foremost, I extend my heartfelt thanks to Désirée for her unwavering support, insightful ideas, guidance, and feedback. Her constant encouragement and constructive suggestions helped me to refine my research and achieve my goals. I would also like to thank Andy for his excellent teaching and contributions on the NuthKaab co-registration, which served as a starting point for my thesis.

I am also thankful to Macro, Eric, and other colleagues in the department for engaging in many detailed discussions, providing valuable feedback and insights, and helping me to stay motivated. Special thanks to Bernd, Karianne, and Henning for organizing such a amazing geomorphology field trip. I would like to acknowledge NVE for providing the validation dataset that was crucial to my work.

I would like to extend my appreciation to Professor Xu for his engaging lectures on geophysical data science and warm-hearted hosting. Additionally, I am grateful to my classmates, Milot, Femke, Gabriele, Stewart, Johanne, Halldor, Helene, Stewart, Gard, Stian, Jørgen, Kira and Satu, as well as Cao Bo, Wenzhe, Tintin, Jiamei, Yaoyao, Ruitang, Nastya, Ryan, Aga, Romvy for making my time in Oslo memorable. I also express my gratitude to my flatmates, Hanna, Jing, Tong, Veera, Cecilia, and Natalia, as well as OSI Friluft and OSEG for their good vibes.

Finally, I began my studies amidst the COVID-19 pandemic, and I would like to express my gratitude to my family for their understanding of my absence during this difficult time, especially my grandfather.

Oslo, May 13, 2022

Zhihao LIU

Summary

Seasonal snow plays a crucial role as water reservoirs and energy balance component, but accurately estimating the depth of the snowpack remains a challenge, particularly in remote areas. ICESat-2 laser altimetry has the potential to provide precise snow depth measurements by comparing satellite-based snow surface elevation profiles with high-quality Digital Elevation Model (DEM) of the snow-free ground. However, the satellite's acquisition pattern is sparse both in time and space, raising the need for additional steps to produce a spatially complete snow-depth map.

This study generated downscaled snow depth maps by employing a machine-learning regressor to combine snow depths derived from ICESat-2 ATL08 product (2018-2022) with ERA-5 Land data. The methodology involves careful co-registering the data, applying bias correction on DEMs using ICESat-2 surface elevations from snow-free conditions as a reference. Subsequently, snow depth maps are generated by statistically downscaling ERA-5 time-series data on snow depth with the derived snow depth, using terrain, vegetation, and wind parameters. Our results are able to reproduce snow depth patterns at the hill-slope scale, achieving an R-Squared value of 0.68 and a Spearman correlation coefficient of 0.81 when compared to lidar-based snow depth acquired in Hardangervidda, Norway. This approach is applicable globally in any location where accurate snow-free DEMs are available.

Additionally, this study contributes a Gradient Descent Co-registration algorithm, which offers possibilities for handling large-scale datasets at a fine resolution. And, a bias correction workflow is utilized to address uncertainties of DEM and ICESat-2, which refines the conventional elevation differencing workflow of producing snow depth.

The thesis is structured as follows: *Chapter 1* provides an introduction to the role of snow cover, observations techniques and thesis objectives. *Chapter 2* reviews the key concepts of the methodologies, including snow depth variability, satellite laser altimetry, DEM uncertainties, co-registration, machine learning, climate reanalysis, downscaling techniques. *Chapter 3* presents a workflow for snow depth retrieval, including gradient descent co-registration, evaluating DEM against ICESat-2 ATL08, applying bias correction, and downscaling of ERA-5 using ICESat-2 derived snow depth. *Chapter 4* presents the methodology's application in mainland Norway, including snow depth validations. *Chapter 5* discusses the uncertainties and limitations. Finally, *Chapter 6* concludes remarks.

Keywords: Snow depth, ICESat-2, Co-registration, Machine learning, Bias correction, Statistical downscaling

Contents

Acknowledgements	i
Summary	iii
Contents	vii
List of Figures	x
List of Tables	xi
1. Introduction	1
1.1. The Role of Snow Cover	1
1.1.1. Energy balance component	1
1.1.2. Climate indicator	1
1.1.3. Water reservoir	2
1.1.4. Ecological habitats	2
1.1.5. Thermal insulator	3
1.1.6. Geohazards	5
1.2. The Challenges of the Snow Observation	5
1.2.1. Snow cover	5
1.2.2. Snow mass	6
1.3. Objectives	8
1.4. Thesis structure	9
2. Scientific Overview	11
2.1. The Variability of the Snow Cover	11
2.1.1. Spatial variability	11
2.1.2. Temporal variability	12
2.2. DEM	14
2.2.1. DEM products	14
2.2.2. DEM uncertainties	15
2.2.3. DEM Co-registration	18
2.3. ICESat-2 Altimetry	19
2.3.1. Addressing DEM errors by ICESat-2	20

Contents

2.3.2. Snow depth from ICESat-2	21
2.4. Machine Learning and XGBoost	22
2.4.1. Hyperparameters of XGBoost	24
2.5. Climate Reanalysis and Downscaling	24
2.5.1. Statistical downscaling and subgrid variability of snow depth	26
3. Study Area and Methodologies	29
3.1. Research Design	29
3.2. The Physical Geographical Setting	30
3.3. The Data Setting	31
3.3.1. ICESat-2 ATL08	31
3.3.2. DEMs	35
3.3.3. Terrain and vegetation features	35
3.3.4. ERA5 Land: snow and wind features	38
3.4. Gradient Descent Co-registration	39
3.4.1. Noise suppression	40
3.5. Regression Model	42
3.5.1. Bias correction	42
3.5.2. Statistical downscaling	43
3.5.3. Hyperparameters and cross-validation	46
3.5.4. Feature engineering	46
3.6. Evaluation of Snow Depth	49
3.6.1. ICESat-2 derived snow depth vs lidar derived snow depth	49
3.6.2. ICESat-2 derived snow depth vs ERA5 Land snow depth	50
3.6.3. Downscaled snow depth vs lidar derived snow depth	50
4. Results	51
4.1. Co-registration	51
4.2. Binning Analysis: DEM Comparison	52
4.3. Feature Importance and Regression Performance of Bias Correction	57
4.4. ICESat-2 Derived Snow Depth and Validation	60
4.5. ICESat-2 Derived Snow Depth vs ERA5 Land	62
4.6. Modeled Snow Depth and Validation	64
4.7. Feature Importance of Downscaling Regression	64
5. Discussion	69
5.1. GDC vs NuthKaab Co-registration	69
5.2. The Uncertainties: ICESat-2, DEM and Bias Correction	70
5.2.1. ICESat-2 ATL08: skewness and curvature	70

Contents

5.2.2. Uncertainties in DEMs: vegetation, resolution, snowpack	72
5.2.3. Assumptions behind bias correction: snow-on vs snow-free	74
5.3. ICESat-2 Derived Snow Depth: Cutting-out	74
5.4. Statistical Downscaling: Noise and Randomization	75
6. Concluding Remarks and Future Outlooks	79
References	81
A. Appendix	95
A.1. Additional Figures	95
A.2. Technical Details of ICESat-2 ATL08	100
A.2.1. Processing and QC Steps of ICESat-2 ATL08 Land products	100
A.3. Recommendation For Winter Canopy Height Correction	102
A.4. The Consistency of ICESat-2 ATL08 Snow-free Segments	103
A.5. Code Repository	105
A.5.1. Offset adjustment: spatial aggregation, linear adjustment and quantile mapping	105
A.5.2. Gradient descent co-registration	106
A.5.3. XSnow	106

List of Figures

1.1. Historical Snow Trend From 1981 to 2018	3
1.2. Snowmelt-dominated Catchments: Present and Future	4
1.3. Remote Sensing of SWE/Snow Depth	6
1.4. Timeline of Satellite Altimetry since the 1990s	8
2.1. Snow Redistribution and Wind	12
2.2. An Overview of Models for Snow Redistribution	13
2.3. Uncertainty Framework for DEM Assessment	16
2.4. Non-stationary and Heteroscedasticity	17
2.5. Mean Absolute Errors of Global DEM Products in the Tibetan Plateau	17
2.6. ICESat-2 and Photon Counting System	20
2.7. A Bias Correction Scheme for DEMs Using ICESat-2 Based on Slopes and Vegetation Coverage	21
2.8. How XGBoost Works	23
3.1. Study Design Schematic	29
3.2. The Climate of Norway	32
3.3. ICESat-2 Segments and Snow Surveys	33
3.4. The Topography Index: Scale-dependent Metric	37
3.5. Quantifying the Relationship between Wind and Aspect in Snow Redistribution	38
3.6. Co-registration Techniques: NuthKaab vs Gradient Descent	40
3.7. Brute-Force Test on Gradient Descent Co-registration	41
3.8. Preventing Overfitting with Cross-validation in XGBoost	47
3.9. Spatial Aggregation of Snow-free and Snow-on Segments.	48
3.10. Snow Survey via Drone Lidar in Hardangervidda, Southern Norway	49
4.1. The Misalignment of DEMs	53
4.2. Statistic Binning Analysis: Co-registration and Bias Correction (ICESat-2 - DTM10)	54
4.3. The Mean Error and NMAD of Quantiles: Co-registration and Bias correction	55
4.4. The Overall Distribution of DEM errors: Co-registration and Bias correction	56
4.5. Bias Correction Map for Snow-on Scenarios	58
4.6. Feature Importance of Bias Correction Regression	59

List of Figures

4.7. Aggregated Snow Depth from ICESat-2 and Four DEMs, 2018-2022	61
4.8. The Validation of ICESat-2 Derived Snow Depth	62
4.9. ICESat-2 Derived Snow Depth vs ERA5 Land Snow Depth.	63
4.10. The Validation of Snow Depth Downscaling: Map and Profile	65
4.11. The Validation of Snow Depth Downscaling: Histogram and Scattering	66
4.12. The Validation of Snow Depth Downscaling: Plan Curvature, Vegetation Coverage, El- elevation and East	67
4.13. Feature Importance of Snow Depth Downscaling	68
5.1. Skewness Variations in Convex and Concave Terrains	71
5.2. Comparison of FAB, GLO30, and DTM1 in terms of Plan Curvature and Canopy Height .	73
5.3. Features Related to Snow-on and Snow-free Surface	76
5.4. The Negative Snow Depth and Cutting-out Value	76
A.1. The Monthly Average of Wind Speed and Direction in Winter (2019-2022)	96
A.2. Statistic binning Analysis of DEM Errors: Co-registration and Bias correction (ICESat-2 - DTM1)	97
A.3. The Validation of Snow Depth Downscaling (2009)	98
A.4. Comparison of DTM1 and DTM10 in terms of Plan Curvature and Canopy Height . . .	99
A.5. Comparison of Clear Sky and Cloudy Sky	99
A.6. ICESat-2 Consistency: Year, Day/night, Strong/Weak, Beams and Pairs	104

List of Tables

2.1. Snow Distribution at Different Scales.	11
2.2. Global DEM Products.	15
3.1. Features in Bias Correction Regressor	43
3.2. Features in Downscaling Regressor	44
4.1. The Improvement of DEM by Co-registration	51
4.2. The Improvement of DEM by Co-registration (<i>subset_te_flag</i> equals 5)	52
4.3. Cross Validation of Bias Correction Regression	60
5.1. Impact of Cutting-out Negative Values on Snow Depth Statistics	74

1. Introduction

1.1. The Role of Snow Cover

Snow cover is a vital component of Earth's seasonal cycle, with approximately 47% of the Northern Hemisphere (NH) land surface being covered by snow during winter (Estilow et al., 2015). Moreover, 1.9 billion people rely on glaciers and seasonal snow packs for their freshwater supply (Barnett et al., 2005; Immerzeel et al., 2020). Changes in snow cover interact with various systems, such as climate, hydrology, geohazards, and ecology (Henderson et al., 2018; Immerzeel et al., 2020), through influencing the surface energy balance, water balance, thermal regimes, vegetation, and trace gas fluxes (Callaghan et al., 2011). Therefore, understanding the role of snow cover and observing its variability is crucial in addressing regional and global challenges.

1.1.1. Energy balance component

Snow cover presence or absence governs Earth's surface heating and cooling patterns. Fresh snow is highly reflective and white, with an albedo of 0.8-0.9, reflecting an enormous amount of sunlight back into space instead of warming the surface. Conversely, without snow cover or in the presence of black carbon pollutants (Réveillet et al., 2022), the ground absorbs more solar energy, leading to more rapid melting and further warming (Riihelä et al., 2021). This effect is called snow and ice albedo feedback (Perovich et al., 2007), which is driving warming trends in cold regions at a rate two-to-four times faster than the global average (Rantanen et al., 2022). This phenomenon is commonly referred to as Arctic amplification.

1.1.2. Climate indicator

Seasonal snow cover is a valuable climate indicator due to its sensitivity to changes in air temperature and humidity, as demonstrated by observations and models (Brown & Mote, 2009; Gulev et al., 2021). For instance, a remarkable linear relationship exhibits that a 1-degree Celsius increase in warming leads to a decrease in NH spring snow cover extent of about 8% relative to the 1995 - 2014 level in the World Climate Research Programme Coupled Model Inter-comparison Project Phase 6 (CMIP6) (L. Mudryk et al., 2020).

1. Introduction

Spring snow cover has significantly decreased in mountain regions, especially at lower elevations, with an average duration decline of 5 days per decade (Hock et al., 2019). In the Alps, there has been a reduction of 5.6% per decade in snow cover duration over the past 50 years (Carrer et al., 2023; Matiu et al., 2021). The Arctic has experienced an average decrease of $-3.5 \pm 1.9\%$ per decade in May and $-13.4 \pm 5.4\%$ per decade in June in snow cover extent between 1967 and 2018 (Meredith et al., 2019), along with a decline of two to four days per decade in snow cover duration since the 1970s (AMAP, 2017).

Besides, various satellite-based passive microwave products and other observations also indicate a general negative trend in NH snow mass between 1981 and 2018, as showed in Figure 1.1. The shift from solid to liquid precipitation is observed during shoulder seasons (L. Mudryk et al., 2020) and in mountain areas (Rhoades et al., 2022; Siirila-Woodburn et al., 2021). The worldwide reduction in glacier mass balance also implies a loss of snow on glaciers and ice sheets, again a partial confirmation of the declining of snow mass (Kraaijenbrink et al., 2021; Kumar et al., 2019).

In North America, there has been a consistent decrease in maximum snow depth since the 1960s (Kunkel et al., 2016). Similarly, the Alps have seen an 8.4% decrease per decade over the last fifty years in maximum snow depth (Matiu et al., 2021). However, regional and temporal variations in snow depth (or snow water equivalent) are significant due to limited data coverage and lack of high-resolution observations for mountain regions (Kunkel et al., 2016; Luojus et al., 2021; Xu et al., 2017). Increasing humidity can result in heavy snowfall, while higher temperatures decrease the probability of snowfall (L. Mudryk et al., 2020; Quante et al., 2021).

1.1.3. Water reservoir

The snowpack plays a crucial role as a surface water storage reservoir, influencing the timing of runoff that sustains the water demands of downstream ecosystems (Immerzeel et al., 2020; Kraaijenbrink et al., 2021). Alterations in snowpacks can significantly impact water availability, particularly in a warmer world (Figure 1.2). This is due to reduced winter precipitation falling as snow and an earlier onset of melting in spring (Musselman et al., 2021). Even without any changes in precipitation intensity, these effects lead to a shift in peak river runoff towards winter-early spring, away from summer and autumn when demand is highest (Barnett et al., 2005). Additionally, the snowpack serves as an essential indicator for flood prediction and provides early warning of hydrologic drought (Livneh & Badger, 2020).

1.1.4. Ecological habitats

Changes in snow conditions can significantly impact vegetation distribution, biodiversity, and productivity in various environments (Callaghan & Johansson, 2021), such as alpine (Wipf et al., 2009),

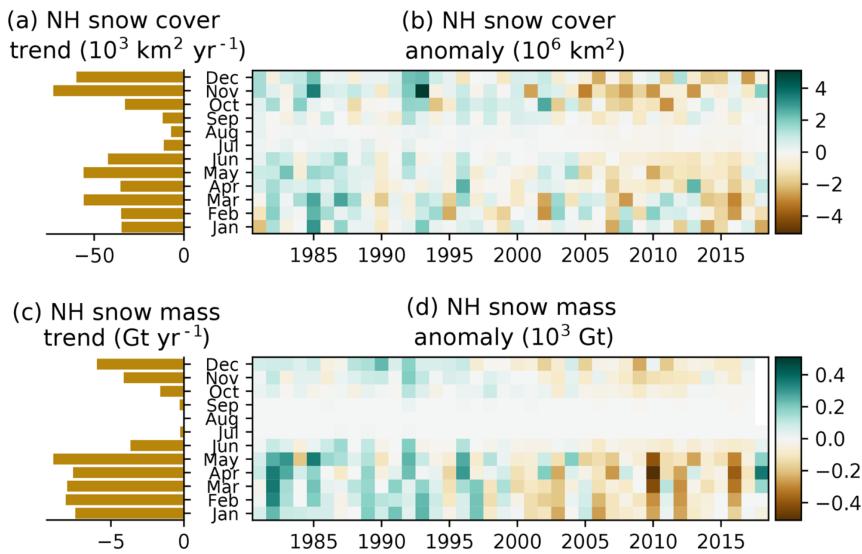


Figure 1.1.: Historical Snow Trend From 1981 to 2018. Trends in snow extent over 1981–2018 are negative in all months and exceed $-50 \times 10^3 \text{ km}^2 \text{ yr}^{-1}$ during November, December, March, and May (a, b). Snow mass trends are approximately -5 Gt yr^{-1} or more for all months from December to May (c, d). Source: L. Mudryk et al. (2020).

Arctic (Sturm et al., 2001), and desert regions (PENG et al., 2010). These changes can affect ecosystem through frost events, soil moisture, and winter soil temperatures, surface energy budgets (Callaghan et al., 2011; Harpold & Molotch, 2015). For example, the Arctic has experienced “greening” in the past decades due to decreases in snow cover duration and an extended growing season (Callaghan et al., 2011), but some areas have reversed to a “browning” trend in recent years with unclear reasons, demonstrating the complexity of snow-vegetation interactions (Myers-Smith et al., 2020).

1.1.5. Thermal insulator

Snow layers’ porous structure makes them excellent thermal insulators, with profound implications for the permafrost thermal regime. For instance, the height of palsas influences their decay rate, as a higher palsas accumulates more snow on its leeward side due to snow drift, than a lower palsas. The thicker the snow depth, the less cold penetration there would be in winter and the more meltwater in spring and summer, leading to increased local permafrost thawing and block erosion (Olvmo et al., 2020; Seppälä, 1994). After accounting for statistically derived snow depths, regional permafrost models suggest nearly twice as large a total permafrost area in Norway compared to those that do not consider snow (Gisnås et al., 2016). The same insulation mechanism applies to float ice (Sturm & Massom, 2016) and winter wheat as well (Zhu et al., 2022).

1. Introduction

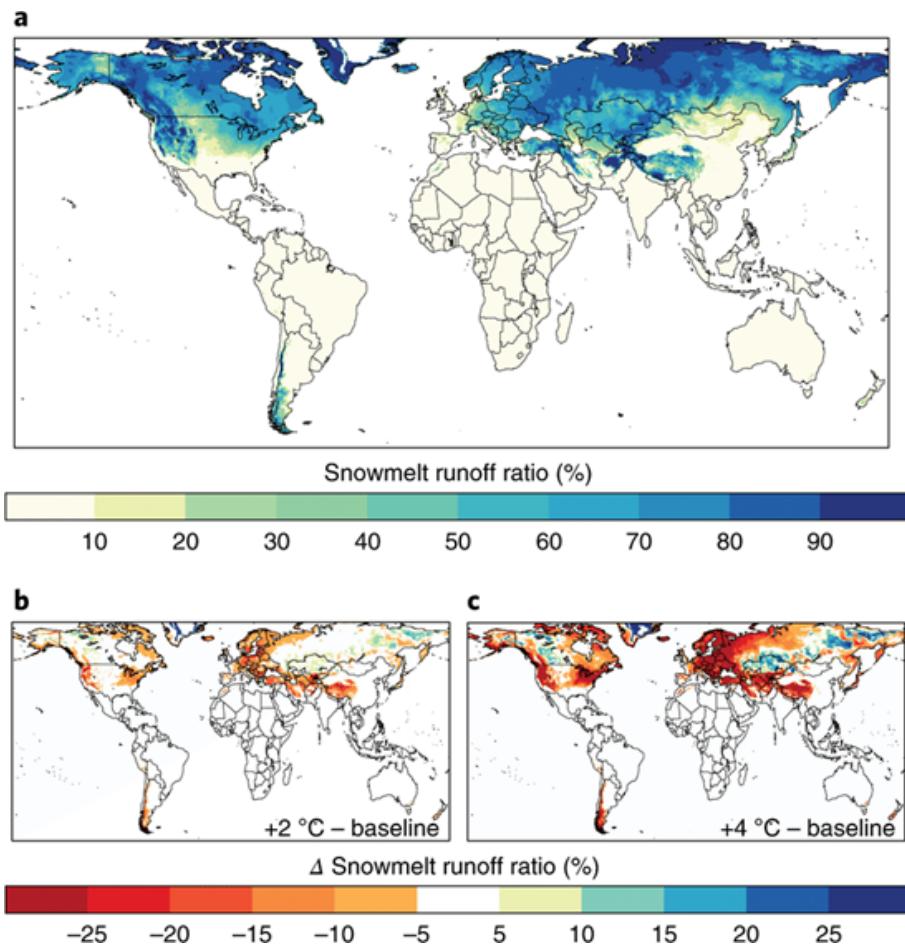


Figure 1.2.: Snowmelt-dominated Catchments: Present and Future. The average ratio of annual runoff from snowmelt based on historical data (1985-2015). It shows that snowmelt contributes to at least half of the runoff across 26% of the global land area, particularly in western US, Tibetan Plateau (TP), Central Asia, the southern Andes, and Scandinavia (a). In addition, in a warmer climate, the proportion of runoff derived from snowmelt significantly declines with an average warming of 2 °C and 4 °C (b, c, respectively). Source: Qin et al. (2020).

1.1.6. Geohazards

Avalanches are natural disasters that occur when a massive amount of snow slides down a slope, often triggered by external factors such as weather changes, earthquakes, or human activities (Schweizer et al., 2003). The impact of an avalanche can be devastating, causing significant damage to infrastructure and loss of life. Yet, predicting a single avalanche event is impossible due to its inherently variable nature (Schweizer et al., 2021). However, observing and modeling snow depth and the persistence of snow accumulation patterns can help identify avalanche risk.

Rain-on-snow events play a crucial role in the runoff regime of snow-dominated catchments, as flood water is formed not only by rain input but also by melted snow. The information of the snowpack was particularly useful in assessing the flooding risk (Würzer & Jonas, 2018). Snowstorms are another geohazard that can pose a threat to society. Heavy snow accumulation can cause structural damage to building roofs and electrical grids. Additionally, heavy snowfalls can block roads or railways, leading to severe transportation disruptions such as traffic congestion and delays (Mo et al., 2016).

1.2. The Challenges of the Snow Observation

Snowpack is characterized via three variables: snow cover extent (SCE), continuous snow cover duration (SCD), and snow accumulation, which can be expressed as either snow depth (SD) or snow water equivalent (SWE).

1.2.1. Snow cover

There are two types of products for mapping snow cover: a binary classification in which each pixel in an image is designated as ‘snow’ or ‘non-snow’ cover area (SCA) or a percentage of snow-cover (SCF, snow cover fraction). The snow cover information is usually derived from optical satellite systems because of their high resolution but with obvious limitations related to persistent clouds, forest cover (Nolin, 2010), lack of solar illumination at higher latitude regions (Fox-Kemper et al., 2021) and rough terrain. The MODIS (Moderate Resolution Imaging Spectroradiometer) provides one of the state-of-the-art snow cover products. It has a high spatial resolution (500 m) and daily coverage that can extend the time series from 2000 to present (Hall et al., 2006). High-resolution imagery from Landsat has also proven useful for validating operational snow-cover products and for applications that require snow cover information at spatial scales of less than 50 m (Macander et al., 2015).

1. Introduction

1.2.2. Snow mass

Snow mass is a challenging parameter to monitor, both globally and regionally, compared to snow cover extent (L. Mudryk et al., 2020). A major persistent gap in contemporary is observing reliable SWE over mountain regions (Bormann et al., 2018). There are several reasons for this:

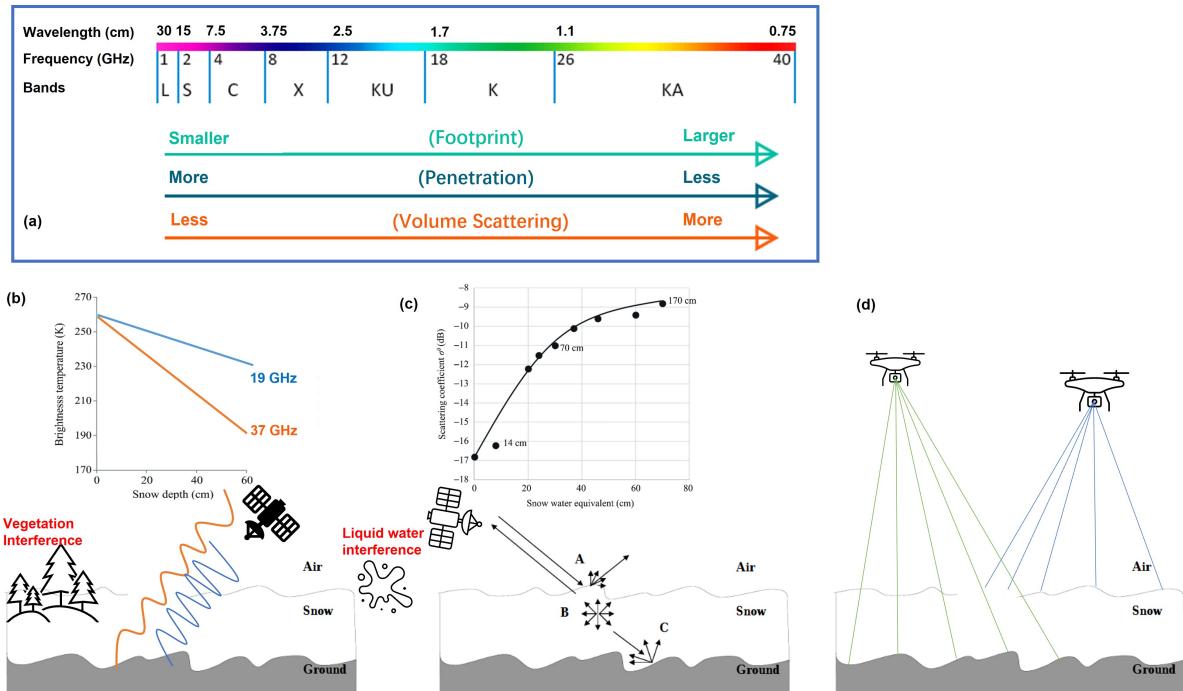


Figure 1.3.: Remote Sensing of SWE/Snow Depth. The figure present radar frequency bands (a), SWE from microwave sensors (b), SWE from active microwave sensors (c) and snow depth from elevation differencing (d). The difference between brightness temperature at 18 GHz and 36 GHz is directly related to snow depth, given constant snow grain size and density (b). The radar backscatter is contributed by scattering at the air-snow interface, volumne scattering from snowcpk and the ground surface. The observed backscattering coefficient can be trasfered into SWE, given emperical coefficients such as polarization, incidence angle, snow wetness (c). Adapted from Stiles & Ulaby (1980)

- (1) Weather stations are generally located in gentle terrain at lower or mid elevation, which makes it difficult to obtain representative measurements for regions with varying topography. Moreover, the precipitation gauges often have errors due to wind effects which can lead to significant and systematic underestimates of up to 50% (Rasmussen et al., 2012), particularly for historical archives.
- (2) Space-borne passive microwave radiometers have been used since 1978 to retrieve SWE, using a microwave retrieval algorithms that calculates the difference at 19 and 37 GHz (Chang et al.,

1.2. The Challenges of the Snow Observation

1987). The overlaying snow mass absorbs and scatters microwave radiation from the ground surface. As a result, the difference in brightness temperature at these two channels can be linked to snow depth or SWE (Figure 1.3 b). However, plants strongly absorb radiation in 37 GHz range. Moreover, the presence of liquid water significantly increases microwave radiation absorption within wet snowpacks. Additionally, limited resolution of passive microwave data (generally ranging from 25 to 50 km) hinders its effectiveness for SWE estimation (Dietz et al., 2012).

- (3) Active microwave sensor (radar), can be utilized to measure SWE by analyzing backscattering signals (Figure 1.3 c). For instance, synthetic aperture radar (SAR) has a lower frequency and a finer spatial resolution than passive microwave radiometry. However, it is important to acknowledge that the low frequency radar (C-band) can penetrate up to 20 meters in dry conditions, but this depth decreases to only 13.8 centimeters with only 1% liquid water content (Rott & Nagler, 1993). Therefore, C-band radar cannot penetrate wet and icy snowpacks, but also ignore backscattering signals from the dry, thin snowpack (Dietz et al., 2012).
- (4) Lidar, which stands for Light Detection and Ranging, does not directly measure snow depth. Rather, lidar data are acquired over an area prior to snowfall, and repeated later when snow is present to create a map of the snow surface (Figure 1.3 d). The height difference between the bare earth and snow surface is inferred to represent snow depth. Airborne lidar is local interest and costly; thus, it is more successful in mapping snow depth at the watershed scale (Deems et al., 2013).

Due to limitations such as sparse observation networks, satellite retrieval algorithm uncertainties, simplified model parameterizations, spatiotemporal inaccuracies in forcing data (e.g., precipitation), peak seasonal snow mass over the northern hemisphere vary as much as 50% between datasets over the 1981–2010 period with even higher uncertainties for mountain regions (L. R. Mudryk et al., 2015). Despite the limitations in measuring seasonal snow mass, there are several remote sensing techniques and data assimilation models that have shown potential in accurately measuring/simulating snow depth at various scales:

With the advent of laser altimetry (Figure 1.4), satellite altimetry (Trechler & Kääb, 2017) can provide elevation profiles of snow surfaces. Similarly, structure-from-motion techniques can be utilized to determine snow depth (Deschamps-Berger et al., 2020) over mountainous areas where a snow-free DEM is available.

A new empirical algorithm uses cross-polarized observations from Sentinel-1 C-band radar to retrieve snow depth in dry condition at a resolution of 500 m to 1 km (Lievens et al., 2019, 2022). A recent approach proposes the use of a combination of Ku-band and X-band frequencies that will enable separation of various volume- and surface-scattering components (Yueh et al., 2009). Nevertheless, the presence of tall shrubs and trees will strongly absorb the radar backscatter leading to impeded retrievals

1. Introduction

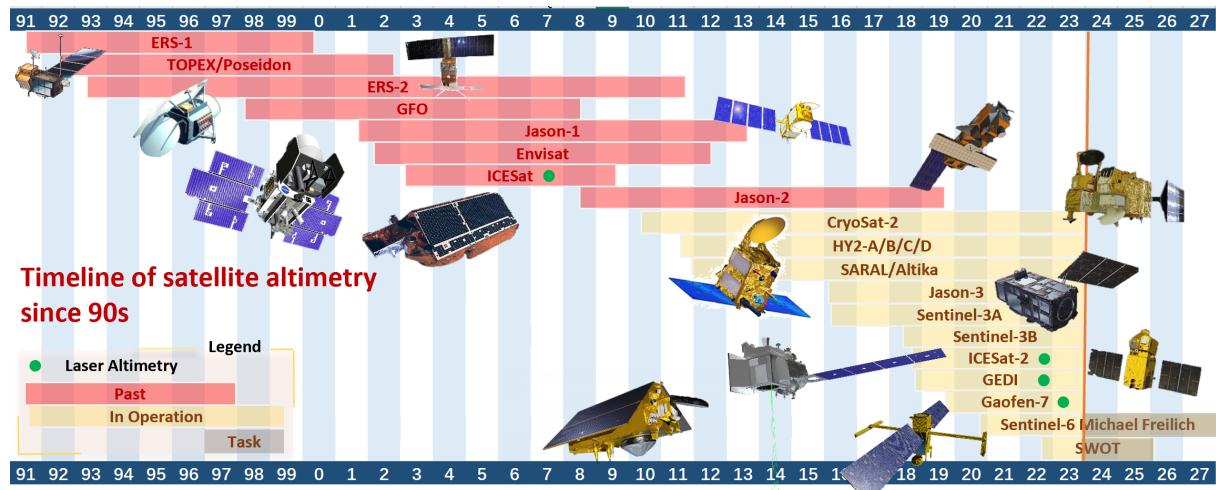


Figure 1.4.: Timeline of Satellite Altimetry since the 1990s. Modified from Abdalla et al. (2021)

when vegetation fraction surpasses 25%. In addition, assessing SWE using radar should consider geometric distortions in mountainous terrain, long repeat orbit time, and the inability to precisely model SWE when snow is vertically inhomogeneous or melting(Nolin, 2010).

With advancements in data assimilation techniques and numerical models, gridded snow datasets are becoming increasingly available. However, they still have limitations in redistributing processes and face challenges at hill-slope scales (more discussion about model products in Section 2.1.1). For regional scales, higher resolution downscaled products offer more realistic patterns of SWE than climate reanalysis (Klehmet et al., 2013). At the hemispheric scale, the latest CMIP6 model has been found to have excessive snow mass and inaccurate simulation of snow extent during the accumulation and snowmelt periods (L. Mudryk et al., 2020). At the global scale, the state-of-the-art snow observation dataset is GlobSnow v3.0. It combines satellite-based passive microwave radiometer datasets with ground-based synoptic snow depth observations using Bayesian data assimilation techniques while incorporating the HUT Snow Emission model (Luoju et al., 2021). However, mountainous areas are still masked out due to complex terrain features (Bormann et al., 2018).

1.3. Objectives

Heavy precipitation and low temperatures at high altitudes result in considerable snow accumulation at elevated locations. Understanding Norway's snow patterns and their variability is crucial for effective natural resource management, climate change adaptation, and mitigation of associated risks. The primary objective of this study is to develop a comprehensive workflow for retrieving and downscaling snow depth using satellite laser altimetry, machine learning, and climate reanalysis. The study

1.4. Thesis structure

will be conducted with a focus on mainland Norway, but the methodology can be applied globally. The specific objectives of this study are as follows:

- (1) Develop a universal co-registration algorithm to efficiently co-register high-resolution DEMs and/or point elevation measurements.
- (2) Benchmark available DEM products using ICESat-2 snow-off segments and develop a workflow to identify and correct biases in the DEM data using machine learning.
- (3) Generate a novel ICESat-2 derived snow depth dataset (2018-2022) covering mainland Norway, incorporating the bias-corrected DEM from objective (2) and the ICESat-2 ATL08 snow-on segments.
- (4) Evaluate the accuracy of the developed workflow by comparing the ICESat-2 derived snow depth dataset with in situ snow depth measurements.
- (5) Develop a statistical downscaling method for snow depth using machine learning techniques and the ERA5 Land climate reanalysis, which can provide a spatially complete snow depth map.

By achieving these objectives step by step, this study will offer a better understanding of snow depth distribution and variation at hillslope scales, particularly in remote areas. Additionally, this study will refine our comprehension of the DEM uncertainties and data quality of the ICESat-2 ATL08 products. Furthermore, this study will contribute a novel training dataset for machine learning that addresses snow distribution at hillslope scales, which is likely the first of its kind across the country.

1.4. Thesis structure

Chapter 1 provides an overview of the role of snow cover, and the challenges associated with snow observation. It also outlines the objectives and structure of the thesis.

Chapter 2 offers a review of the key concepts involved in the variability of snow cover, and the laser altimetry derived snow depth workflow. This includes satellite laser altimetry, DEM uncertainties, co-registration, machine learning, climate reanalysis, downscaling techniques.

Chapter 3 details the workflow process. This chapter covers data preparation, co-registration of DEM and ICESat-2 to generate differential datasets, accounting for DEM uncertainties using snow-off segments from ICESat-2 data, DEM bias correction using ICESat-2 as a reference, snow depth retrieval and downscaling using the XGBoost algorithm. This chapter also includes the introduction of field data.

Chapter 4 presents the result for mainland Norway with in situ validation. The results are then discussed in *Chapter 5* with an emphasis on uncertainties.

Finally, *Chapter 6* offers future prospects and concluding remarks.

2. Scientific Overview

2.1. The Variability of the Snow Cover

As a consequence of the Earth's seasonal cycles, climate patterns, and long-term climate change, snow cover is one of the most variable land surface conditions, both spatially and temporally. Observing and simulating these changes is challenging due to the high degree of region-to-region and season-to-season variability in trends (L. Mudryk et al., 2020).

2.1.1. Spatial variability

The spatial variability of snow accumulation is influenced by various factors such as climate, topography, vegetation, and wind at different scales. While climate is a major contributor to large-scale variations of snow accumulation, smaller scales are affected significantly by the effects of topography, vegetation, and wind (Table 2.1). The relationship between vegetation and snow is complex; for instance, forests can reduce horizontal snow redistribution and trap more snow than open areas. Furthermore, vegetation affects snow interception by altering surface albedo and reducing wind speeds (Freudiger et al., 2017; Trujillo et al., 2007).

The topography of a region plays a critical role in shaping the distribution of snow as it influences various factors such as temperature, precipitation, solar radiation and wind speeds and direction. Wind-driven processes influence the redistribution of snow, ranging from orographic precipitation at large scales to preferential deposition of snowfall, and wind-induced transport of snow on smaller scales. Additionally, avalanches are also a significant agent (Mott et al., 2018) (Figure 2.1).

Table 2.1.: Snow Distribution at Different Scales.

Scales	Distance	Factors
Macroscale	10 km - 1000 km	Dynamic meteorologic effects
Mesoscale	100 m - 10 km	Wind redistribution, terrain and vegetation.
Microscale	10 m - 100 m	Air flow patterns and transportation

2. Scientific Overview

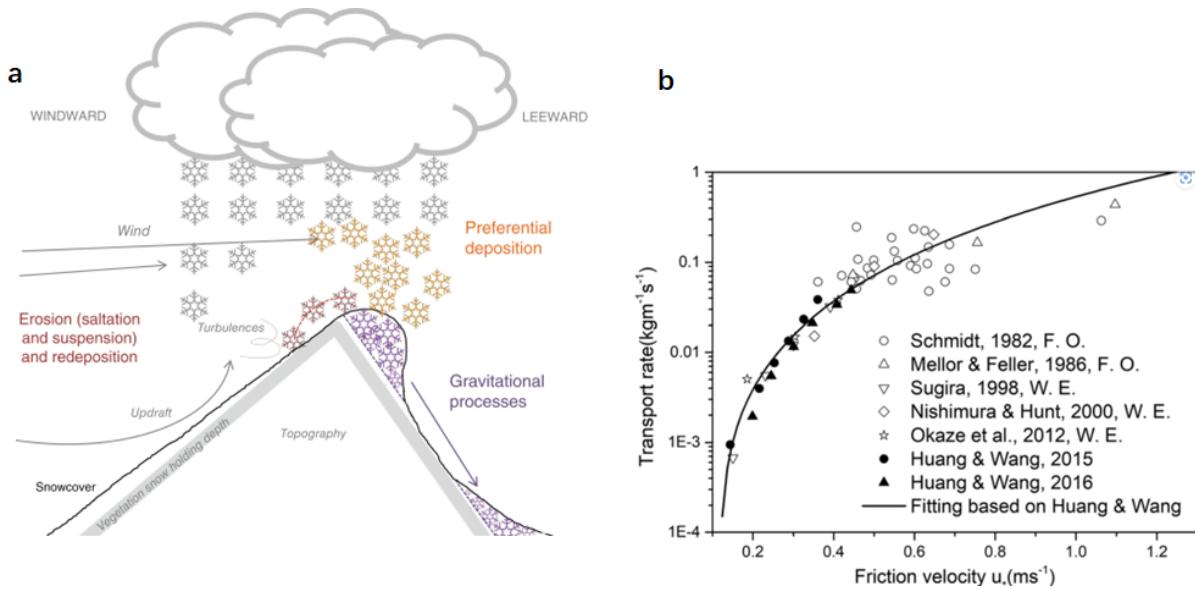


Figure 2.1: Snow Redistribution and Wind. Wind erosion, deposition and gravity processes (avalanche) drive snow redistribution (a). The aeolian research suggests that as wind speed increases, the transport rate increases exponentially with a power of 3 (b). Adapted from Freudiger et al. (2017); G. Li et al. (2018).

Accurately capturing snow depth variability at hill-slope scales remains a challenging task. This is largely due to the complex vegetation-snow interactions and wind redistribution. Additionally, the unavailability of forcing data at hillslope scales, such as precipitation and wind in remote mountains, further complicates the modeling process. As a result, physically based models face many challenges and limitations. They either have limited coverage area or too coarse resolution to address subgrid distribution (Figure 2.2). The statistical model is less complicated, so it can describe the overall spatial variability of the snow cover over large catchment, but also in very fine scales (less than 100 m) to address the subgrid variability (Freudiger et al., 2017).

Meanwhile, the distribution patterns of snow cover exhibit a notable resemblance year after year due to their dependence on topography, vegetation, and consistent synoptic weather patterns (Sturm & Wagner, 2010; Trujillo et al., 2007). This similarity lends support to an empirical approach known as statistical downscaling for snow depth distribution (further discussion in Section 2.5.1)

2.1.2. Temporal variability

Snow cover is subject to vary over time. On average, in January, snow covers about 47% of the land surface in the Northern Hemisphere, while in August, only 3% is covered by snow (Estilow et al., 2015). Snow accumulation changes from year to year. Decadal variability refers to fluctuations that occur

2.1. The Variability of the Snow Cover

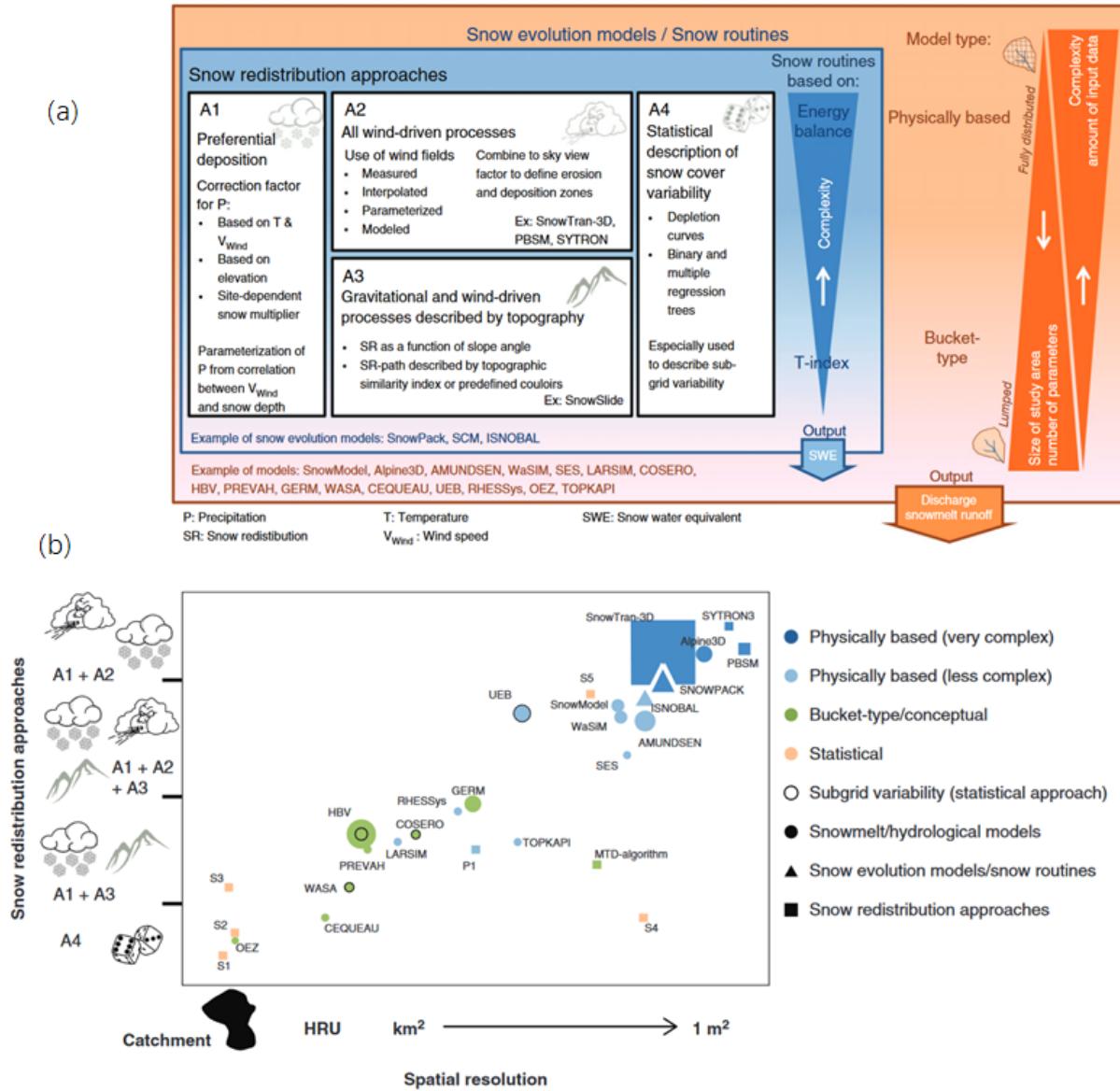


Figure 2.2.: An Overview of Models for Snow Redistribution. These approaches can be broadly classified into four categories: (A1) the correction of precipitation data; (A2) wind-driven processes; (A3) gravitational transportation and/or wind-driven processes; (A4) statistical description (a). The more complex the physical model (the size of objects), the more forcing data is required, and the smaller the study area. Bucket-type hydrological models are usually based on a temperature-index approach. The areas with similar hydrological response, called hydrological response units (HRUs) are aggregated to speed up computation. Statistical models do not specifically model snow redistribution processes but describe the spatial variability of the snow cover in at least catchment scales. Adapted from Freudiger et al. (2017)

2. Scientific Overview

over periods of ten years or more and are frequently associated with long-term climate variability and climate change (see Section 1.1.2).

Over the time, the development of the snowpack is driven by melting energy, which can be expressed in terms of the snowmelt energy balance (Anderson, 1968). This serves as the fundamental principle for numerical modeling of snowpack:

$$Q_m = Q_{ns} + Q_{nl} + Q_h + Q_e + Q_p + Q_g \quad (2.1)$$

Where all Q refer to heat fluxes (W m^{-2}). Q_m is the total heat flux (positive or negative) available for snow melting, Q_{ns} is the short-wave radiation, Q_{nl} is the long-wave radiation, Q_h is the sensible heat flux, Q_e is the latent heat flux caused by water phase changes, Q_p is the heat supplied by precipitation, Q_g is the heat supplied by the ground.

The terrain parameters, such as elevation, are related to temperature, and the aspect is related to solar radiation. The roughness of the terrain and vegetation cover also plays a crucial role in determining wind speeds, which in turn affects the heat flux of the snowpack.

2.2. DEM

2.2.1. DEM products

Digital elevation models (DEMs) are widely used representation of the Earth's surface in a fixed grid, triangulated irregular network (TIN), or point cloud. Two primary types of DEM are the digital terrain model (DTM) and the digital surface model (DSM). The DTM represents the bare Earth without any vegetation or man-made structures, while the DSM includes these features (Rosen et al., 2000). These models can be generated using various space-borne remote sensing techniques such as stereoscopic satellite imagery or Synthetic Aperture Radar (SAR). They can also be created using airborne remote sensing methods like lidar, photogrammetry or ground-based survey methods like laser scanning (X. Liu, 2008; Westoby et al., 2012) at local and regional scales.

Most publicly available DEMs (Table 2.2) have a resolution of 1" (~30 m), which may not be suffice for high-resolution applications. The high-resolution DEMs are now becoming more readily available, such as the Arctic DEM/Earth DEM with a resolution of 2 m (Morin et al., 2016). Furthermore, lidar-based DEMs can offer resolutions as fine as 1 m in many countries. It is essential to note that radar can penetrate dry snow surfaces, resulting in nearly snow-free DEMs. Photogrammetry and lidar-based DEMs offer higher resolution than other radar but are not entirely free of snow in some area, leading to higher surface estimation. Additionally, lidar-based DEM may encounter quality issues when dealing with vegetation or other surface objects.

Table 2.2.: Global DEM Products.

DEM	Acquired	Primary Source	Resolution	
			(m)	Coverage
SRTM (v3)	2000	C band Radar	30, 90	Global except high latitudes
ASTER GDEM	2000-2013	Stereo imagery	30	Global except high latitudes
TanDEM-X	2010-2011	X band Radar	90	Global
Copernicus DEM	2010-2015	TanDEM-X, WorldDEM	30, 90	Global
NASA DEM	2000	Reprocessed SRTM	30	Global
MERIT	2000-2013	Reprocessed SRTM, ASTER	90	Global
ArcticDEM, EarthDEM	2007- Present	Stereo imagery	2	Global (in development)

2.2.2. DEM uncertainties

DEMs suffer from various types of errors and biases from the acquisition to processing stages. Despite the wide range of applications of DEMs and the technological advances used for their creation, quality assessment of DEM remains an open question as there are no specific guidelines available for it (Mesa-Mingorance & Ariza-López, 2020). For instance, an error distribution is typically described by its mean and variance. However, this description cannot fully account for non-stationary errors that are more commonly observed in real-world scenarios (Oksanen & Sarjakoski, 2006). Moreover, the sampling points utilized cannot be considered a reliable representation of the overall accuracy of the DEM. As a result, estimation becomes increasingly challenging under such conditions.

One of the practical approach to evaluate the quality of DEM is to compare it with a reference elevation over stable terrain, excluding snow cover, glaciers, landslides, volcanic uplifts and dense forest, and further analysis residual differences. When referring to discrepancy as errors, outliers, noises or biases; we are either referring to “systematic errors” (accuracy or bias) or “random error” (precision or variance). Figure 2.3 provides a framework of the uncertainties associated with DEMs (Hugonnet et al., 2022).

The systematic bias can result by 3D affine, referring to changes in the translation, rotations or scaling relative to a reference system (Figure 2.3). This type of bias is common across all types of elevation data, but it can be minimized through co-registration techniques (Nuth & Kääb, 2011) (see Section 2.2.3 for further discussion). Additionally, specific systematic errors may arise due to instrument calibration issues (Girod et al., 2017), physical conditions such as snow cover or vegetation, or errors

2. Scientific Overview

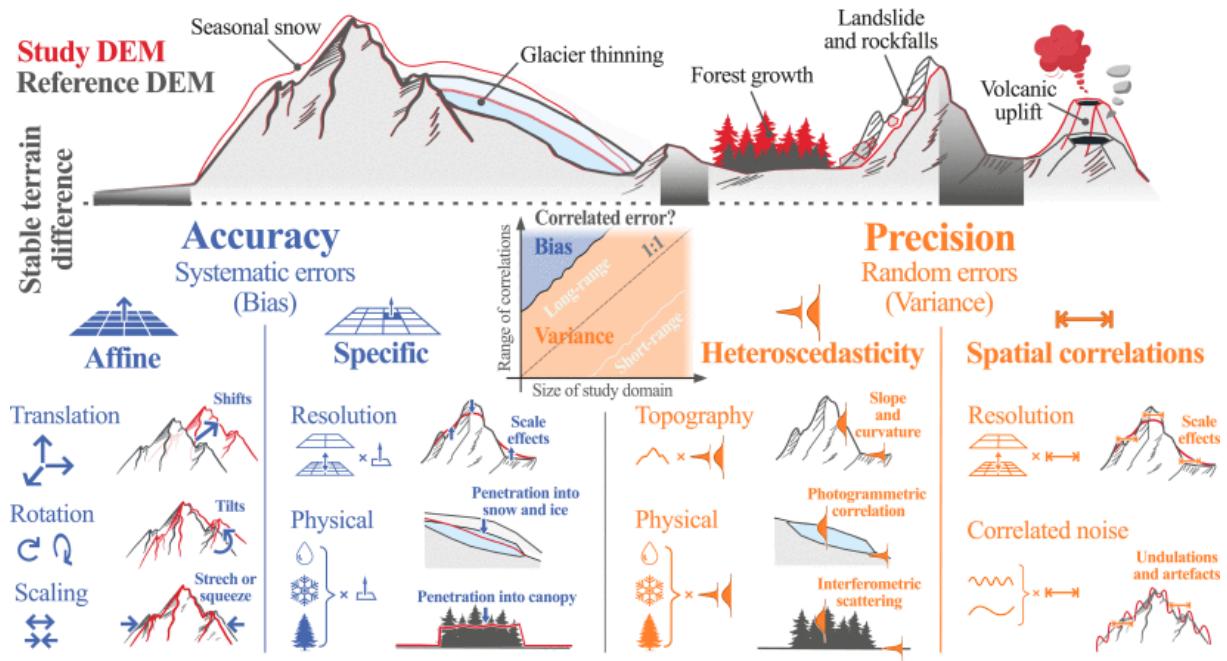


Figure 2.3.: Uncertainty Framework for DEM Assessment. Source: Hugonet et al. (2022)

introduced during data processing. By accurately identifying and characterizing these specific systematic errors, it may be feasible to apply bias correction techniques to improve the overall accuracy of the DEM.

There are various metrics that can be utilized to describe bias, such as the mean error (μ) and median error. Another metric, skewness, is determined by considering the relationship between mean and median. This metric indicates the asymmetry of a probability distribution: whether data is skewed to the left or right of the mean. However, it does not provide information about the magnitude of errors. In practice, applying an overall bias correction (or adjustment) can make median error or mean error to technical zero. However, this does not imply that the dataset is entirely free of bias. Additionally, researchers assume that a ‘bias-free’ dataset should also have minimal skewness, which requires a specific bias correction to adjust the error distribution (Figure 2.4).

Heteroscedasticity is a phenomenon where the variance changes across different levels of the independent variables (Figure 2.4), such as surface slope (Figure 2.5), curvature and other terrain parameters. Unlike the mean, which can indicate the direction of error, heteroscedasticity does not exhibit any direction of error but instead reveals the magnitude of the spread of errors. While improving heteroscedasticity is often challenging; it is possible to identify and eliminate measurements that do not conform to ideal standards. For example, in practice, outliers are often identified using $3 \times \text{RMSE}$ (Root Mean Squared Error) or $3 \times \sigma$ (Standard Deviation). Additionally, describing the variance under different conditions is necessary to represent the precision of the DEM.

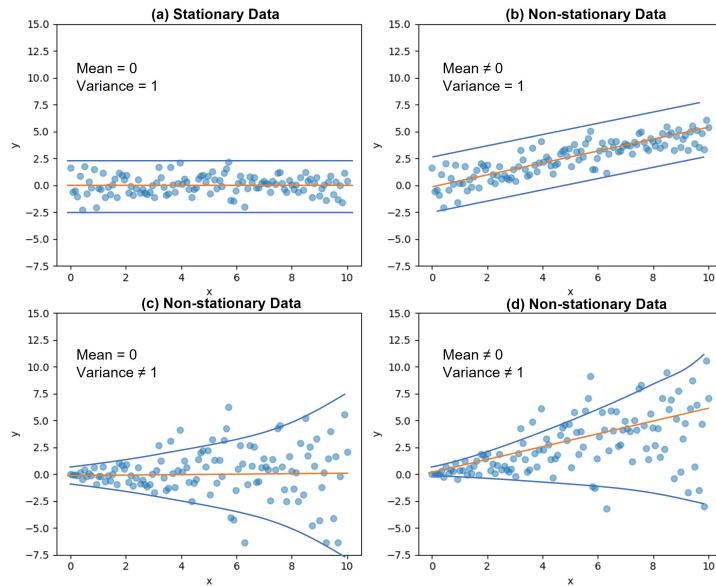


Figure 2.4.: Non-stationary and Heteroscedasticity. The stationary errors follows the normal distribution (a). More often, the presence of nonstationary error can lead to biased estimates (b,c,d). If the relationship between the independent variables and bias can be explained, the bias can be corrected (b,d). Heteroscedasticity refers to changing variances in the data, which can also lead to biased estimates (c,d). It is important to identify and address non-stationarity and heteroscedasticity in accurate analysis.

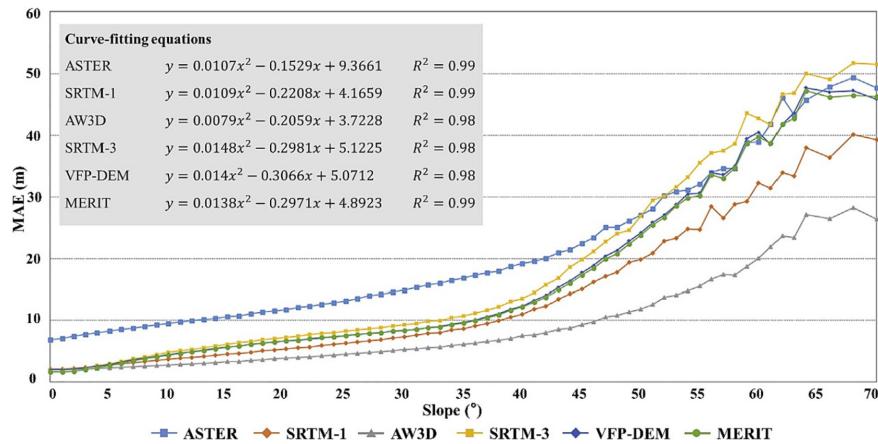


Figure 2.5.: Mean Absolute Errors of Global DEM Products in the Tibetan Plateau. Surface slope is an important factor affecting DEM errors. The ASTER, SRTM, AW3D, VFP-DEM and MERIT are global DEM products. Source: Xiong et al. (2022).

2. Scientific Overview

It is important to note that certain measures, such as mean, root mean square error (RMSE), and standard deviation, are significantly susceptible to the presence of outliers. As a result, to ensure more accurate and reliable results, robust estimators such as the median, mean absolute error (MAE), and normalized median absolute deviation (NMAD) are recommended (Höhle & Höhle, 2009). The NMAD is proportional to the median of the absolute differences between errors and the median error.

To perform uncertainty analysis, the following metrics are used in this study:

$$\text{Mean Error } (\mu) = \frac{1}{n} \sum \delta h \quad (2.2)$$

$$\text{Pearson's second skewness} = \frac{3 \cdot (\mu - m_{\delta h})}{\sigma} \quad (2.3)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum \delta h^2} \quad (2.4)$$

$$\text{MAE} = \frac{1}{n} \sum |\delta h| \quad (2.5)$$

$$\text{STD } (\sigma) = \sqrt{\frac{1}{n-1} \sum (\delta h - \mu)^2} \quad (2.6)$$

$$\text{NMAD} = 1.4826 \cdot \text{median}(|\delta h - m_{\delta h}|) \quad (2.7)$$

where δh is the sample of elevation difference and n is the number of samples. $m_{\delta h}$ is the median of the samples.

2.2.3. DEM Co-registration

DEM often suffer from misalignment, which poses a challenge for accurate change detection and analysis (Paul et al., 2015). When comparing two DEMs in elevation differencing workflow, co-registration of two datasets is the initial step. One commonly used approach from computer vision is the Iterative Closest Point (ICP) algorithm (Besl & McKay, 1992). This algorithm is capable of co-registering two sets of rigid point clouds into a single coordinate system and can also be applied to DEM co-registration. This algorithm iteratively matches corresponding points, computes the transformation that aligns the points, and then applies the transformation to the points until convergence. There are numerous co-registration methods available in computer vision, differing mainly in how they minimize differences due to mismatching. For instance, the Least Squares 3D

2.3. ICESat-2 Altimetry

Surface Matching (LS3D) algorithm treats it as a least square process of the Euclidean distances of overlapping surfaces (Gruen & Akca, 2005). Feature point-based methods reduce computation by solving the problem on selected points such as centroids of sub-watersheds (H. Li et al., 2017). All these methods require costs to establish correspondences.

The NuthKaab co-registration method (Nuth & Kääb, 2011) elegantly solves co-registration issue by identifying that statistic patterns of elevation differences that are related to slope and aspect. This method has become a standard procedure in high-accuracy cryosphere applications (W. Chen et al., 2022; Hugonnet et al., 2021; Vacaflo et al., 2022) due to its efficiency and effectiveness. However, this method is not efficient for high-resolution DEMs and initially does not support rotations. Grid search methods shift the DEM stepwise in a predefined window but are considered rarely used due to their brute-force search process being computationally expensive (T. Li et al., 2022). This study recognizes these shortcomings and addresses them with a new algorithm (see Section 3.4).

2.3. ICESat-2 Altimetry

The major goal of satellite altimetry is to measure the shape of the Earth, observing the elevation changes on Earth surface, by the round trip travel time of a short electromagnetic pulse sent from an orbiting spacecraft toward the planetary body (Abdalla et al., 2021). The first altimetry measurements were made by radar instruments on board the Seasat-1 and Geosat satellites, which launched in 1978 and 1985 respectively. In 2003, National Aeronautics and Space Administration (NASA) launched the Ice Cloud and Land Elevation Satellite (ICESat) (Schutz et al., 2005), which is the first laser altimeter, marked a major advancement in satellite altimetry. It reveals unprecedented details, such as sea-ice freeboard and thickness (Kwok & Cunningham, 2008), Antarctic ice-sheet loss (Pritchard et al., 2012), subglacial lake drainage (Palmer et al., 2015). The laser altimetry has several advantages over radar altimetry, including smaller footprints, higher spatial resolution, and ability of detecting both ground surfaces and canopies, and less penetration on snow surface (Deems et al., 2013). However, laser altimetry is more susceptible to atmospheric conditions and has a narrower field of view that limits its coverage range (Wang et al., 2011).

As the successor to ICESat, ICESat-2 was launched in September 2018 by NASA and provides world-wide photon counting lidar data over land surfaces, vegetation canopies, water bodies, sea ice, and ice sheets (A. Neuenschwander & Pitts, 2019). ICESat-2 operates at an orbit height of approximately 486 km with a near-polar orbit (92° inclination) and a 91-day cycle. During this cycle, the ground track shifts laterally to increase coverage. The instrument has three beam pairs separated at 3.3 km, with each beam having a strong and weak beam in 90 m (Figure 2.6). There is a power differential of approximately 4:1 between the strong and weak beams that allows for high- and low-reflectance surfaces to be measured (Markus et al., 2017).

2. Scientific Overview

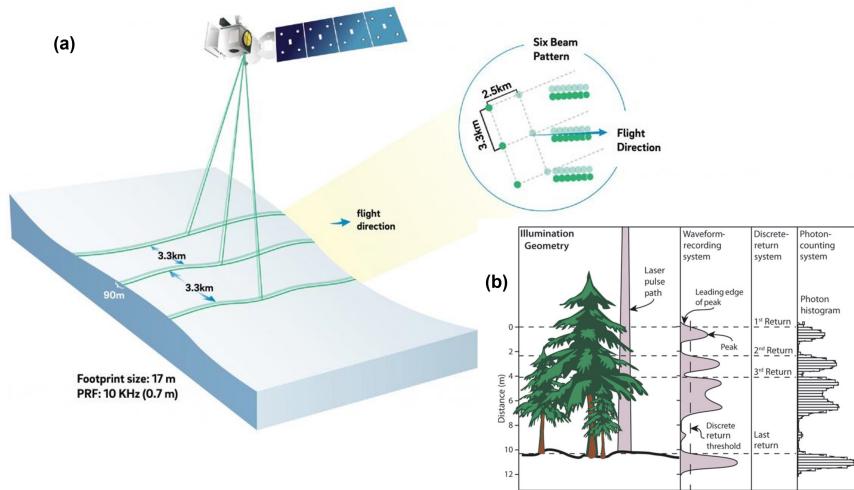


Figure 2.6.: ICESat-2 and Photon Counting System. (a) ICESat-2 is equipped with six laser beams that are arranged into three pairs, allowing for improved coverage (adapted from Deems et al. (2013)) (b) Unlike its predecessor, ICESat-2 employs a low-power system known as photons counting lidar to manage the signals generated by the laser beams. This system uses probability distribution function (PDF) analysis to process the signals (adapted from Smith et al. (2019))

When ICESat-2 flies over an area, it emits 532 nm laser pulses at a repetition rate of 10 kHz from the Advanced Topographic Laser Altimeter System (ATLAS). Approximately 10^{14} photons travel to hit the surface with a spread of approximately 13-17 m. The instrument measures surface elevations every 70 cm by detecting photons returning back to ATLAS telescope's focal plane as many photon events (Figure 2.6) (Markus et al., 2017; A. Neuenschwander et al., 2022; A. Neuenschwander & Pitts, 2019). The individual photon events are geolocated into an XYZ point by measuring traveling time and the precise location of satellite. To reproduce the surface elevation profile, individual photon events must be accumulated and statistically analyzed (Figure 2.6 b). Hence, the elevation precision depends on the number and distribution of returned photons, which varies according to surface geometry, reflectance, atmospheric effects, and solar background noise (Neumann et al., 2021; Smith et al., 2019). Higher availability of photons translates to higher signal-to-noise ratio and typically better elevation precision.

2.3.1. Addressing DEM errors by ICESat-2

The laser altimetry technique is highly precise in providing ground elevation control points for DEMs, as demonstrated by previous studies involving ICESat (Gruber et al., 2012; Yamazaki et al., 2017). Due to its limited spatial coverage, it cannot be used independently to create moderate resolution DEMs. Nevertheless, the latest ICESat-2 mission presents new opportunities for assessing DEM uncertainties

2.3. ICESat-2 Altimetry

(W. Chen et al., 2022; H. Li et al., 2022). Recent research proposes a two-dimensional bias correction method (Magruder et al., 2021) for SRTM using data obtained from ICESat-2 ATL08 product and vegetation product from Landsat, leading to enhanced accuracy as shown in Figure 2.7.

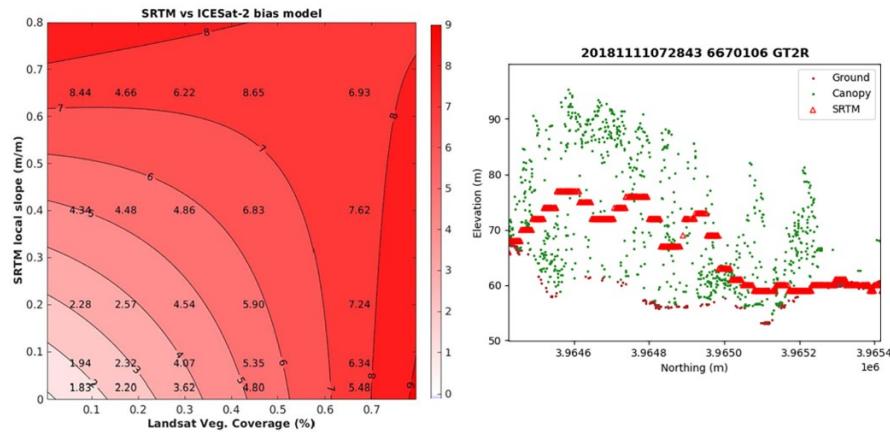


Figure 2.7.: A Bias Correction Scheme for DEMs Based on Slopes and Vegetation Coverage. Flat terrain with sparse vegetation have lower bias (a). An example transect of the comparison demonstrate that (b). Source: Magruder et al. (2021)

However, Tian & Shan (2021) discovered that the ATL08 product tends to underestimate terrain height in areas with significant terrain slopes while overestimating it in regions covered by dense vegetation. This was observed through a comparison with a 1m lidar reference DEM. Another case study (A. Liu, 2021), utilizing the lidar reference DEM in the same year of acquisition, has revealed that the ATL08 mean terrain height exhibit a error correlated to slope. Specifically, the RMSE increases from 0.6 m to 7.5 m for slopes ranging from less than 5° to greater than 30°. These findings emphasize the need for comprehensive evaluation before utilizing the ICESat2 products in snow depth retrieval, particularly when dealing with complex terrain.

2.3.2. Snow depth from ICESat-2

Since the first attempt on ICESat (Treichler & Käab, 2017) to derive snow depth in mountain areas by differencing satellite laser snow-on measurements with snow-off DEM, there are several attempts on ICESat-2:

- Snow depth derived from ATL06 yielded an accuracy of 0.2 m (bias), a precision (NMAD) of 0.5 m for low slopes and 1.2 m for steeper areas, compared to airborne lidar snow depth over the upper Tuolumne basin (California, USA) (Deschamps-Berger et al., 2022).
- Hu et al. (2022) found that the ‘crossover differences’ (snow-on-snow-off-differencing) shows accurate snow depth with RMSE of 4.20 cm over flat area (slope < 1.5°) and thin snowpack (< 40

2. Scientific Overview

cm) in Altay, northwest China.

- A case study in Alaska's Kenai Mountains (Enderlin et al., 2022) found the snow depth had MAD varying from 0.2 m for slopes $<5^\circ$ to over 1 m for slopes $> 20^\circ$. The ATL08 are only suggested to use in unglacierized mountain regions with relatively low-slopes and sparse vegetation cover.

However, Deschamps-Berger et al. (2022) do not correct bias or inconsistency between the reference DEM and ICESat-2 elevation measurements, which introduces uncertainties into the final results. Additionally, using Copernicus GLO30 for snow depth retrieval was considered too imprecise (Enderlin et al., 2022). While slope-dependent adjustments partially remove the negative skewness of ICESat-2 ATL08 in some cases, large uncertainties remain (Enderlin et al., 2022), indicating that a more comprehensive approach is needed to address nonstationary bias.

Most remote sensing techniques, including ICESat-2, are constrained by temporal and spatial limitations. The poor temporal and spatial coverage of ICESat-2 derived snow depth limits its utility. One potential solution to this issue is to assimilate snow observations into a snow model (Alonso-González et al., 2022; Girotto et al., 2020) or use it in dynamical downscaling with climate models (Wang, 2022) to produce spatially complete products. Alternatively, statistical downscaling can be used to develop an empirical relationship between the subgrid snow distribution and the snow depth from climate models or climate reanalysis (Further discussion in Section 2.5.1).

2.4. Machine Learning and XGBoost

Machine learning (ML) is a type of artificial intelligence that involves the development of systems capable of learning from training dataset to identify patterns and relationships, and subsequently make predictions or decisions on new data.

Two main types of machine learning exist: supervised and unsupervised. In supervised learning, a mapping function (F) is established based on labeled input variables (X) and output variables (y):

$$y \sim F(X) \quad (2.8)$$

Where the mapping function can be as simple as linear regression or as complex as tree structures:

- Linear regression. The function describe the linear relationship between X and y .
- Logistic regression. The function based on linear relationship but output binary results.
- Decision trees. The function compose of multiple leaf of the tree (node), each internal node represents a decision or test on input variable. The goal is to find a set of decision rules that minimize the error and make a correct decision (Kingsford & Salzberg, 2008).

- Random forests (Breiman, 2001). The function combines multiple decision trees in parallel to improve performance and reduce overfitting. As it average the predictions of all the models, so this ensemble learning algorithms call “bagging” (bootstrap aggregating).
- GBDT (Gradient-Boosting Decision Tree). An ensemble learning algorithm that uses sequential weak trees and updates input data point weights based on residuals for improved performance. “Boosting” refers to the fact that each weak learners contribute sequentially.
- XGBoost (Extreme Gradient Boosting). A specific implementation of GBDT that uses optimization techniques (Rashmi & Gilad-Bachrach, 2015) to accelerate training processes while improving accuracy (T. Chen & Guestrin, 2016).

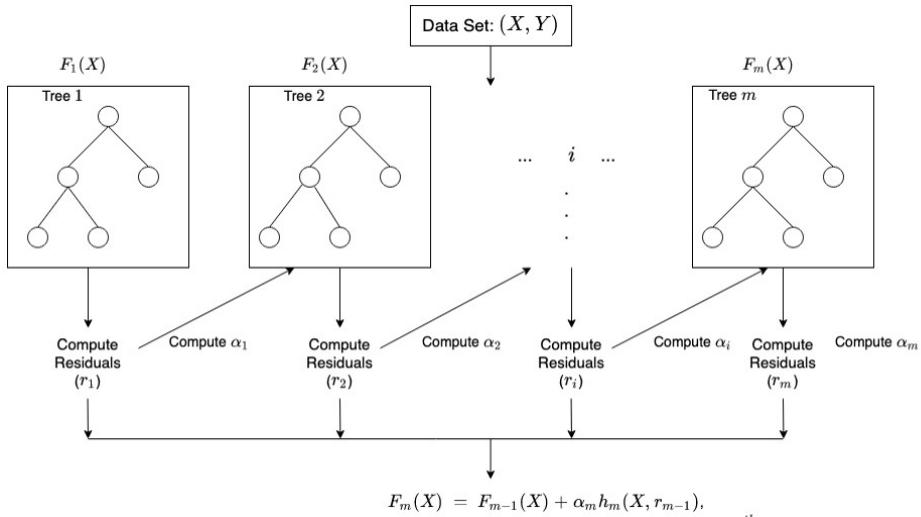


Figure 2.8.: How XGBoost Works. XGBoost builds a series of decision trees iteratively. Each decision tree tries to correct the residual errors γ_i of the previous tree, so that the final prediction $F_m(X)$ is a combination of all the trees. Regularization parameters α_i are added to the objective function as penalty terms to discourage overfitting and promote better generalization of the model. The function is solved by compute. Source: Developer guide by Amazon (2023)

XGBoost builds an ensemble of decision trees. Each tree consists of nodes that represent decision rules, which split the data into smaller subsets based on the input features. The objective function measures the difference between the predicted values and labeled actual values, and is optimized through gradient boosting. The final prediction is the sum of predictions from all the individual trees in the model (Figure 2.8). Additionally, XGBoost contains additional regularization techniques, including L1 and L2 regularization terms, and a built-in mechanism for handling missing values in the input data. These regularization techniques help prevent overfitting and improve the generalization performance of the model.

2. Scientific Overview

2.4.1. Hyperparameters of XGBoost

By default, XGBoost uses the *reg:squarederror* objective function, which aims to minimize the mean squared error (MSE) between the predicted and actual values. This objective function is generally suitable for most regression tasks. On the other hand, *reg:absoluteerror* is another objective function that minimizes the mean absolute error (MAE) between the predicted and actual values. MAE is more robust to outliers than MSE. For this reason, this study has opted to use *reg:absoluteerror* in the context of DEM error and snow depth. Furthermore, there are several hyperparameters that can be tuned to improve the performance of XGBoost:

- *n_estimators* (default 100) – This parameter determines the number of trees in the ensemble. Increasing this value results in more weak learners contributing towards the final output but can significantly slow down training time.
- *max_depth* (default 6) – This parameter controls the complexity of the algorithm. A lower value decreases the algorithm's ability to detect patterns (underfitting), while a higher value risks overfitting by making the model too complex.
- *min_child_weight* (default 1) – Overfitting can be prevented by limiting tree depth, which is achieved with this parameter. A higher value reduces chances of overfitting on training data.
- *learning_rate* (default 0.3) – The rate at which the model learns is inversely proportional to its accuracy. Lowering this parameter improves a model's ability to detect patterns but requires longer training time. Setting it too low may hinder convergence.
- *gamma* (default 0) – This regularization parameter influences how strongly regularization should be applied, with higher values reducing chances of overfitting but increasing risk of underfitting if set too high.
- *colsample_bytree* (default 1.0) – This parameter determines what fraction of features/predictors should be used in each tree during training. As different trees may use different features, overfitting can be reduced and training speed improved. The value ranges from 0 to 1.
- *subsample* (default 1.0) – Similar to *colsample_bytree*, this parameter determines what fraction of instances should be used in each tree during training. It also reduces overfitting and improves training time.

Note that these hyperparameters may require adjustment based on the dataset being utilized. A popular method for hyperparameter tuning is grid search.

2.5. Climate Reanalysis and Downscaling

Climate data archives and regional observations often suffer from gaps, inconsistencies, biases, and uncertainties due to various sources and methods. To provide a comprehensive understanding of

2.5. Climate Reanalysis and Downscaling

past and present climate conditions, several climate reanalyses have been developed through the assimilation of data from diverse sources into numerical models that simulate the Earth's atmosphere, land, and oceans. These reanalyses offer a physically consistent framework for describing long-term weather patterns and climate variability (Dee et al., 2014; Laloyaux et al., 2016).

ERA-Interim, ERA5 and ERA5 Land are widely used reanalysis datasets produced by the European Centre for Medium-Range Weather Forecasts (ECMWF). ERA-Interim is the previous generation of ECMWF reanalysis that covers the period from 1979 to 2019 with a spatial resolution of about 80 km. ERA5 is the latest generation of ECMWF reanalysis that covers the period from 1950 to present with a higher spatial resolution of about 30 km and more atmospheric levels (Dee et al., 2011; Hersbach et al., 2020). ERA5 Land is a subset of ERA5 that produces a total of 50 variables describing the water and energy cycles over land, globally, hourly, and at a spatial resolution of 9 km (Muñoz-Sabater et al., 2021).

However, reanalysis datasets still have quality issues due to sparse and unreliable observations, and bias from inconsistency of models. The coarse resolution also limits its applicability for regional studies, which has motivated downscale of global reanalysis data into finer-scale regional reanalysis data (Wilby & Wigley, 1997). There are two main methods of climate reanalysis downscaling: dynamical and statistical.

- (a) Dynamical downscaling involves using physically based models that simulate atmospheric dynamics and physics at finer resolutions, with global reanalysis or global climate model data serving as boundary conditions. While this approach can provide a full suite of climate variables with good spatial coherence and physical consistency, it is computationally expensive. (Fowler et al., 2007). The accuracy of this method is affected by uncertainties from forcing data and model parameterizations.
- (b) Statistical downscaling is a method that utilizes empirically based formulas, such as regression models, to establish a statistical relationships between large-scale reanalysis and local observations of climate variables (Maraun & Widmann, 2018). This approach is sensitive to the quality and representativeness of the input data and may not capture the complex physical processes and feedback at the local scale, especially for extreme events (Fowler et al., 2007).

The statistical downscaling approach assumes that there exists a stationary empirical relationship between large-scale and local observations, which is commonly referred to as “perfect prognosis (PP)” by climatologists. The downscaling can be applied not only to climate reanalysis but also to model-based atmospheric variables, referred to model output statistics (MOS). MOS focuses on downscaling model-based predictors that are often biased from observations. Hence, this approach is known as bias correction (BC) downscaling (Maraun & Widmann, 2018).

In summary, both dynamical and statistical downscaling have advantages and limitations in providing high-resolution climate data. The choice of method should depend on specific research goals,

2. Scientific Overview

available computational resources, quality of input data, and other relevant factors such as trends and extremes.

2.5.1. Statistical downscaling and subgrid variability of snow depth

Previous studies have successfully downscaled topography-based variables such as temperature, humidity, wind speed, and radiation (Fiddes et al., 2022). However, the high spatial variability and limited observations of snow depth have hindered its downscaling. To address this issue, Fiddes & Gruber (2014) simulated forcing data first and utilized it in snow models to derive downscaled snow depth. Pflug et al. (2021) presents a snow deposition downscaling routine that utilizes pattern-based approaches. The method involves assuming interannually repeatable snow patterns and unbiased mean snowfall. The authors applied the technique in the California Tuolumne River Watershed, where they downscaled 6 km snowfall from an atmospheric model to 25 m resolution using snow depth pattern data from seven different years. The simulations resulted in a correlation coefficient of $r=0.76$ during peak-snowpack timing. This downscaling only account for snowfall so that the snow deposition downscaling that was too spatially heterogeneous.

Grünewald et al. (2013) used high-resolution lidar snow depth in seven different mountain regions and applied statistical modeling to assess the spatial variability of the snow depth. The results showed that by aggregating snow accumulation on a larger scale, up to 91% of the snow depth variability can be explained by local models calibrated to specific study areas. Elevation, slope, northing, and wind sheltering parameters were found to be good predictors of snow distribution. However, coefficients differed among the catchments. A general model combining all data from different regions could only explain 23% of the variability. Models developed in one peak snow season were found to be good predictors for other peak snow seasons.

Some studies do not downscaling the snow depth but focus on the subgrid distribution of snow depth. For instance, Melvold & Skaugen (2013) conducted a snow survey in Norway's Hardangervidda mountain plateau during the peak of the snow accumulation season in 2008 and 2009. The study fount significant variability in local snow depth, ranging from 0 to 10 meters, but similar spatial patterns of accumulation between both years. Although the SeNorge model, a degree-days-based snow model with a resolution of 1 km, captured the general trend, it exhibited notable differences at the 1 km scale. This suggests that SeNorge is unable to account for physical processes associated with subgrid-scale variability.

Another study by Gisnås et al. (2016) found that the subgrid distribution of snow more closely follows a gamma distribution than a lognormal distribution. The covariance can be utilized to describe this subgrid distribution. The authors statistically derived snow distributions within 1 km² grid cells and applied them in permafrost models.

2.5. Climate Reanalysis and Downscaling

In summary, there is a growing interest in investigating the wind redistribution of snow and understanding its accumulation patterns at a subgrid level. Previous studies have relied on airborne laser technology to investigate these patterns, which have shown that they tend to repeat over time (Grünewald et al., 2013; Melvold & Skaugen, 2013). However, these studies have limitations as they only survey the maximum snow depth at the end of the snow season and cover limited study areas that may not be generalizable to other catchments due to data availability and climate differences (Grünewald et al., 2013).

To overcome these limitations, ICESat-2 derived snow depth data is a good data source as it offers a high-resolution dataset that covers a wide range of areas. And a machine learning based downscaling methods has emerged as a promising approach for downscaling climate variables (Hobeichi et al., 2023; Sachindra et al., 2018). This approach effectively manages large datasets with numerous predictors and captures complex relationships between predictors and responses while accounting for nonlinearities in the data. Therefore, it can offer a potential solution for the subgrid snow distribution.

3. Study Area and Methodologies

3.1. Research Design

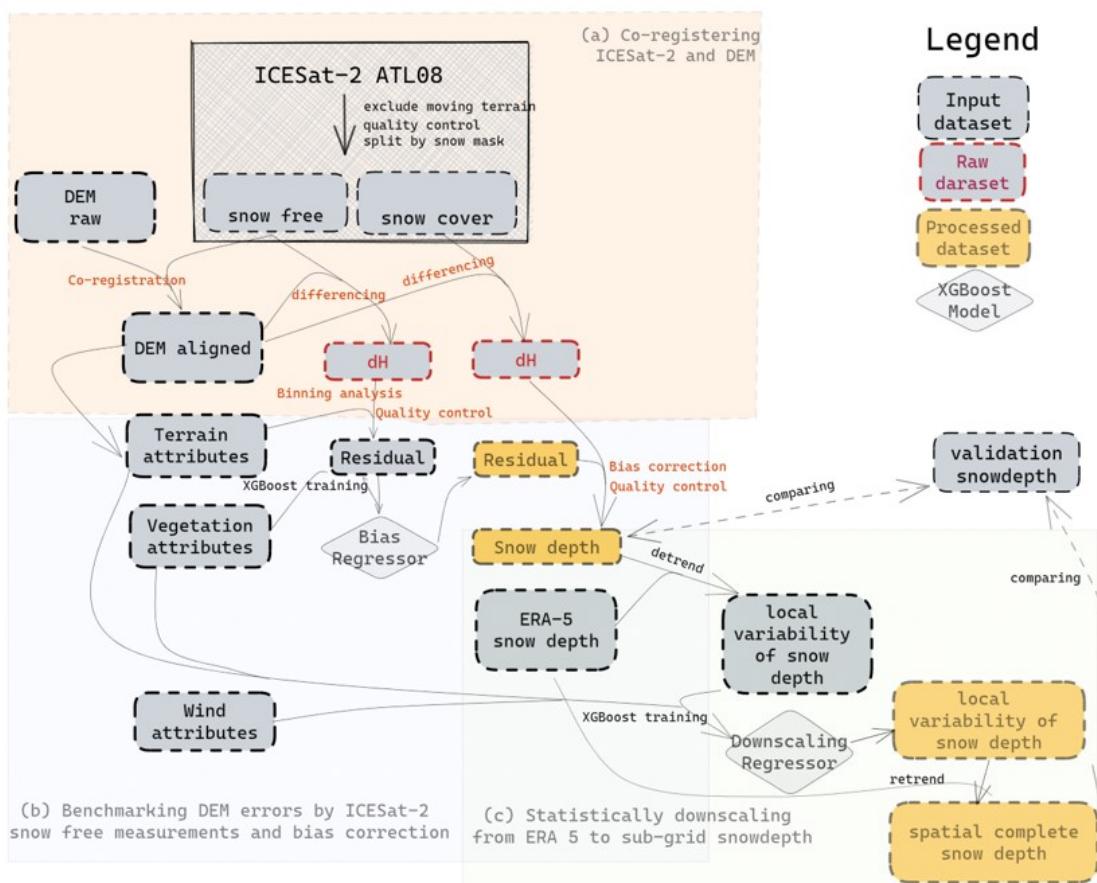


Figure 3.1.: Study Design Schematic. Firstly, DEMs are co-register to generate raw elevation differences (a). The snow-off segments, terrain and vegetation features are utilized as training data for a regressor that can eliminate biases for snow-on segments (b) Finally, another regressor for downscaling is trained and implemented to predict local variability of snow depth in any location and at any time (c).

The calculation of snow depth from lidar data requires two co-registered surface elevation data col-

3. Study Area and Methodologies

lections, one for snow-free bare-ground and one for snow-covered, followed by elevation differencing (Deems et al., 2013). This approach also works in the case of satellite laser altimetry. To implement this approach, several technical challenges need to be addressed.

- The magnitude of seasonal snow surface elevation change is relatively small compared to the uncertainties associated with DEMs. Hence, it is crucial to use the most advanced, bias-free DEM products to achieve accurate measurements. However, obtaining such data may not always be feasible due to factors such as cost and accessibility. Furthermore, this methodology requires a snow-off DEM; thus high-resolution Arctic DEMs are not applicable.
- While the commonly used NuthKaab co-registration (Nuth & Kääb, 2011) is effective in most cases, it struggles with different resolutions and computationally-intensive tasks. Particularly for fine resolution and large scale DEMs required by this workflow. And this process can be time-consuming.
- Biases are widely present in DEMs. It calls the need of quality control and bias correction.
- The satellite's acquisition pattern is sparse both in time and space resulting in a need for additional data to produce a spatially complete snow-depth map.

To address these challenges effectively, three key steps are necessary: co-registration of datasets, bias correction, and statistical downscaling as shown in Figure 3.1. The ICESat-2 data is first split into two categories: snow-free and snow-on segments. Subsequently, the snow-free segments are used in co-registration and fed into bias correction of DEMs. The same bias corrections are then applied to derive snow depth. Finally, the derived snow depth is used in downscaling of snow depth from ERA5 Land . The terrain parameters, vegetation and wind filed from ERA5 Land are used for regression features.

3.2. The Physical Geographical Setting

Norway is located on the western side of the Scandinavian Peninsula in northwestern Europe, situated between latitudes 57° 58' N and 71° 11' N and longitudes 4° 40' E and 30° 58' E. The country's heterogeneous climate is influenced by its diverse topography, ranging from lowland valleys to high mountains. This section provides an overview of Norway's climate characteristics, including patterns of snow distribution, precipitation, wind, and temperature fluctuations.

Snow distribution and variability are primarily influenced by moisture and temperature and vary based on altitude and latitude. The North Atlantic Current, moderates the range of temperatures throughout the year, particularly along the coastal areas of Norway (Kaspi & Schneider, 2011; Keil et al., 2020). During winter, the air temperature is usually lower than the water temperature of the current, resulting in higher evaporation rates (Seppälä, 2005). Moist air masses are carried with the westerly winds, leading to heavy precipitation averaging from 1000 to 3000 mm per year in coastal areas when moist air is uplifted by the mountains. Conversely, inland regions experience drier and

3.3. The Data Setting

colder conditions due to the continental climate, such as Dovrefjell (Fokstugu: 435 mm) (Ketzler et al., 2021). This gradient of climate from the western coast to the eastern inland is reflected on distribution of glaciers. In southwestern Norway, the largest glacier in mainland Europe, Jostedalsbre Glacier, covering an area of approximately 487 square kilometers. The North Norwegian coast also witnesses remarkable snow accumulation with an extended duration of snow extent. Here, Glacier Svartisen is the second-largest glacier in Norway. Furthermore, snow accumulation varies from year to year; for instance, winter 2020-2021 experienced significant low snow conditions compared to recent years (Figure 3.2 a,b,c,d).

Coastal regions are exposed to strong winds due to differences in temperature between land and sea. Mountains act as barriers deflecting wind and creates calmer conditions inland. Among these mountains, Hardangervidda is the largest mountain plateau in Europe, with an average elevation of over 1,100 meters. The relatively flat terrain on top of the plateau makes it prone to high winds with intense snow conditions. Finnmarksvidda Plateau, located in the northernmost part of Norway, is also known for its strong winds, particularly in the winter with extreme cold temperatures (Ketzler et al., 2021) (Figure 3.2 and A.1).

3.3. The Data Setting

3.3.1. ICESat-2 ATL08

The ATL08 product (version 5) from ICESat-2 (A. L. Neuenschwander & Sheridan, 2021) provides global geodetic measurements of terrain surface and relative canopy heights. There are 3968 ATL08 granules across mainland Norway from 2018 Oct 14 to 2022 Oct 12. After dropping NaN, there are 14,801,830 segments, where 5,443,945 snow-free segments are on land, excluding moving terrain (permanent ice and inland water), and 9,213,030 segments are with snow cover over land (Figure 3.3). The measurements consist of fixed segment sizes of 100 meters along the ground track. The following features of the ATL08 segments have been utilized:

- *latitude and longitude*: representing the location of the midpoint of the segment.
- *date*: representing the acquisition date.
- *h_te_best_fit*: representing the best-fit terrain elevation at the 100-meter segment midpoint location, which is determined by polynomial fitting to terrain photons with slope correction and weighting applied (more technical details in Appendix A.2). Some studies (A. Neuenschwander & Pitts, 2019; Tian & Shan, 2021) found that *h_te_best_fit* has slightly higher accuracy and smaller variation than the *h_te_interp*. This study confirmed this by comparison (not shown).
- *n_te_photons*: representing the number of terrain photons within the segment.
- *h_te_skew*: skewness of terrain height within the segment.

3. Study Area and Methodologies

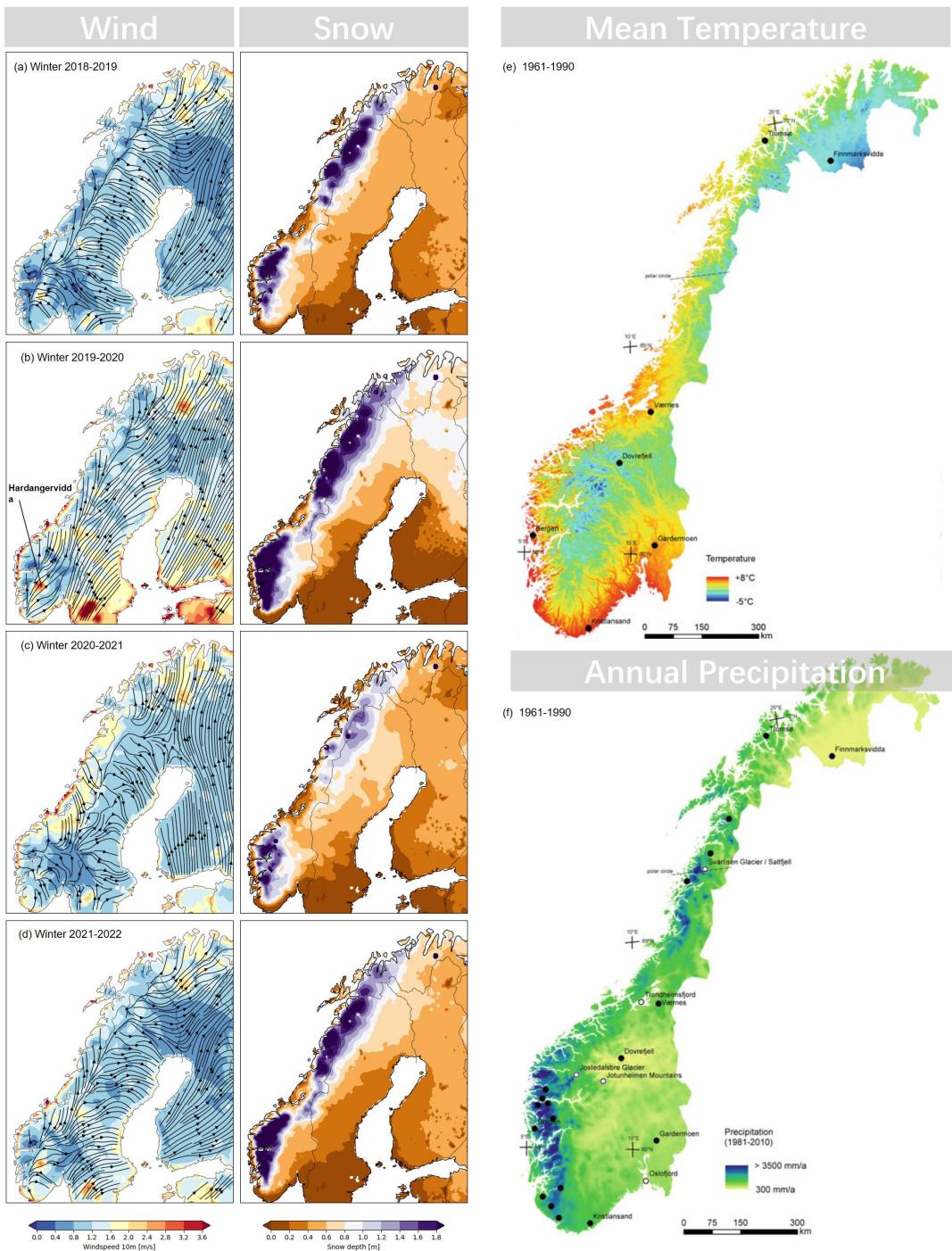


Figure 3.2.: The Climate of Norway. Wind patterns and snow distribution during winter (December to March) from 2018 to 2022 on the Scandinavian Peninsula (a, b, c, d). The average air temperature for the period of 1961–1990 is shown in (e), while precipitation is based on station data and geostatistical modelling for the same period (f). Notably, the ten most intense precipitation events from 1950 to 2008 are marked with black dots in the western coastal foreland. Figures e and f are from Ketzler et al. (2021). Data source: Lussana (2018); Muñoz Sabater (2021b).

3.3. The Data Setting

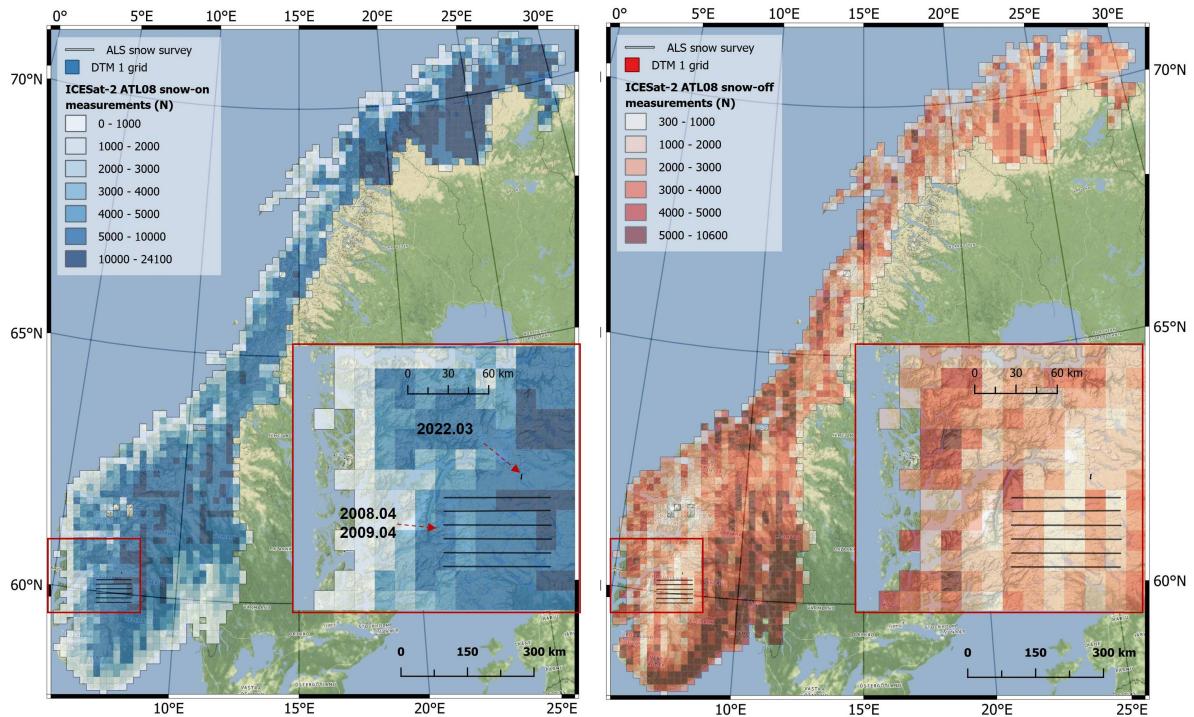


Figure 3.3.: ICESat-2 Segments and Snow Surveys. The number of ICESat-2 Segment in DTM1 tiles, include snow-on segments in blue (left) and snow-free segments in red (right). The validation dataset from NVE's ALS snow survey conducted in 2008.04 and 2009.04 in Hardangervidda, Southern Norway, as well as the field trip conducted in 2022.03 by this study, were used for validation purposes. NVE data source: Melvold & Skaugen (2019)

3. Study Area and Methodologies

- h_{te_std} : standard deviation of terrain height within the segment.
- $h_{te_uncertainty}$: the sum of uncertainties such as geolocation, pointing angle, timing, etc, are used to describe the quality, status of elevation of segment.
- *snowcover mask*: dividing the dataset into snow-free land surface and snow surface using daily snow-cover masks from NOAA daily snow cover products (A. Neuenschwander et al., 2022).
- *segment_watermask*: water mask from *Global Raster Water Mask* at 250-meter spatial resolution (Carroll et al., 2009).
- *segment_landcover*: land-cover labels from *Copernicus Global Land Cover* at 100-meter spatial resolution (Buchhorn et al., 2020).
- h_{mean_canopy} : for the segments where the proportion of photons classified as canopy photons exceeds 5%, this value represents the mean canopy height of the segments.(A. Neuenschwander et al., 2022).
- $h_{te_uncertainty}$. Uncertainty of mean terrain height incorporates all systematics uncertainties, such as timing, orbits, geolocation, slope.(A. Neuenschwander et al., 2022).
- *canopy_openness*: quantifying the standard deviation of relative heights for all canopy photons within a segment. This parameter potentially indicates canopy openness, as greater standard deviation indicates greater penetration of laser energy into the canopy.
- *segment_cover*: woody vegetation fractional cover derived from *Copernicus Global Land Cover* at 100-meter spatial resolution (Buchhorn et al., 2020).
- *subset_te_flag*: each segment comprises five geosegments, and *subset_te_flag* denotes the number of geosegments for each. With a value of 5, all geosegments are available and continuous.
- *brightness_flag*: indicating the snow cover and other high reflective object. This is a good alternative when the *snowcover* mask is not accurate enough (A. Neuenschwander et al., 2022).

Prior to differencing the DEMs, this study excludes moving terrain based on the *segment_watermask* and *segment_landcover* masks. Moving terrain refers to unstable topography during the period of DEM production and the ICESat-2 visit, including inland water bodies, permanent snow and ice, and open sea. However, potential biases may arise from surface mass movement such as landslides, significant erosion or deposition. These biases are more likely to occur in outdated DEMs and can considerably impact regional accuracy.

The ICESat-2 segments are divided into snow-free or snow-on categories using the *snowcover* mask. Suspicious snow-on measurements are further excluded from the snow-free category using the *brightness_flag*. To eliminate ICESat-2 data of low reliability, segments with few $n_{te_photons}$ ($N < 10$) or $h_{te_best_fit}$ labeled by NaN (no data) are excluded. As part of the quality control procedures, this study retained only the segments in which the *subset_te_flag* equals 5.

3.3.2. DEMs

Four DEMs were used as reference ground for the entire mainland of Norway: DTM1, DTM10¹, Copernicus GLO-30 (GLO30)², and FABDEM (Forest And Buildings removed Copernicus DEM, FAB)³. GLO-30 (European Space Agency & Sinergise, 2021) is a 30-meter-resolution Digital Surface Model acquired from December 2010 to January 2015 during the TanDEM-X mission using Synthetic Aperture Radar interferometry. It covers the entire Earth's landmass and fills all voids with multiple sources. The products are provided in Geographic Coordinates with the horizontal reference datum of World Geodetic System 1984 (WGS84) and vertical reference datum of Earth Gravitational Model 2008 (EGM2008).

FAB represents the DTM variant of GLO30 (Hawker et al., 2022), which effectively eliminates buildings and trees by utilizing the random forest algorithm. As reported by Hawker et al. (2022), FAB provides improved accuracy, making it a suitable comparison with GLO30.

DTM1 was developed by Kartverket's National Detailed Height Model project and launched in 2016. It has a resolution of 1x1 meter and uses the ETRS89 - UTM33N coordinate system and NN2000 as the vertical reference datum. DTM10, with a spatial resolution of 10 meters, is derived from DTM1.

DTM1 is divided into approximately 15 x 15 km tiles, while DTM10 and COP30 are divided into approximately 50 x 50 km tiles. FABDEM is divided into 1° x 1° tiles. All DEMs were transformed to the WGS 84 - UTM33N projection and vertical datum EGM2008 by PROJ⁴

3.3.3. Terrain and vegetation features

The bias of DEMs varies depending on the condition of the surface, such as slope, vegetation, as previously discussed in Section 2.2.2. The spatial variability of snow is also related to both topographic factors and vegetation, as discussed in Section 2.1.1. In light of this, the following parameters or features are prepared for regression:

- (1) *The elevation difference (dh) of DEM to ICESat-2 elevation is defined as:*

$$dh = H_{\text{ICESat-2}} - H_{\text{dem}} \quad (3.1)$$

When the ICESat-2 segments ($H_{\text{ICESat-2}}$) is snow-free, the elevation difference is elevation residuals, when it is snow-on, the elevation difference is snow depth. Due to the difference

¹The DTM can be accessed from ©Kartverket <https://hoydedata.no/LaserInnsyn2/>

²The Copernicus DEM can be accessed from <https://spacedata.copernicus.eu/fr/collections/copernicus-digital-elevation-model>

³The FABDEM can be accessed from <https://data.bris.ac.uk/data/dataset/25wfy0f9ukoge2gs7a5mqpq2j7>

⁴For PROJ, visit <https://proj.org/index.html> and for datum grids, visit <https://github.com/OSGeo/PROJ-data/>

3. Study Area and Methodologies

in structure between the gridded DEM and the ICESat-2 segment midpoints, the DEM is interpolated linearly to the point by `scipy.interpolate`⁵.

- (2) The *slope* of the DEM describes the gradient of each pixel in relation to its neighbours.
- (3) The *aspect* describes the orientation of the strongest slope.
- (4) The *curvature* refers to the change in slope. Curvature can be further categorized into planform curvature (*planc*), perpendicular to the direction of the slope, and profile curvature (*prof*), parallel to the direction of the slope. A negative profile curvature indicates upward convexity, while a positive value indicates upward concavity, where water flow accelerates flowing down. Negative plan curvature means sideward concavity with water flow convergence, while positive plan curvature is sideward convexity.

Specifically, xDEM⁶ (*Xdem*, 2021) calculates terrain attributes based on DTM10 in a 3x3 pixel window and assigns the result to the central pixel. This process uses the Zevenbergen & Thorne algorithm (Zevenbergen & Thorne, 1987).

- (5) The *Topographic Position Index (TPI)* is a metric that determines slope position based on Weiss's algorithm (Weiss, 2001). It measures the difference between the elevation of a central pixel and the average of its neighboring pixels (3 x 3 pixels).

A *TPI* value of zero or near-zero indicates a flat or nearly continuous slope. Positive *TPI* values indicate that the central pixel is much higher than the surrounding areas, forming a ridge or hill. Negative *TPI* values indicate that the central pixel is much lower than its surroundings, forming a valley or gully. The *TPI* has been shown to be effective in predicting snow distribution in alpine environments (Re-vuelto et al., 2014). It is worth to note that *TPI* is a scale-dependent metric (Weiss, 2001). To represent landforms at a coarse scale, this study introduces two additional indices: *TPI_9* (calculated in 9 x 9 pixel windows) and *TPI_27* (using a window size of 27 x 27 pixels). See Figure 3.4 for typical mountainous terrain of Norway.

The vegetation parameters are derived from ICESat-2 ATL08 product (Section 3.3.1), which includes (6) *h_mean_canopy* and (7) *canopy_openness*. These parameters are obtained by studying the photon distribution within a 100m segment of the canopy. Moreover, (8) *segment_cover* signifies the proportion of woody vegetation that is sourced from a third-party database (Buchhorn et al., 2020). However, ICESat-2 cannot accurately measure canopy parameters during snowy conditions, necessitating the use of supplementary corrections or third-party vegetation databases. For more details, refer to Section 5.2.3, while Appendix A.3 provides a recommended solution.

⁵The SciPy manual <https://docs.scipy.org/doc/scipy/reference/interpolate.html>

⁶The xDEM library <https://xdem.readthedocs.io/>

3.3. The Data Setting

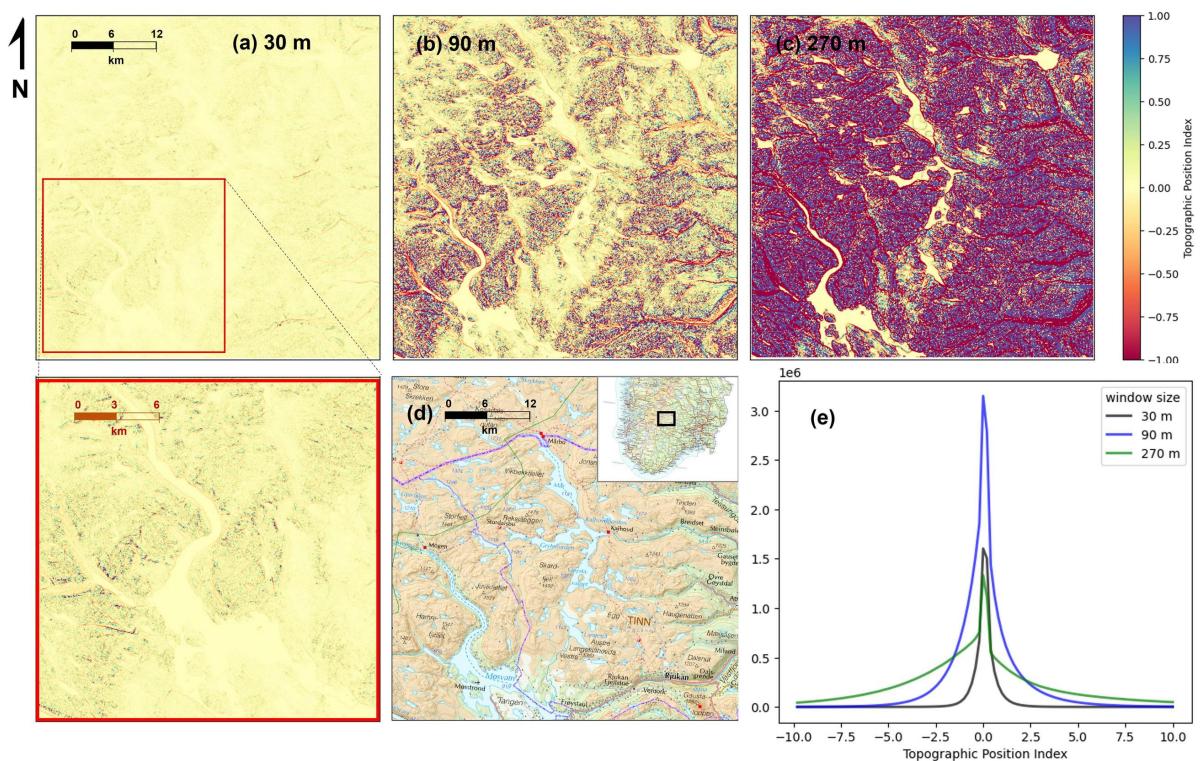


Figure 3.4.: The Topography Index: Scale-dependent Metric. Window size in 30 m (a); window size in 90 m (b) ; window size in 270 m (c), the map of east Hardangervidda, southern Norway (d); the distribution of TPI at different scales (e).

3. Study Area and Methodologies

3.3.4. ERA5 Land: snow and wind features

ERA5 Land (version 5) (Muñoz Sabater, 2021a) provides climate reanalysis data gridded on a regular grid of $0.1^\circ \times 0.1^\circ$ (~ 9 km) hourly. ERA5 Land's snow depth data represents an instantaneous average of the snow thickness on the ground, excluding snow on the canopy. The snow depth from ERA5-Land couples with ICESat-2 derived snow depth at daily resolution. Additionally, ERA5 Land monthly dataset also contributes snow melt, snowfall, and wind features accumulated for the water year period (September to September). These calculations aid in characterizing the grid in the degree of snow accumulation and snow melting, which in turn helps with statistical downscaling:

- (1) Snowfall accumulation (s_{fall_acc}) is a monthly proxy for snow precipitation.
- (2) Snowmelt accumulation (s_{melt_acc}) is a monthly proxy for melting energy.
- (3) *Wind-aspect factor* (W_f) (Bennett et al., 2022; DVORNIKOV, 2015) is a proxy for snow accumulation on the lee side of topographic obstacles. It is positive value to the leeward but negative value to windward of the topographical features. Essentially, the wind-aspect relationship can be described by a cosine function ranging from -1 to 1 for any prevailing direction (Figure 3.5):

$$W_f = -\cos(aspect - dir_{wind}) \quad (3.2)$$

Where dir_{wind} is defined by the direction of the wind origin, i.e, northerly wind blows from the north to the south and referred to as 0° .

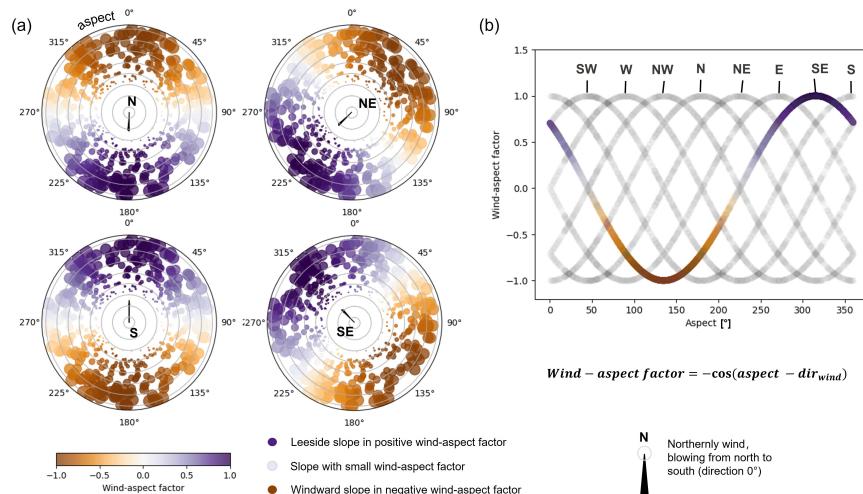


Figure 3.5.: Quantifying the Relationship between Wind and Aspect in Snow Redistribution. The prevailing wind, e.g. from N, NE, S, SW result in negative value on windward side and positive values on leeside (a). And the relationship are described by a cosine function. Eight cardinal directions are plot but the function works on any wind direction (b).

3.4. Gradient Descent Co-registration

As the scouring and loading effect can not cancel each other, so this study further split W_f into the leeside factor ($W_{uf_{positive}}$) and the windward factor ($W_{uf_{negative}}$), and multiplied with the wind speed in power of 3 (Figure 3.5) to get accumulative effect of wind redistribution over the winter period.

$$W_{uf_{positive}} = \sum W_{f_{positive}} u_{wind}^3 \quad (3.3)$$

$$W_{uf_{negative}} = \sum W_{f_{negative}} u_{wind}^3 \quad (3.4)$$

Where u_{wind} is the monthly average wind speed at 10 meters from ERA5. The accumulation starts in September when the value from 0 and continues until the next September when the value reaches its maximum. The value does not accumulate when the monthly average snow depth falls below 0.1 meters during the annual cycle.

3.4. Gradient Descent Co-registration

While the commonly used NuthKaab co-registration (Nuth & Kääb, 2011) efficiently handles georeferencing errors on the pixel and sub-pixel levels, it is limited in the computationally-intensive tasks, which is a typical issue using DTM1 at a national scale. This study presents a robust gradient descent-based co-registration (GDC) approach that treats DEM co-registration as a bound-restricted minimizing problem with random noise:

$$\min_{x \subseteq \theta} f(x) = \min_{x \subseteq \theta} \mathbb{E}[F(x, \xi)] \quad (3.5)$$

where x is the parameter that minimizes $f(x)$ within the bounds θ . And ξ represents the noise.

The goal of the minimization problem is to find the parameter of x that minimizes the expected value of $F(x, \xi)$, which represents the optimal decision under uncertainty. In the context of co-registration. The function to be minimised is the NMAD of elevation difference. The parameters is the shift matrix (E, N). The primary source of noise ξ is from DEM and ICESat-2 elevations. And the optimization technique employed in this method is stochastic gradient descent algorithm (Spall, 1998). At each step, GDC gains at a learning rate and ceases when it meets convergence. Without computing terrain parameters, GDC executes co-registration iteratively with lower computing consumption. If there are no optimization techniques, GDC can be regarded as an infrequently used ‘grid search algorithm’ (Section 2.2.3 discussed other algorithm). The implementation of the GDC method has been incorporated to a Python repository xDEM [Xdem (2021)]⁷.

⁷The xDEM library <https://xdem.readthedocs.io/>

3. Study Area and Methodologies

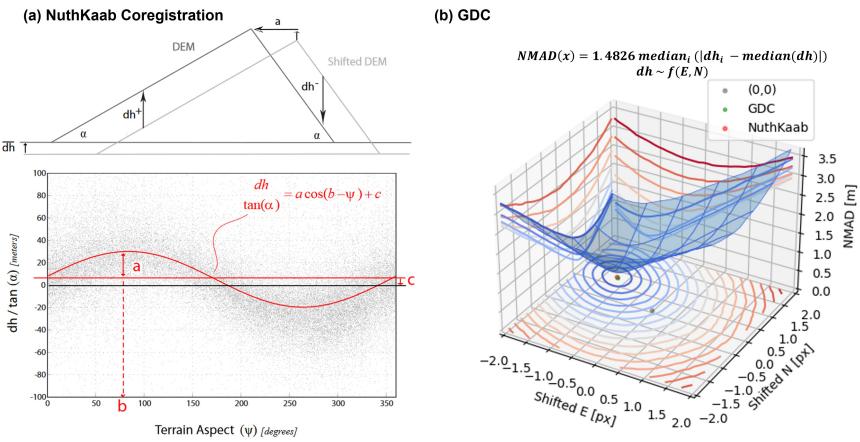


Figure 3.6.: Co-registration Techniques: NuthKaab vs Gradient Descent. The elevation difference (dh) induced by mismatch can be expressed in a 2-dimentional curve line (a). This curve suggests a statistical solution where the dispersion is related to aspect (ψ), slope (α), shifted distance (a), shifted direction (b) and vertical bias (c). By solving the equation, it gives a vector (a, b, c) to shift the DEM back over iterations Nuth & Käab (2011). On the other hand, GDC present dh in a 3-dimentional curved surface, it aims finding the shift matrix using gradient descent algorithm, resulting in the local minimal of NMAD (b)

This study aligned DEMs with ICESat-2 snow-free segments tile by tile. The co-registered DEM tiles exhibit the minimal NMAD compared to reference ICESat-2 snow-free segments. In some study cases of co-registration, the remaining median error is usually considered a vertical bias. However, this study cautions against applying vertical bias adjustment if there is no confirmation that the DEM was produced under snow-free conditions since snowpack on surfaces can lead to incorrect median error. This study did not carry out vertical bias adjustment during co-registration, and left it for the next step of bias correction. Additionally, in some tiles, there were inadequate sample sizes to achieve reliable solutions, which addressed noise suppression concerns.

3.4.1. Noise suppression

The process of registering DEMs using GDC requires special attention to the noise problem. If the DEM is intentionally shifted, the co-registration between the original and shifted DEM will result in a noise-free problem, and the NMAD of the final match will be technically zero. However, in practice, the datasets are produced using different techniques, through different processing, transformations, interpolations, at different dates. The noise exhibits on both datasets and interferes with the real difference produced by misalignment. Therefore, excluding moving terrain and suppressing noise to achieve high-quality co-registration are critical steps.

3.4. Gradient Descent Co-registration

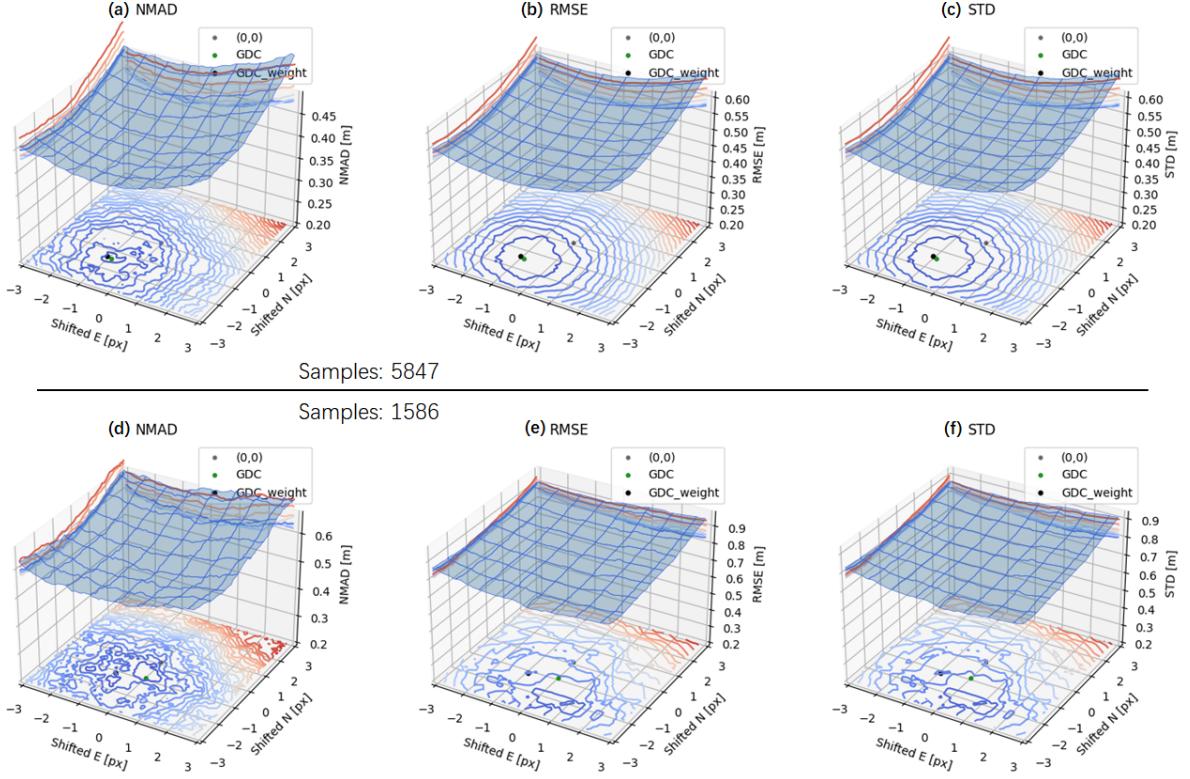


Figure 3.7.: Brute-Force Test on Gradient Descent Co-registration. The plot presents the co-registration of DTM1 with 5487 points ICESat-2. It shows that NMAD offers a steeper converging surface, compared with RMSE and STD (a,b,c), where each contour line represent 3% improvements from outside to inside, and the final results of GDC and weight-enhanced GDC has ineligible difference (in green and black dots). When the number of points decreases to 1586 points (d,e,f) the surface is ‘not’ converging anymore as more noise interference. However, the weight-enhanced GDC still has a higher chance to give the similar results as before.

3. Study Area and Methodologies

However, limited snow-free measurements are available in high altitude areas when DTM1 tiles have relatively small coverage, making the noise severe. The Figure 3.7 is a brute force test on DTM when there are not enough sample points, which goes through all possible choices until a solution is found. This test suggests several practices to mitigate noise interference:

- Filtering by standard deviation or quantiles can also exclude outliers.
- Increasing the number of sampling points: Each ATL08 segment has five subsegments. When there are not enough points ($N < 5000$), this study used the subsegments as well.
- Assigning weights based on quality metrics of points. This study adopted weight for points according to ICESat-2 quality features, *subset_te_flag*.
- Using a robust estimator to reduce the influence of outliers. NMAD is better than RMSE or STD (Figure 3.7) as a cost function under critical conditions. It gives sharper and more robust results compared to RMSE or STD, which are sensitive to outliers.

Overall, suppressing noise during co-registration is crucial for achieving accurate results. Employing these practices can help to mitigate noise interference and ensure high-quality co-registration results.

3.5. Regression Model

3.5.1. Bias correction

In this study, XGBoost⁸, a supervised regression model (Section 2.4), was used for bias correction. The model corrected specific system errors from the DEM and ICESat-2 data after co-registration. To obtain unbiased elevation difference, a two-dimensional bias correction approach was proposed in previous studies (Section 2.3.1). However, errors can vary across multiple dimensions. Therefore, this study employs a regression model to predict the bias using various features such as terrain parameters, vegetation, and quality metrics from ICESat-2 (see Table 3.1). These features were chosen based on their potential influence on elevation measurements and were obtained from multiple data sources (see Sections 3.3.1 and 3.3.3). The features are expected to explain the bias from both datasets.

The XGBoost regression model requires adequate training data to make reliable bias correction predictions. Mainland Norway, with over 5 million snow-free segments, should provide sufficient data. XGBoost has a feature importance metric that helps identify critical features and guide feature engineering. Model tuning and cross-validation will be discussed further in Section 3.5.3.

⁸The XGBoost library can be accessed at <https://xgboost.readthedocs.io/>

Table 3.1.: Features in Bias Correction Regressor

Data source	Features	Descriptions
ICESat-2 ATL08	E	the east grid coordinates
	N	the north grid coordinates
	h_te_best_fit	the elevation in snow-off condition (m)
DTM10	slope	slope angle (degree)
	aspect	slope aspect (degree)
	curvature	curvature (m^{-1}), scalar value without direction
	planc	curvature (m^{-1}) in the direction of the steepest slope
	profcc	curvature (m^{-1}) in the direction perpendicular to the steepest slope
	TPI	topographic position index in 3 pixels (30m window size)
	TPI_9	topographic position index in 9 pixels (90m window size)
	TPI_27	topographic position index in 27 pixels (270m window size)
ICESat-2 ATL08	h_te_std	quality metrics, see Section 3.3.1
	h_te_skew	
	h_te_uncertainty	
	subset_te_flag	
	segment_cover	vegetation parameters, see Section 3.3.1
	h_mean_canopy	
	canopy_openness	

3.5.2. Statistical downscaling

$$y \sim F(X) \quad (3.6)$$

Statistical downscaling is to bridge the coarse resolution products to fine resolution by establishing an empirical link between a set of predictors X and predictands y by a statistical model (F) (eq. 3.6). The stationary assumption behind is that the predictor have to be realistically and bias free in present climate, and Informative predictors need to be selected that explain a large fraction of local variability on all timescales of interest (Maraun & Widmann, 2018).

This study uses ICESat-2 derived snow depth minus mean snow depth from ERA5 Land at the same location and day as the predictand y , while predictors X are chosen based on their influence on snow

3. Study Area and Methodologies

melting and redistribution processes. The downscaling process comprises three steps: detrending, regression, and re-trending.

(a) Detrending

$$y = \text{Snowdepth}_{ICESat_2} - \text{Snowdepth}_{ERA5} \quad (3.7)$$

The subtracted snow depth is the mean value interpolated from ERA5 Land. The detrended snow depth, denoted by y , is the deviation from mean snow depth. The variance of y represents the local variability of the snow depth. The mean value of y indicates systematic bias between the two datasets.

Before the training, quality control measures are implemented to ensure reliable results. Specifically, measurements $\text{Snowdepth}_{ICESat_2}$ with negative values are excluded as they are considered pure noise. Additionally, mean snow depth Snowdepth_{ERA5} less than 0.1 m are excluded as these values indicate that the snow depth is close to zero and therefore meaningless for downscaling purposes.

(b) Regression

As a consequence of snow redistribution, exposed windward slopes are often has less snow accumulation. In the opposite, wind-sheltered areas, surface depressions, receive large amounts of snow. So multiple terrain parameters, slope, aspect, curvature (normal curvature, profile curvature and horizontal curvature), TPI (in 30 m, 90 m, 270 m scales) are input as predictors of snow distribution at local scales. Vegetation is also an important factor, being described by canopy openness, vegetation height from ICESat-2 (Table 3.2).

The wind-aspect factors used in this study are the leeside factor $W_{uf_{positive}}$ and the windward factor $W_{uf_{negative}}$, which were accumulated during the snow season (Section 3.3.4). Snow melt accumulation from ERA5 Land monthly was used to determine the scale of snow melt for each grid, while snowfall as well (Section 3.3.4). It should be noted that these variables do not work on a pixel level (10 m), but rather characterize wind redistribution and snow dynamic patterns at the grid level (~9 km). Furthermore, the availability of some features, such as canopy information from ICESat-2, will decide the coverage of output snow depth.

Table 3.2.: Features in Downscaling Regressor

Data source	Features	Resolution	Descriptions
ICESat-2 ATL08	E	-	the east grid coordinates
	N	-	the north grid coordinates
	h_te_best_fit	-	the surface elevation of middle point of segment (m)

Data source	Features	Resolution	Descriptions
DTM10	slope	10 m	slope angle (degree)
	aspect	10 m	slope aspect (degree)
	curvature	10 m	
	planc	10 m	
	profcc	10 m	
	TPI	10 m	
	TPI_9	10 m	
	TPI_27	10 m	
ICESat-2 ATL08	segment_cover	100 m	(optional) the percentage of vegetation cover.
	h_mean_canopy	100 m	(optional) the mean height of canopy.
	canopy_openness	100 m	(optional) the openness of canopy.
ERA5 Land	sfall_acc	9 km	(optional) snow fall accumulation
	smelt_acc	9 km	(optional) snow melt accumulation
	wf_negative	9 km	wind-aspect factor
	wf_positive	9 km	wind-aspect factor
-	month	-	

(c) Re-trending

The study's final step involves predicting local-scale snow depth by adding the predicted y to the mean snow depth obtained from ERA5 Land. The variance of the predicted y represents the snow depth distribution at local scales, while the mean value is normally offset due to inconsistency between ICESat-2 derived snow depth and ERA5 Land's snow depth.

The coverage and resolution of the predicted y depend on input predictive variables X . This study produced two types of downscaled snow depth products: point-based and grid-based. The point-based product has the same coverage as the ATL08 product, with almost 20 million points for mainland Norway. On the other hand, the grid-based snow depth product was generated based on DEM coverage; hence no vegetation parameter from ICESat-2 was included as input features. The native resolution of the predicted y is 10 m. Since some variables like wind are obtained from ERA5 Land (~9 km) and cannot explain changes at 10 meters, this study only regards output resolution as hill-slope scales.

(d) Offset adjustment (optional)

There are three methods that can be applied to address the significant difference between downscaled snow depth and true snow depth. The first method is the linear adjustment, which assumes a

3. Study Area and Methodologies

linear relationship in the discrepancy. The second method is the spatial aggregation method, which assumes symmetrical discrepancy. The third method is quantile mapping, which requires knowledge of the true distribution of y . Code implementations for both approaches can be found in Appendix A.5. All adjustments only change the scale of values and do not affect the Spearman correlation coefficient and R-Squared value of the results. By default, this study does not apply any offset adjustment except for Figure 4.10 and Appendix A.3 where linear adjustment has been applied.

3.5.3. Hyperparameters and cross-validation

XGBoost has many hyperparameters that can significantly impact its performance, including the objective function, number of trees, learning rate, and regularization terms (Section 2.4.1). It is important to tune these parameters carefully and use a validation set to avoid overfitting. The optimal parameters for the regression task in this study have been identified by grid search as follows (Listing 3.1). After identifying the optimal hyperparameters, this study takes measures to prevent overfitting by dividing the dataset into training (80%) and testing subsets (20%), and applying cross-validation as demonstrated in Figure 3.8. In this example, as the number of boosting rounds increases, the training and testing MAE decrease and eventually converge to a stable value (0.793 and 0.756 respectively). It indicates that the model is not overfitting, and $n_estimators$ is reasonable at a value of 250.

Listing 3.1: The Optimal Hyperparameters by Grid Search.

```
1  params = {  
2      'objective': 'reg:absoluteerror',  
3      'max_depth': 10,  
4      'learning_rate': 0.1,  
5      'n_estimators': 250,  
6      'min_child_weight': 1,  
7      'subsample': 0.7,  
8      'colsample_bytree': 1,  
9      'gamma': 0.1,  
10     }
```

3.5.4. Feature engineering

Feature engineering is the process of selecting, extracting, and transforming useful features from raw data to new input variables that can improve the performance of machine learning models. Specifically, the regression model of bias correction is influenced by selected features such as vegetation height and canopy openness.

The figure depicted in 3.9 reveals that the height of vegetation differs significantly between snow-free and snow-covered conditions, with the mean canopy height being substantially higher under snow-free conditions. This is because when the ground surface is covered with snow, the elevation measurement from ICESat-2 will also include the snow depth, thus resulting in a shorter canopy height than

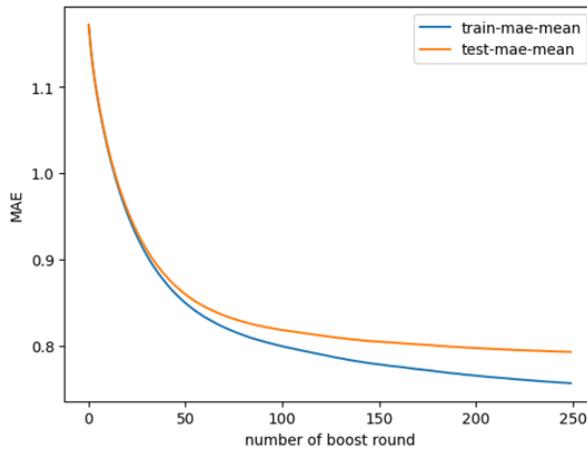


Figure 3.8.: Preventing Overfitting with Cross-validation in XGBoost. MAE is the objective function, and the number of boost rounds is the ‘n_estimators’ parameters. The MAE of the training datasets (blue) and testing datasets (orange) decreases and eventually converges to a stable value (0.793 and 0.756, respectively), which indicates that the model is not overfitting. If the training MAE has a significantly lower trend than the test MAE, then increasing the number of boost rounds will result in more severe overfitting.

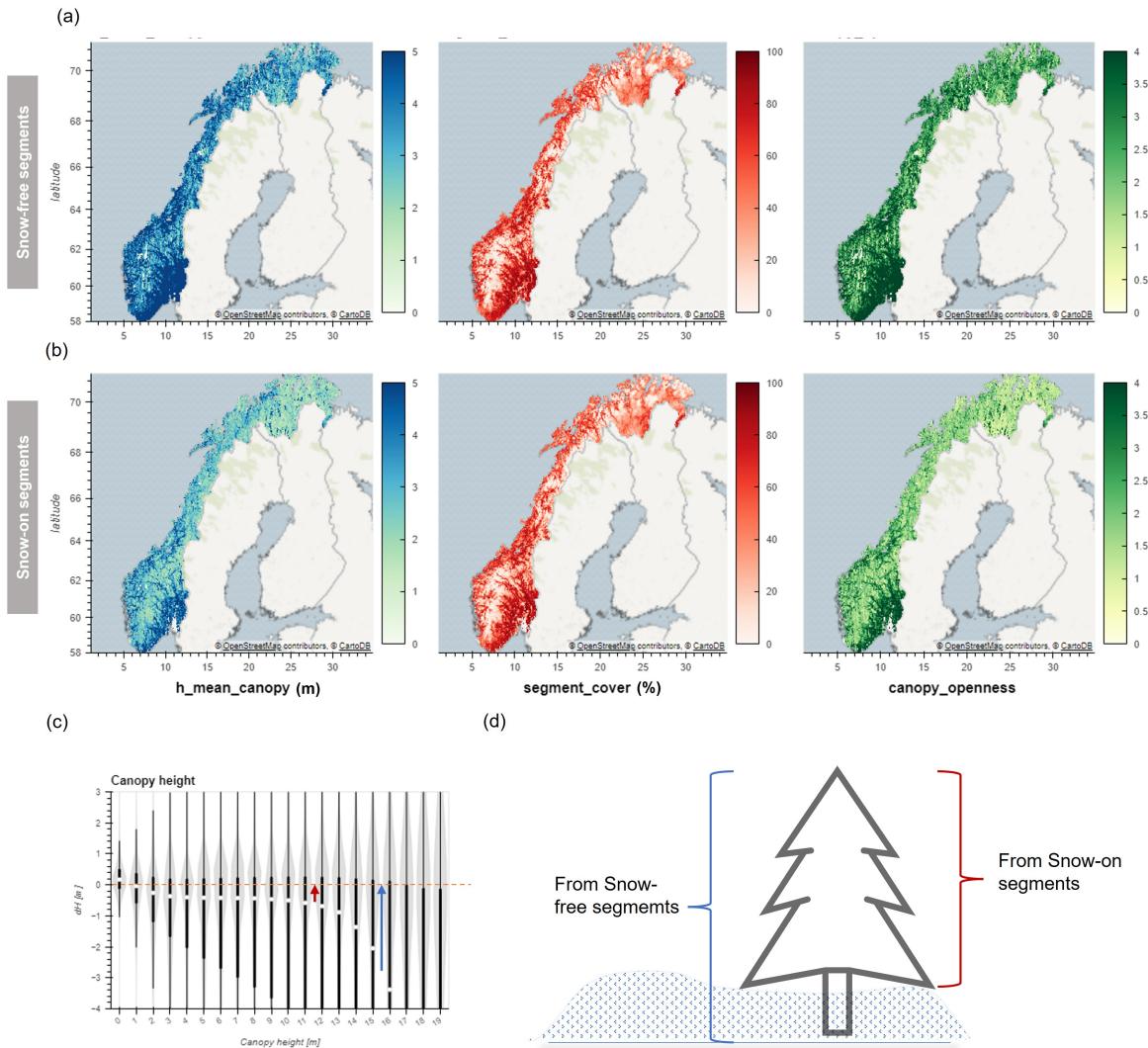
the actual vegetation canopy. Moreover, canopy openness is greater under snow-free conditions, implying that fewer photons are classified as canopy photons.

For this reason, it is crucial to correct for differences in canopy height between snow-on and snow-free segments. However, this task poses a challenge, since the amount and distribution of snow can vary significantly within a segment. In this study, we applied additional regression models to both *h_mean_canopy* and *canopy_openness* for snow-on segments before incorporating them into our model (see Appendix A.3). This allowed us to obtain a more accurate estimation of canopy height while accounting for variations caused by differences in snow cover.

On the other hand, after careful comparison, it is not recommended to use features such as the number of photons (*n_te_photons*) and signal-to-noise (*snr*) in the regression model. These features are efficiently explaining the accuracy and precision of ICESat-2, for example, ICESat-2 has more photons available and yield lower uncertainties over snow-on surface. However, their behavior is highly dependent on the presence or absence of snow cover, but does not represent the same bias pattern for the snow-off condition. As such, including them in a regression model would not yield reliable predictions.

In summary, to achieve an accurate regression model, one must consider and understand the potential biases present within both ICESat-2 and DEM. The predictors may vary significantly depending on external factors like snow cover. As a result, feature engineering is key when building an effective

3. Study Area and Methodologies



3.6. Evaluation of Snow Depth

model.

3.6. Evaluation of Snow Depth

3.6.1. ICESat-2 derived snow depth vs lidar derived snow depth

If a dataset has three times better precision than another, in error propagation practice, it can explain 90% of the variance in the output of elevation differencing. This high precision dataset can serve as a true surface to assess the biased dataset (Höhle & Höhle, 2009; Maune, 2007), and validation surfaces are typically obtained from ground-based surveys or airborne lidar.

A near real-time snow survey was conducted using DJI Zenmuse L1 in Hardangervidda on March 11, 2022, one week after ICESat-2 flew over the same location (Figure 3.10, 3.3). The survey covered only a strip of snow-on surface measuring 3,700 meters in length and 700 meters in width with relatively flat terrain, sparse vegetation. And DTM1 is used as the reference ground.

The purpose of the validation data is to investigate various factors that impact the accuracy and precision of ICESat-2 derived snow depth. These include figuring out the impact of the resolution and quality of the reference DEM, correctness of bias correction, and fitness of ICESat-2 derived snow depth.

Accuracy is quantified using mean, median, while variance is quantified using NMAD. The fitness of the ICESat-2 derived snow depth to validation data is determined by Spearman correlation coefficient and R-square value. It is important to note that since validation datasets only cover a small area, they only account for performance under specific conditions. Further analysis is necessary to determine the impact of slopes, vegetation, and other terrain parameters.



Figure 3.10.: Snow Survey via Drone Lidar in Hardangervidda, Southern Norway. The field trip took place on March 11, 2022, a week later after ICESat-2 flying over the same location. Photo: Eivind Torgersen, UiO

3. Study Area and Methodologies

3.6.2. ICESat-2 derived snow depth vs ERA5 Land snow depth

The purpose of the comparison is to determine if there are any discrepancies in snow depth values between the two datasets, specifically whether ERA5 Land overestimates snow depth. Such a phenomenon has been noted in previous studies (Orsolini et al., 2019). Additionally, this comparison helps to identify differences in using different reference DEMs in snow retrieval. To achieve this, the two datasets were spatially and temporally coupled at daily resolution to obtain the snow depth value from ERA5 Land. The snow depth was aggregated to a mean value and visually compared to assess bias at a national scale. A scatter plot with the Spearman correlation coefficient will be utilized. Additionally, a histogram will be employed to depict the distribution of the snow depth differences.

The evaluation of snow depth from different sources, at different scales, and with varying representations is a challenging task. ERA5 Land provides coverage for glaciers, lakes, and fjords in a 9 km resolution while ICESat-2-derived snow depth excludes all moving terrain and water areas where reference DEM is unavailable. These disparities in data source must be taken into consideration when understanding the difference (further discussion in Section 5.3).

3.6.3. Downscaled snow depth vs lidar derived snow depth

Airborne lidar data was collected for a 240 km² area on Hardangervidda mountain plateau using a Leica ALS50-II instrument. The data was collected during two different time periods in 3-21 April 2008 and 21-24 April 2009 (Figure 3.3). Each time, six flight-lines of lidar data were collected, with each flight-line being 80 km long, following a west-east orientation and having a scanning width of approximately 1000 m. The snow depth is calculated on 2 m resolution bare ground DEM, where some small perennial snow patches still exist (Melvold & Skaugen, 2013).

The evaluation of downscale snow depth focuses on determining whether it can accurately represent the distribution of snow depth on a hill-slope scale. The assessment of similarity is quantified by the R-Squared value and Spearman correlation coefficient. If successful, it is essential to identify which features dominate the subgrid distribution the most. Conversely, if it fails to do so, it is necessary to identify which conditions cannot be reproduced. The downscaling technique can generate snow depth values at any specific location and time, allowing this study to use snow depth surveys from 2008 for validation. However, this process also introduces uncertainties due to the regression training dataset and initial conditions (the ERA5 Land dataset). The training data for regression were based on observations from the years 2018-2022. Furthermore, predictor features such as terrain, vegetation, and climate conditions may have undergone changes since 2008.

4. Results

4.1. Co-registration

The GDC corrects the horizontal misalignment of DEMs and minimizes the NMAD of samples. Table 4.1 presents the results of DEM co-registration. The NMAD shows a 2% improvement for DTM1, while DTM10, GLO30, and FAB exhibit an improvement of 22%, 21%, and 35%, respectively. Subsequently, after co-registration, the percentage of points with a deviation less than 0.5 m from ICESat-2 snow-free segments is calculated to be 48.9%, 49.13%, 23.7%, and 27.2% for DTM1, DTM10, GLO30, and FAB, respectively.

To diminish the effect of noise, this study employed different weightings on elevation differences based on the quality metric *subset_te_flag* from ICESat-2 data (see Section 3.4.1 for further discussion). Nevertheless, final statistics, such as NMAD were calculated evenly across all data points. It should be noted that the ‘minimal NMAD of samples’ may not always guarantee the best co-registration if there is severe noise interference in samples. Instead, the performance of co-registration may have been underestimated due to noisy sample points.

Table 4.1.: The Improvement of DEM by Co-registration

Metrics	Co-registration	DTM10	DTM1	GLO30	FAB
N	-	5,443,945	5,443,945	5,443,945	5,443,945
dh <0.5m	before	41.84 %	48.73 %	18.11 %	18.99 %
dh <0.5m	after	48.92 %	49.13 %	23.72 %	27.30 %
dh <1.0m	before	59.79 %	65.44 %	31.08 %	34.00 %
dh <1.0m	after	65.63 %	65.68 %	37.97 %	45.78 %
NMAD (m)	before	0.992 m	0.774 m	2.533 m	2.381 m
NMAD (m)	after	0.772 m	0.760 m	2.058 m	1.611 m

4. Results

Table 4.2.: The Improvement of DEM by Co-registration (*subset_te_flag* equals 5)

Metrics	Co-registration	DTM10	DTM1	GLO30	FAB
N	-	3,609,933	3,609,933	3,609,933	3,609,933
dh <0.5m	before	51.63 %	60.14 %	23.12 %	22.88 %
dh <0.5m	after	60.69 %	60.71 %	31.61 %	33.93 %
dh <1.0m	before	71.63 %	78.42 %	38.78 %	40.13 %
dh <1.0m	after	78.86 %	78.75 %	49.15 %	55.37 %
NMAD	before	0.682 m	0.511 m	2.007 m	1.961 m
NMAD	after	0.507 m	0.498 m	1.450 m	1.228 m

Figure 4.1 shows that all four DEMs indicate a misalignment towards the west and north directions. Most DTM10 tiles experienced an improvement through a ‘half’ pixel (5 to 10 m) shift, as did GLO30 and FAB. However, when DEM resolution comes to 1 m, the misalignment of one to three pixels (1 to 3 m) may not be significant relative to the native accuracy of the DEM (Figure 4.2 a and Appendix A.2 a).

Additionally, approximately 67% of the ICESat-2 segments have *te_subset_flg* equals to 5, indicating a better match with the DEMs (Table 4.2). This demonstrates that accuracy can be achieved with NMAD values of 0.51 m and 0.50 m for DTM1 and DTM10, respectively, and 1.5 m and 1.3 m for GLO30 and FAB, respectively, which are better than previous studies (W. Chen et al., 2022; Deschamps-Berger et al., 2022; H. Li et al., 2022). However, it also indicates uncertainties in ICESat-2 ATL08 under inferior observation conditions (more discussion in Section 5.2.1).

These findings emphasize the significance of accurate co-registration for DEMs and highlight the potential for enhancing accuracy by utilizing quality metrics of ICESat-2 data. Further analysis must be exercised for residual errors.

4.2. Binning Analysis: DEM Comparison

The residual errors (*dh*) between DEM and ICESat-2 snow-free segments are investigated by statistical binning. Figure 4.2 illustrates an example of DTM10 in each processing step, including raw, co-registration, and bias correction. The residual errors are clustered into bins based on different attributes, such as elevation, aspect, slope, curvature, and TPI. In Figure 4.2 a, the aspect-dependent error, or the fingerprint of misalignment, is depicted and removed through co-registration, where the

4.2. Binning Analysis: DEM Comparison

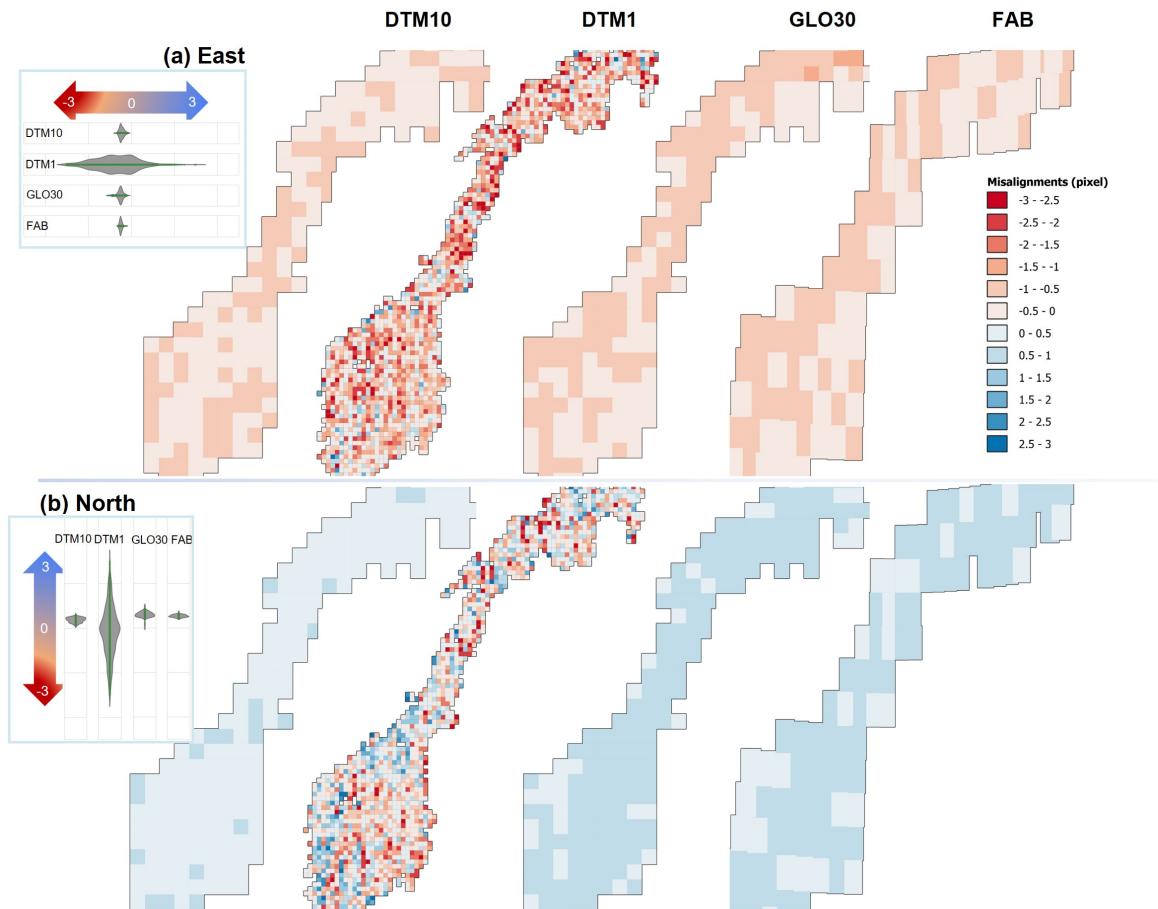


Figure 4.1.: The Misalignment of DEMs. The shift pixels in East (a) and North (b) direction ranges from -3 to 3, which can cause significant errors in the accuracy of the DEMs. Note, the shift is in pixels rather than in meters.

4. Results

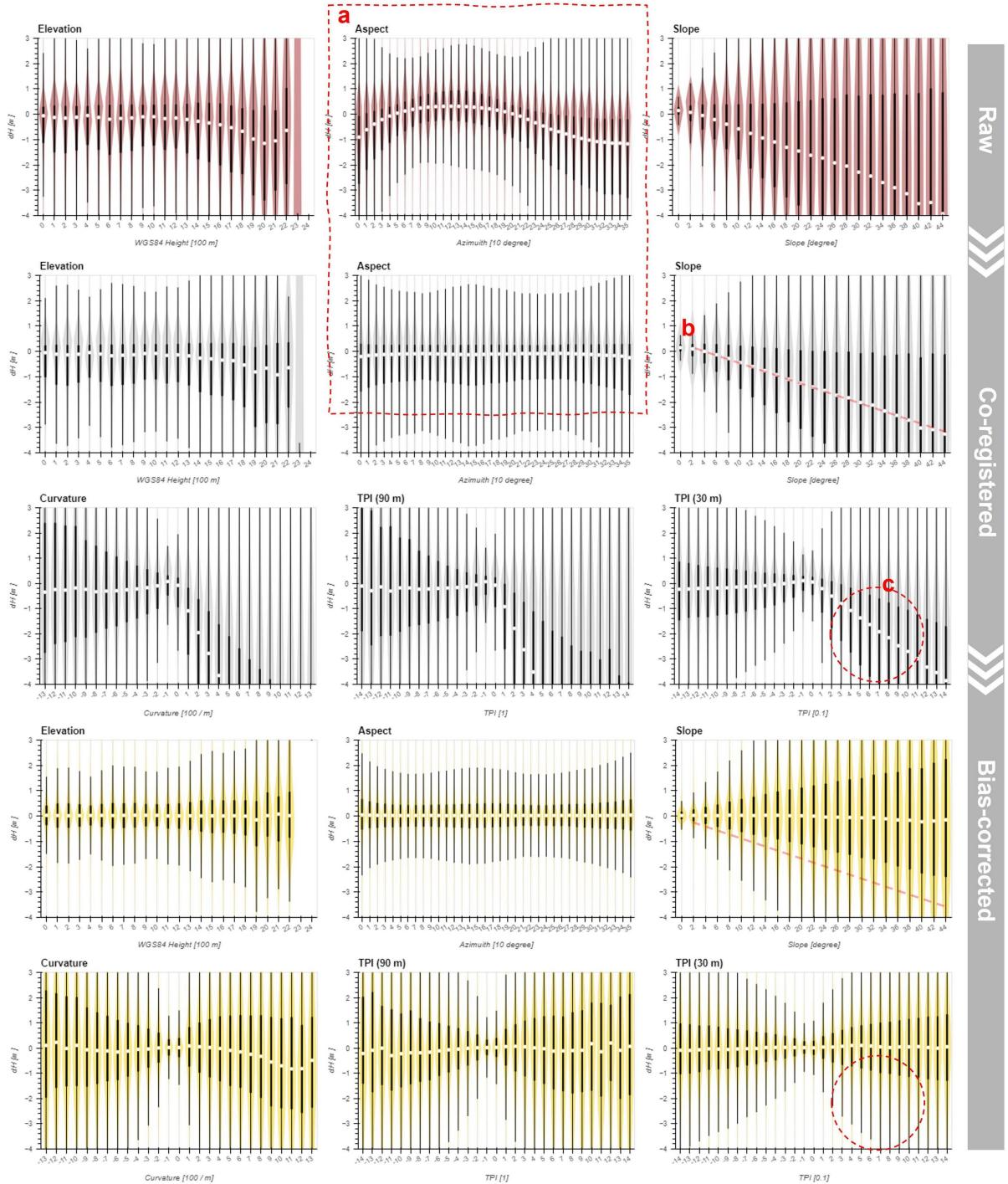


Figure 4.2.: Statistic Binning Analysis: Co-registration and Bias Correction (ICESat-2 - DTM10). The dataset was divided into categories at three stages, and for each category, a violin plot displays the distribution of data. The median of each dataset is represented by a white dot at the center of the violin. The thick bar illustrates the 25% and 75% quantiles, respectively. If the violin is skewed to one side, it indicates a deviation from normal distribution. The aspect-dependent bias is identified and eliminated by co-registration (a); The negative bias are widespread after co-registration (a, b, c).

4.2. Binning Analysis: DEM Comparison

median value (Q2) is located close to zero for all bins. However, the 25% quantile (Q1) and the 75% quartile (Q3) are still skewed negative (i.e., Q3-Q2 < Q2-Q1).

Significant negative bias has been identified for large slopes (Figure 4.2 b) as there exists a nearly linear relationship between median and slope. Some studies have applied vertical bias correction based on slope (Enderlin et al., 2022), slope and vegetation (Magruder et al., 2021), or laser incidence angles and vegetation (Tian & Shan, 2021). However, this study shows that the bias does exist over convex terrain but is not significant over concave terrain, indicated by TPI and curvature bins (Figure 4.2 c).

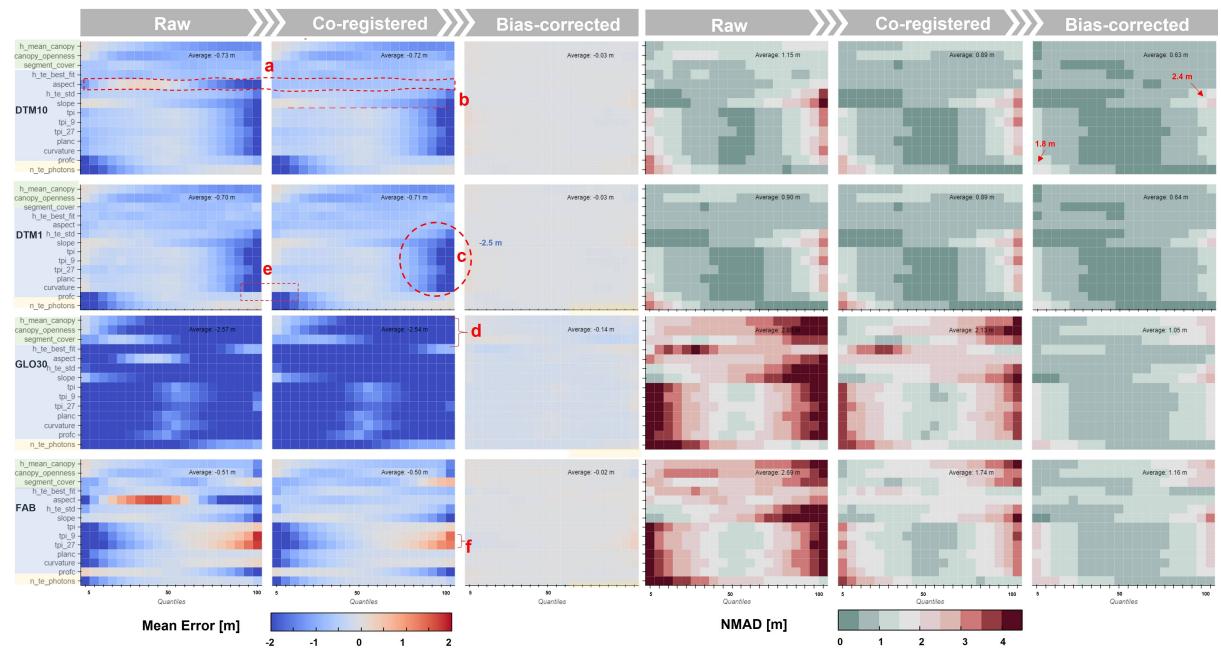


Figure 4.3.: The Mean Error and NMAD of Quantiles: Co-registration and Bias correction. Three vegetation features, ten terrain parameters and the number of photons within ICESat-2 ATL08 segments are used to bin the residual errors. The same patterns has been noticed on Figure 4.2 (a,b,c). GLO30 exhibits a widespread negative bias due to vegetation (d). The profile curvature behavior differently with curvature, plan curvature and TPI (e). So that, this study is going to discuss the negative bias further from a geomorphology perspective (Section 5.2.1). FAB has a positive bias over convex terrain, but a negative bias over concave terrain (f). From left to right, each grid represents 5% measurements. And total measurements (N) is 5,443,945.

Figure 4.3 quantifies the residual errors in terms of mean error and NMAD. Each grid represent 5% quantiles, ranging from the smallest value of each attribute to the highest. As shown in Figure 4.2, prior to co-registration DTM10 has aspect-dependent mean error (Figure 4.3 a), after co-registration, there is still a slope-dependent bias (Figure 4.3 b), and a negative bias over convex terrain (Figure 4.3 c). GLO30 exhibits a widespread negative bias due to vegetation (Figure 4.3 d). FAB has a positive bias

4. Results

over convex terrain, but a negative bias over concave terrain (Figure 4.3 f).

After the bias correction, the comparison between DTM1 and DTM10 indicates that they both achieved a high level of accuracy on low slope categories, with less than 1 m and 0.5 m NMAD in 65% and 35% of cases, respectively. In contrast, the worst results are observed in the top 5% of *slope* quantiles and the least 5% of *n_photons* quantiles, which resulted in an NMAD of 2.4 m and 1.8 m, respectively. The bias correction applied to GLO30 produced better results than for FAB, with approximately 40% of measurements showing an NMAD of less than 1 m on the low slopes category (Figure 4.3). These improvements demonstrate the effectiveness of our approach in eliminating biases.

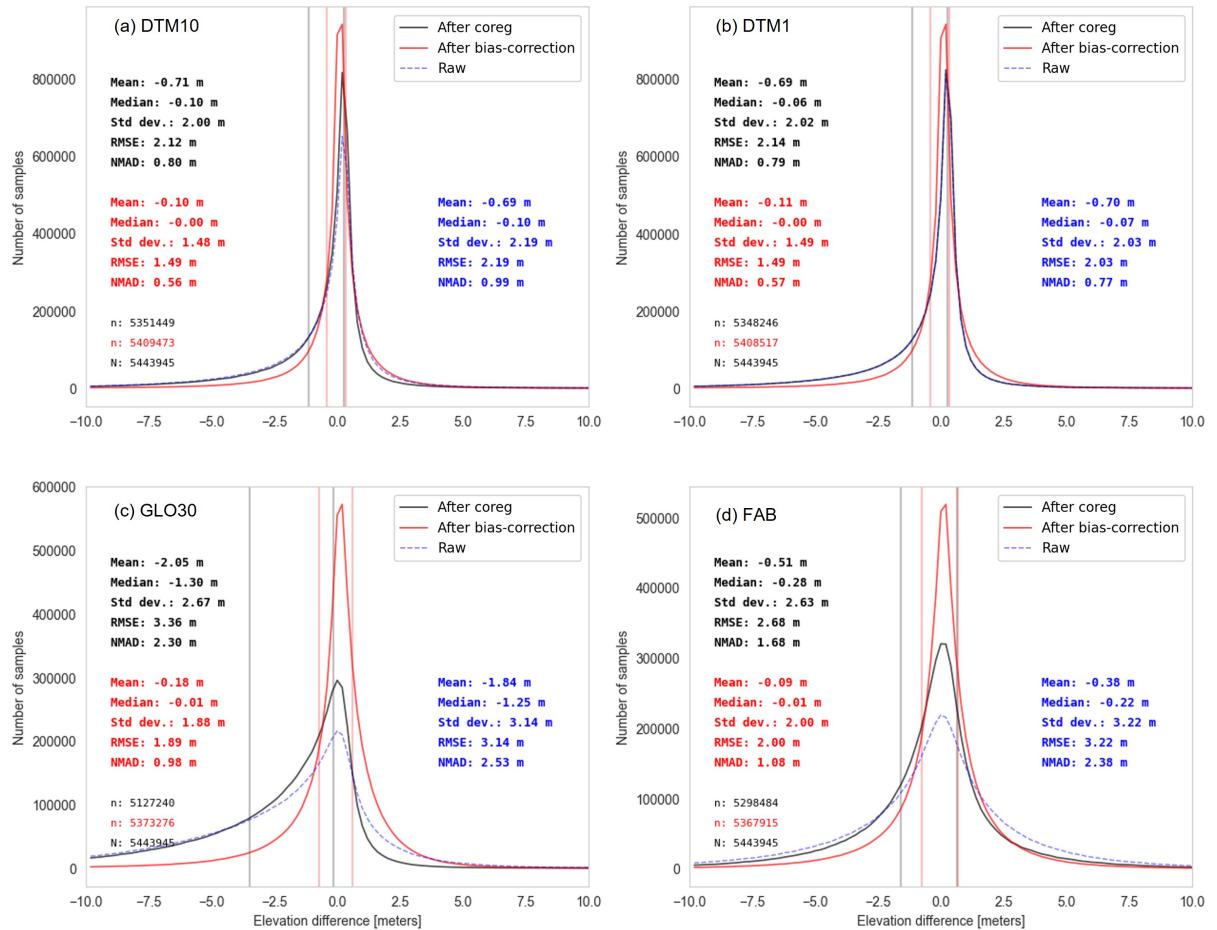


Figure 4.4.: The Overall Distribution of DEM errors. The statistical data for raw, after co-registration, and bias-correction are represented by blue, black, and red colors respectively. The 25% quantile (Q1) and the 75% quantile (Q3) are displayed from left to right along the x-axis. After bias correction, Q1 has moved closer to the median, indicating a reduction in negative skewness. Please note that a window of -10 m to 10 m was used to filter outliers in the metrics employed. And, the number of measurements is given by n (DEM) and N (total).

4.3. Feature Importance and Regression Performance of Bias Correction

The overall dh distributions (Figure 4.4) exhibit a negative skewness with long tails on the left side. Prior to bias correction, GLO30 had a median of -1.30 m, while FAB, which is a version of GLO30 with vegetation and building removed, exhibited less skewness. Although DTM1 and DTM10 had similar shapes, their 25% quantiles (Q1) suggested the presence of systematic biases. After implementing the bias correction, there was a noticeable enhancement in the overall symmetry observed across all four scenarios. This was supported by the fact that Q1 moved nearer to the median, while Q3 remained relatively stable. Ultimately, there is an overall NMAD of 0.56 m and 0.57 m for DTM10 and DTM1, respectively. GLO30 and FAB had higher NMAD of 0.98 m and 1.08 m. It is evident that bias correction led to a significant improvement in the accuracy of the results.

4.3. Feature Importance and Regression Performance of Bias Correction

The Figure 4.5 maps significant variations of predicted bias correction across different regions, depending on factors such as slope, vegetation cover, and the average of photons numbers within segments. Negligible differences are observed between DTM10 and DTM1, with bias patterns being very similar to the slope at an aggregated scale. For instance, Hardangervidda and Finnmarksvidda are relatively flat areas with minimal bias on all four DEMs. In contrast, GLO30 needs to be corrected heavily for vegetation removal in the eastern valleys, while FAB requires extra correction over vegetation area as well for unclear reasons.

The dataset's most informative features for bias correction model are presented in Figure 4.6. The feature importance scores indicate the usefulness of each predictor, with higher scores implying more important features. Among these features, *terrain slope* (slope along the ICESat-2 track) is the most important for DTM10 and DTM1 models, following with slope (calculated from DEM), the terrain roughness (*t_te_std*) or canopy height. As canopy height and openness are essential but not easily quantifiable, it is necessary to pay extra attention to errors related to canopy parameters. Conversely, metrics like *night_flag*, *pair* and *beam* of ICESat-2, and *region* (ascending or descending pass of satellite) are less important (refer to Appendix A.4 for further discussion). For GLO30, top features include *East* and *North*, suggesting that most vegetation is located in the east and south.

The performances of the bias correction on 20% test dataset presented in Table 4.3. The DEM10 model exhibits slightly better performance than the DTM1 model, with an NMAD of 0.549 m compared to 0.551 m for DTM1. This difference in performance could be attributed to the fact that the terrain features used as input were calculated based on DTM10, and some of them may not be able to explain differences at finer scales. The bias of GLO30 is more predictable than that of FAB, with NMAD of 0.901 m to 1.036 m.

4. Results

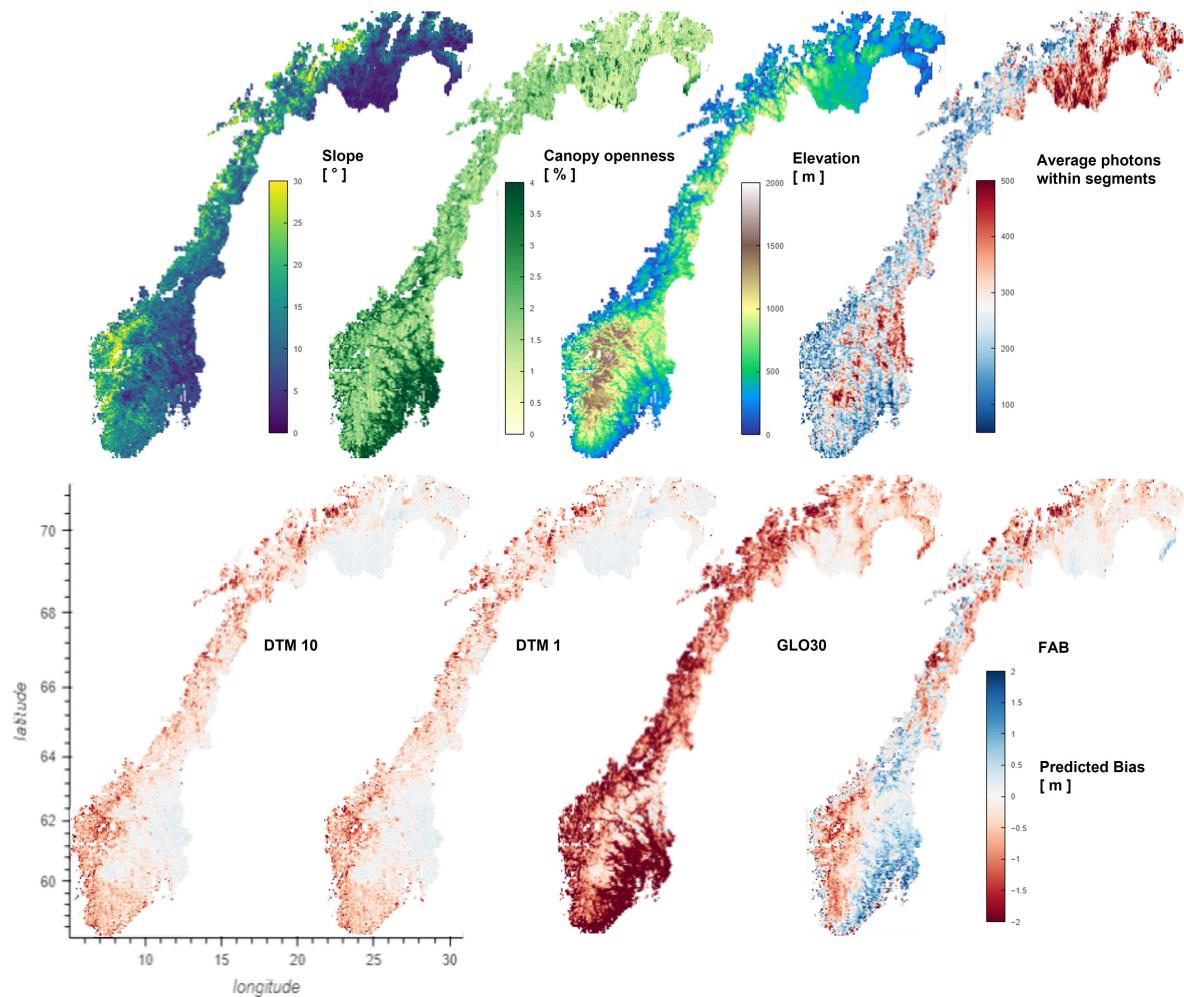


Figure 4.5.: Bias Correction Map for Snow-on Scenarios. This figure shows variations in predicted bias correction across different regions, depending on factors such as slope, vegetation cover, and the number of photons. The map are aggregated results.

4.3. Feature Importance and Regression Performance of Bias Correction

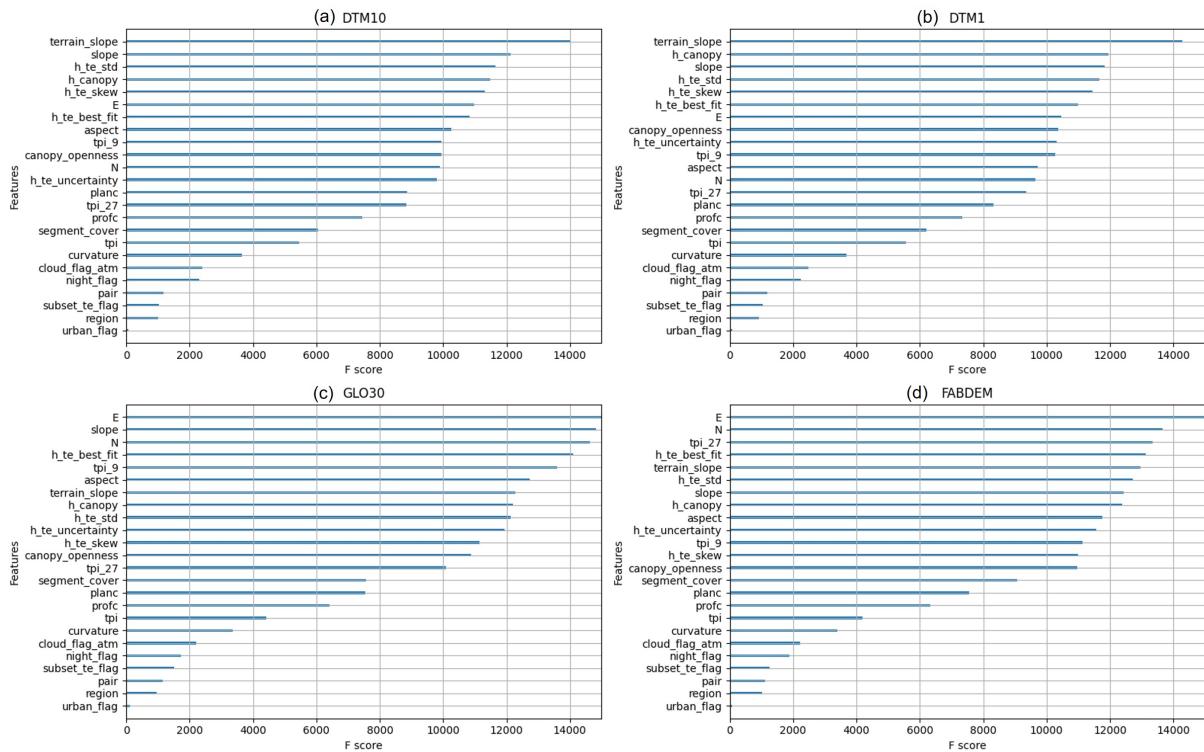


Figure 4.6.: Feature Importance of Bias Correction Regression. The F-score is a metric used to evaluate the performance of a machine learning model. Among these features, terrain slope is obtained from ICESat-2, describing the slope along the track of the segment. Further explanation regarding the name of feature can be found in Section 3.3.1.

4. Results

Table 4.3.: Cross Validation of Bias Correction Regression

Model	Tree Depth	N_estimators	NMAD (m)	RMSE (m)	MAE (m)
DTM1 - ICESat2	10	250	0.551	1.345	0.778
DTM10 - ICESat2	10	250	0.549	1.342	0.774
GLO30 - ICESat2	10	250	0.901	1.640	1.042
FAB - ICESat2	10	250	1.036	1.894	1.219

4.4. ICESat-2 Derived Snow Depth and Validation

The aggregated map of snow depth observed from ICESat-2 from 2018 to 2022 is presented in Figure 4.7. The snow depth retrieval using DTM10 and DTM1 as reference ground exhibit similar patterns. GLO30 and FAB exhibit a residual bias, causing an overestimation of snow depth in vegetated areas located in the eastern valleys. Conversely, relative flat areas with sparse canopies, such as Hardangervidda and Finnmarksvidda, display high consistency over the four DEMs.

In Hardangervidda, the drone trip on 11th March 2022 corresponded 68 valid snow-on-surface measurements to the ATL08 segments on 5th March 2022. Figure 4.8 shows a scatter plot indicating a very good agreement, where the ICESat-2 - DTM1 model has an R-Squared of 0.94 and RMSE of 0.22 m. It is important to note that the validation snow depth is calculated based on DTM1 as reference ground, so that the DTM10 (has a resolution of 10-meters) got an R-Squared value of 0.77 with an RMSE of 0.40 m. On the other hand, global DEMs such as GLO30 and FAB achieve good fitness after applying bias correction, with R-Squared values of 0.76 and 0.74 and RMSEs of 0.49 m and 0.53 m, respectively.

It should be noted that Hardangervidda is a unique area with relatively flat terrain, thick snowpacks, and sparse vegetation. Therefore, the bias correction does not significantly improve the results over DTM1 and DTM10. Additionally, the field trip also covered a limited area and cannot represent the overall performance of ICESat-2 derived snow depth.

Overall, the validation results suggest that the snow-on surface derived from ICESat-2 exhibits a high level of consistency with that obtained from lidar. However, different DEMs acting as reference can impact the accuracy performance of the snow depth product. To mitigate this effect, bias correction is necessary and useful. Aggregating the final results from GLO30 and FAB could also be a solution to increase the reliability and accuracy of the date.

4.4. ICESat-2 Derived Snow Depth and Validation

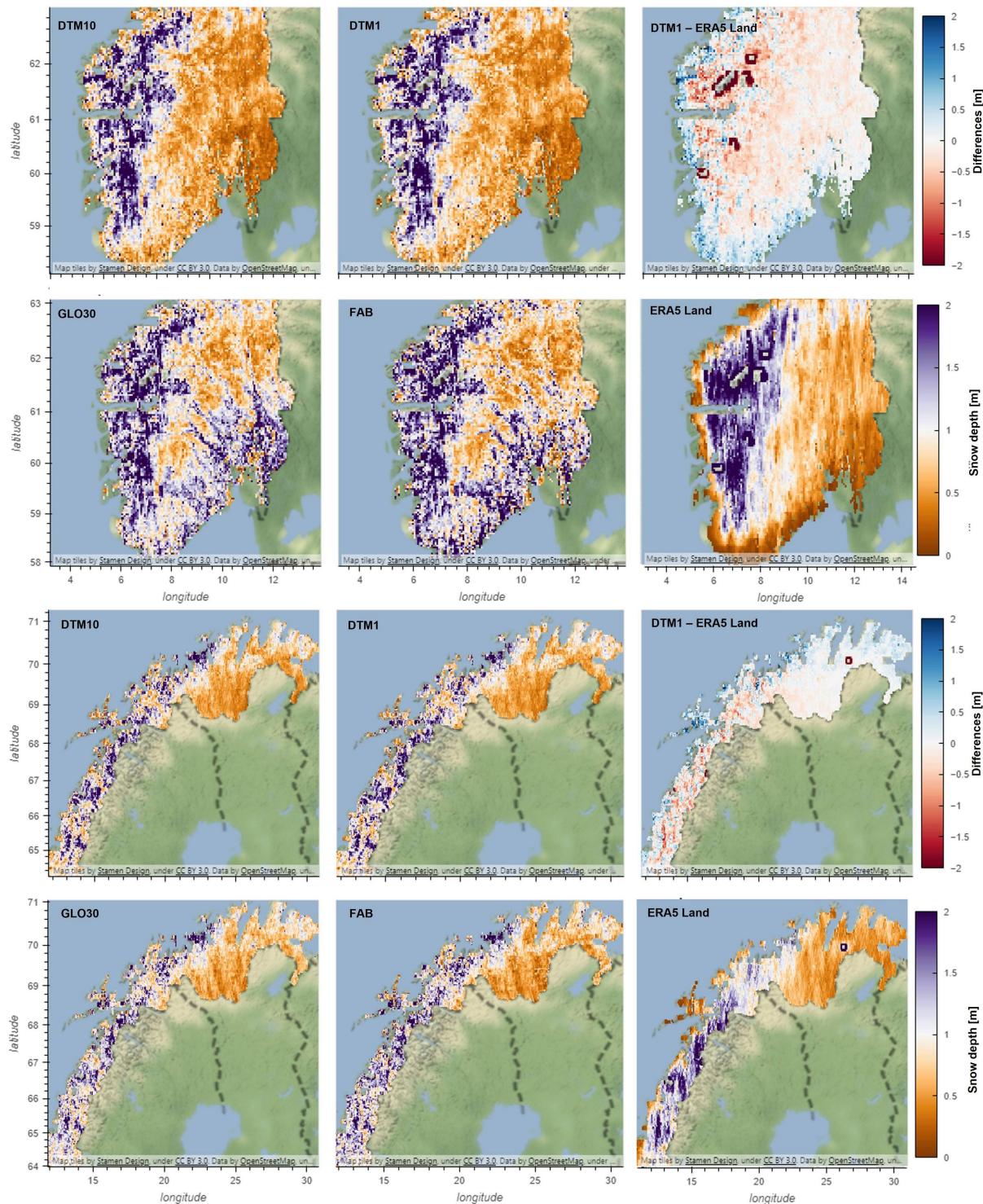


Figure 4.7.: Aggregated Snow Depth from ICESat-2 and Four DEMs, 2018-2022. This figure shows the mean snow depth observed from ICESat-2 over a four-year period, using four DEMs, aggregated at $0.1^\circ \times 0.1^\circ$ grid. Purple areas indicate heavy snow accumulation, while orange indicates minimal or no snow. The cutting-out values used for DTM10 and DTM1 were -0.1m, while GLO30 and FAB used -0.5m (Further discussion about cutting-out in Section 5.3).

4. Results

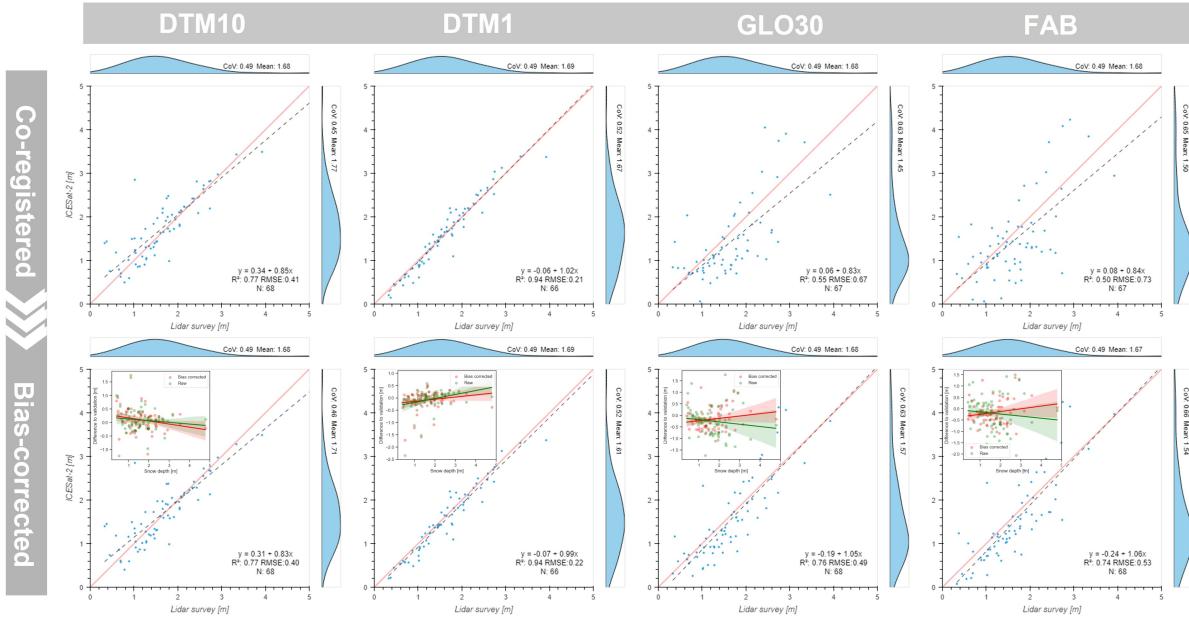


Figure 4.8.: The Validation of ICESat-2 Derived Snow Depth. The snow depth validation are captured by lidar using DTM1 as reference ground. The upper panel is the snow depth derived from DEM after co-registration; the lower panel is the snow depth after bias corrected. In this case, in terms of accuracy, DTM10 is worse than DTM1, slightly better than GLO30, FAB after bias correction.

4.5. ICESat-2 Derived Snow Depth vs ERA5 Land

Figure 4.9 show the difference between ICESat-2 derived snow depth (2018-2022) with snow depth coupled from ERA5 Land reanalysis. The aggregated mean snow depth represents mean bias at the national scales. Reddish pixels represent the overestimation of ERA5 Land snow depth, or could also be interpreted as underestimation of ICESat-2 derived snow depth. The scatter plot indicates that as the snow depth increases, the magnitude of the difference also increases.

The difference can be attributed to several factors. Firstly, due to the edge effect of a 9 km resolution grid, pixels located outside the glacier mask still contain thick snow from aggregation, leading to overestimation (Figure 4.9 a). Similarly, in areas with a transition from snow-free coast to thick snow mountain, ERA5 Land underestimates the snow depth for the same reason (Figure 4.9 b). On the other hand, the ICESat-2 approach in areas with thin snow coverage could also lead to overestimation (Figure 4.9 c). This is related to ‘negative snow depth’ a typical phenomenon from elevation differencing, which results in overestimation if not deal with appropriately. Section 5.3 will further discuss it. However, this comparison does not determine which dataset is more accurate due to the significant differences in representatives and scales.

4.5. ICESat-2 Derived Snow Depth vs ERA5 Land

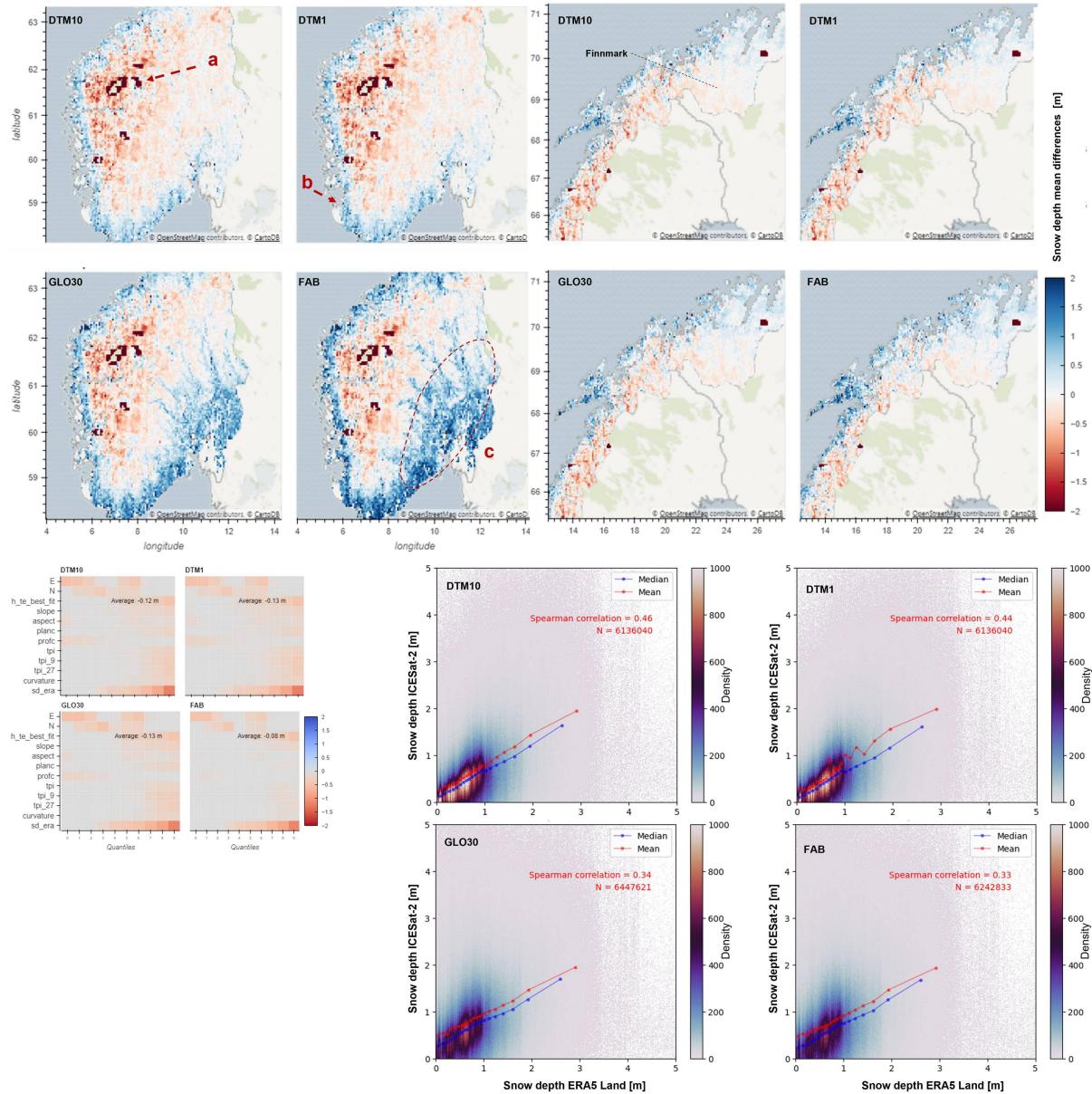


Figure 4.9.: ICESat-2 Derived Snow Depth vs ERA5 Land Snow Depth. The figure illustrates the difference between ICESat-2 derived snow depth (2018–2022) and snow depth from ERA5 Land. The comparison indicates a systematic bias between the two datasets with variations observed in different regions such as glacier areas (a), coastlines (b), and tree coverage areas (c). The Finnmark region shows minimal differences in snow depth measurements. Overall, ICESat-2 derived snow depth is lower than snow depth from ERA5 Land, with Spearman correlation coefficient 0.33 (FAB) to 0.46 (DTM1). And the difference is related to convex terrain. The cutting-out values is 0 for all four models.

4. Results

4.6. Modeled Snow Depth and Validation

The downscaled snow depth reproduces the ‘normal distribution’ of the snow depth at hillslope scales. The validation strip is located in the northwest of the survey area conducted by NVE in 2008, where the hill-shade presents many depressions ranging from 10 to 30 m wide, acting as a shelter with thick snow accumulation (>6 m in validation data), and the hilltops have a thin snowpack instead (Figure 4.10 e, b). Before downscaling, there is no variability in snow depth from ERA5 Land. After downscaling and offset adjustment, the map and profiles demonstrate a good match between model output and validation snow depth. The model differs most for thick snowpacks in depressions (Figure 4.10 a, c, f). The same validation for 2009 can be found in Appendix A.3.

The ATL validation dataset comprises 7,127 sample points at 10 m resolution. By training in a region (100 km wider in all directions than the validation area), all DEMs exhibit R-Squared values ranging from 0.66 to 0.68. The linear relationship between the bias and snow depth can be noticed in all DEMs. Notably, DTM1 demonstrated a Spearman coefficient of 0.81, indicating that the model captured relative snow depth despite significant bias (Figure 4.11 b). However, histograms revealed that the predicted local variability of snow depth is concentrated with narrow variance over four DEMs (Figure 4.11 a), which calls for offset adjustment. Furthermore, the validation area is relatively flat with sparse vegetation, resulting in very similar results for downscaled snow depth from four DEMs.

In Figure 4.12, the distribution of validation snow depth, downscaled snow depth (from ICESat-2 - DTM1 model) and ERA5 Land snow depth are compared. The plan curvature reveals that there is more snow accumulation over concave terrain (negative plan curvature). The same pattern is also captured by downscaled snow depth. The similarity between validation snow depth and downscaled snow depth is shown by vegetation coverage, elevation and east distribution.

In summary, the downscaling of snow depth enables a normal pattern of snow distribution at hillslope scale. Regional regression yields better results than global regression (not shown). Thick snowpacks over complex terrain pose challenges where predicted value loses variance in some cases. Possible optimizations will be discussed in Section 5.4.

4.7. Feature Importance of Downscaling Regression

Four DEMs show remarkable similarity in feature importance. Figure 4.13 shows that elevation ($h_{te_best_fit}$) is the most important predictor for local variability of snow depth. This is followed by slope, aspect, east, north, and snow depth (sd_{era}), TPI (270 m). Even on such local scales, E and N still affect variations of snow depth. In contrast, other predictors such as snow melting accumulation ($smelt_acc$) and snowfall accumulation ($sfall_acc$) were not found to be useful (not shown) and were

4.7. Feature Importance of Downscaling Regression

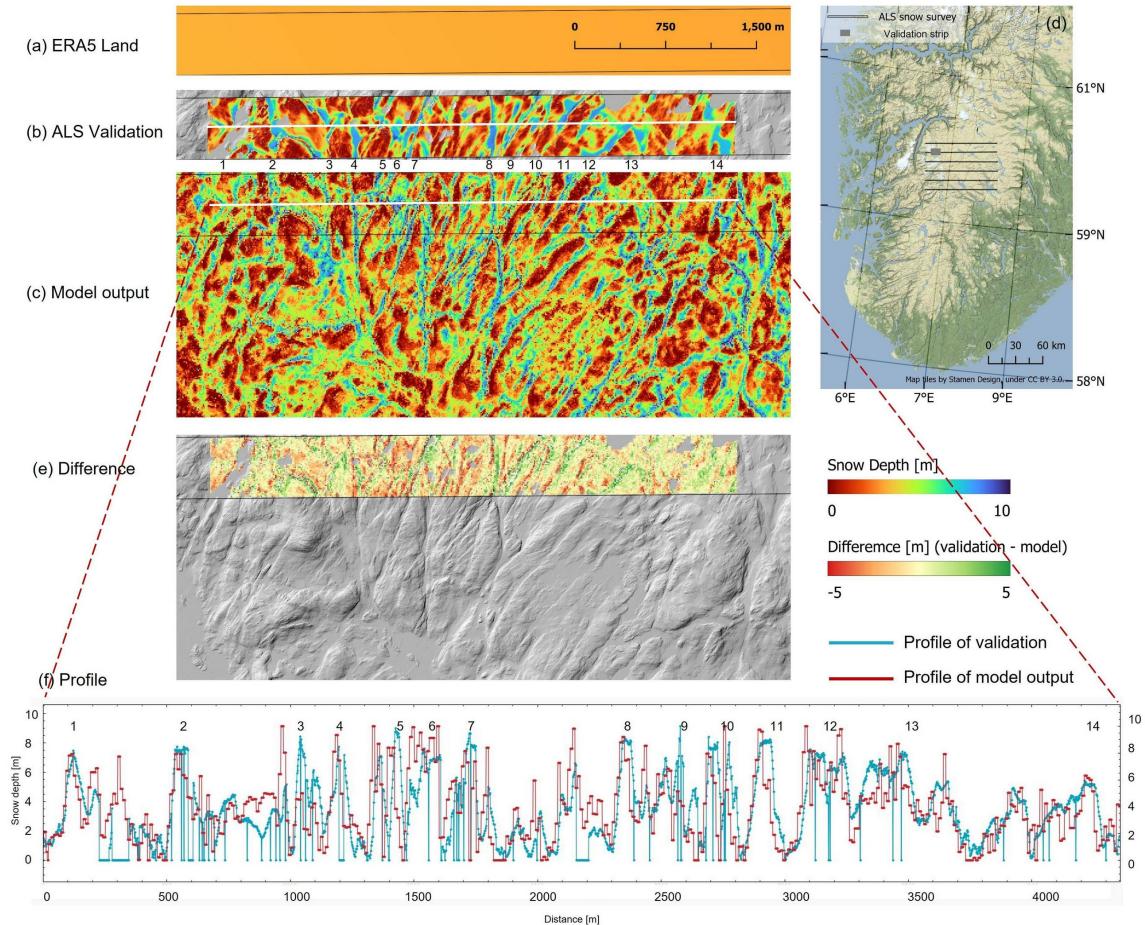


Figure 4.10.: The Validation of Snow Depth Downscaling: Map and Profile. The snow depth from ERA5 Land that needs to be downscaled shows no variability at this scale (a). The NVE's ALS snow survey conducted in April 2008 provided validation strip, which is located in the northwest of the survey area (d) and shows significant variations across depressions and hilltops. The same strip is also displayed in Melvold & Skaugen (2013)'s Figure 3 (b). The downscaled snow depth from the ICESat 2 - DTM1 model reproduced most of the variability (c). Moreover, there were differences between the model and validation, which are shown in (e), while the profile difference is visible in (f). The depressions are numbered 1 to 14, from west to east. Note: The validated snow depth is at a resolution of 2 meters and excludes the water surface, where values are zero. The model is trained in a region that is 100 km wider than the validation area in all directions. The downscaling model output has a resolution of 10 meters and includes water surfaces where values may not be reliable.

4. Results

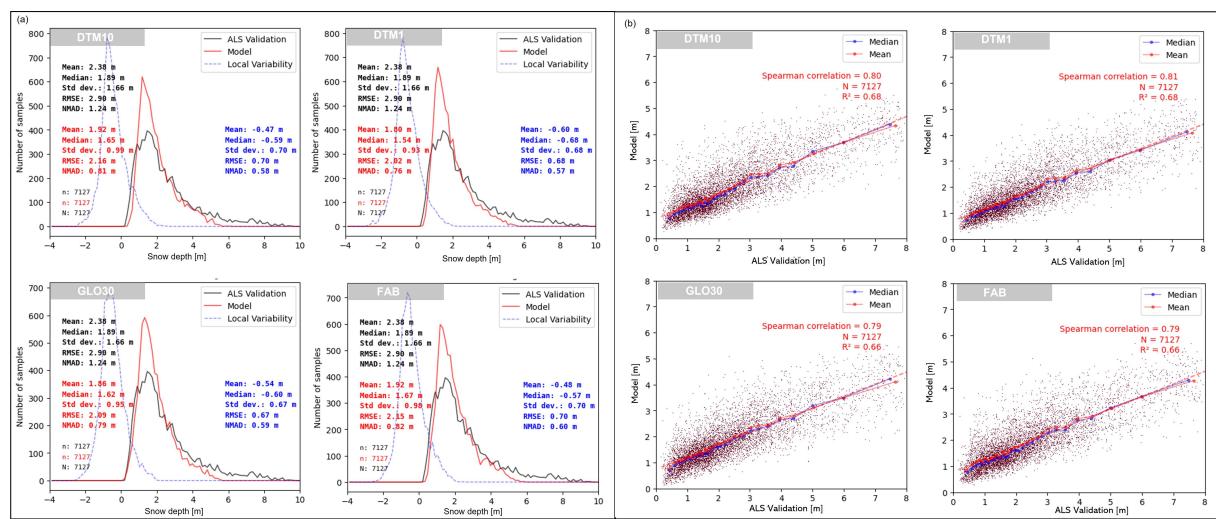


Figure 4.11.: The Validation of Snow Depth Downscaling: Histogram and Scattering. The figure compared four DEMs by histogram (a) scatter plot (b). The histogram of the data showed a narrow distribution, indicating a lack of variability. The scatter plot exhibit a strong Spearman correlation coefficient (0.79 to 0.81), but bias with a linear relationship to validation snow depth. Note: The validated snow depth data is processed to be a 10-meter resolution, which is the same as the model output. Furthermore, the models were trained in a region that is 100 km wider in all direction than validation area.

4.7. Feature Importance of Downscaling Regression

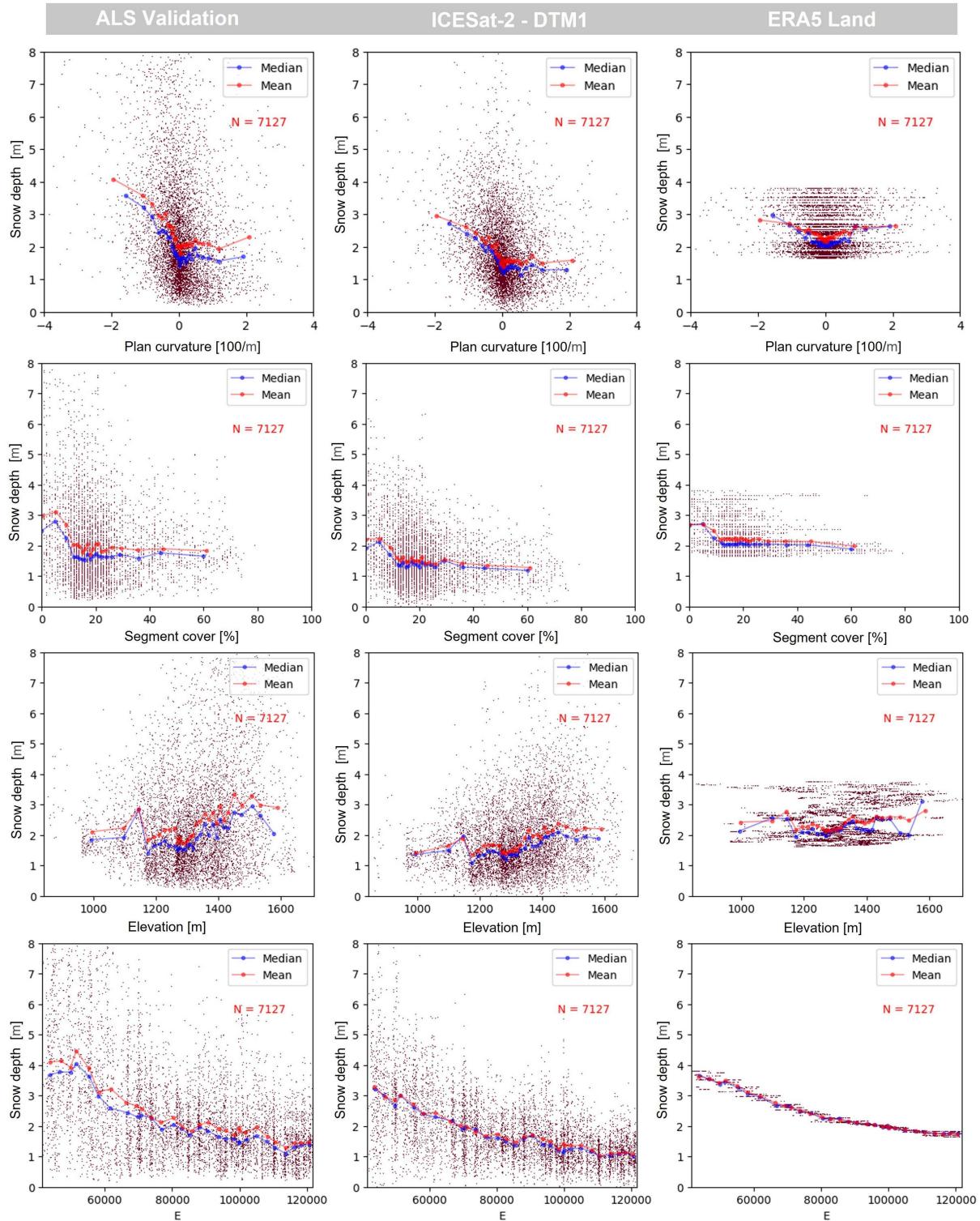


Figure 4.12.: The Validation of Snow Depth Downscaling: Plan Curvature, Vegetation Coverage, Elevation and East. This figure illustrates how downscaling can replicate the variability of snow depth and how it differs from validation. “Segment coverage” refers to the percentage of vegetation in each segment.

4. Results

therefore excluded. *TPI* (270 m) is more important than *TPI* (90 m) and *TPI* (30 m). The *month* is the least important predictor, but cannot be ignored (not shown).

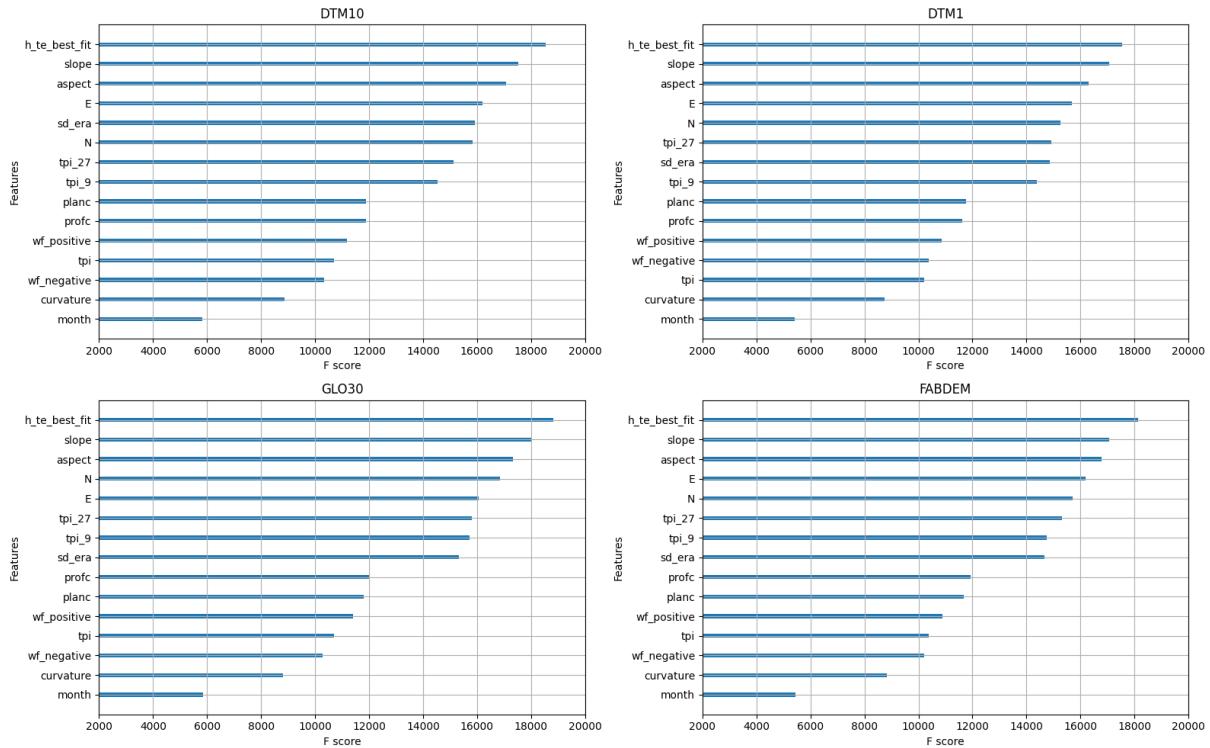


Figure 4.13.: Feature Importance of Snow Depth Downscaling. The elevation (*t_te_best_fit*) is the most significant feature for local variability of snow depth, followed by *slope*, *aspect*, and *E*, *N*, snow depth (*sd_era*), or *TPI_27*. Even on a local scale, *E* and *N* still affect variations of snow depth within smaller areas. For a full explanation of each feature's name, please refer to Section 3.3.1

5. Discussion

5.1. GDC vs NuthKaab Co-registration

GDC (Gradient Descent Co-registration) and NuthKaab (Nuth & Kääb, 2011) address the same problem but using distinct techniques. These methods work when geo-referenced DEMs are not perfectly aligned. In cases where geo-referencing information is absent, LS3D (Least Squares 3D Surface Matching) (Gruen & Akca, 2005), or other ICP-based algorithm, are preferred. GDC and NuthKaab solve the co-registration problem as a bound-restricted minimizing problem with random noise. This problem is often referred to as sub-pixel co-registration, which is true in most cases where the resolution is 10 m or coarser, with a misalignment of less than 10 m. However, at resolutions of 1 m or 2 m, it is no longer considered a sub-pixel problem. In such cases, it is recommended to solve it within the bounds of 3-5 pixels, as suggested by this study (Figure 4.1).

NuthKaab solves the co-registration problem by minimizing NMAD statistically (as implemented in xDEM) and requires calculations of terrain parameters such as slope, aspect, and sometimes curvature at least once. It operates efficiently for low workload tasks with coarse resolution or small coverage of DEMs. On the other hand, GDC solves the same problem more quickly and without requiring terrain parameters calculation. The gradient descent algorithm is used to expedite searching. During co-registration with NuthKaab, the DEM to be aligned is iteratively shifted by updating the transformation parameters in each iteration, requiring interpolation to calculate elevation differences until convergence. However, GDC only performs interpolation task on selecting points.

From a technical perspective, for scenarios requiring fine resolution DEM or extensive coverage involving large workloads, GDC is highly recommended over NuthKaab. Additionally, GDC is more suitable for co-registering two DEMs with different resolutions or for co-registering a DEM with points. Although some implementations of NuthKaab (*Xdem*, 2021) can also handle point co-registration, it has critical requirements regarding the number of points.

This study did not perform track-by-track co-registration of ICESat-2 data, which could lead to even greater improvements in accuracy of co-registration. Further development of gradient descent co-registration could support rotation, which would enhance the algorithm's versatility and applicability.

5. Discussion

5.2. The Uncertainties: ICESat-2, DEM and Bias Correction

5.2.1. ICESat-2 ATL08: skewness and curvature

The most critical issue with ATL08 products identified in this study, as well as previous studies, is from vertical sampling error (other quality issues please see Appendix A.4). The ground-finding algorithm of ICESat-2 relies on a probability distribution function (PDF) of the reflecting surface. So that, dense vegetation and the presence of cloud (Moudrý et al., 2022) can hinder photons from accurately reaching the ground surface, leading to this error (A. Neuenschwander & Magruder, 2016). This error can also impact canopy height determination consequently. Additionally, slope naturally translates horizontal shift (i.e., footprint or geolocation error) into vertical shift (Deems et al., 2013), which may be related to negative skewness noted in many studies (Enderlin et al., 2022; Moudrý et al., 2022; Tian & Shan, 2021). This study highlights that negative bias exists on all pairs of ATL08 snow-free - DEMs in terms of slope, with GLO30 and FAB exhibiting more severe negative bias than DTM10 and DTM1 (Figure 4.2, A.2). However, the significant negative bias is observed only on convex terrain while being totally different over concave terrain indicated by positive TPI, curvature, plan curvature and negative profile curvature (Figure 4.2, 4.3).

To further investigating the negative skewness, Figure 5.1 plots nine curvature combinations categorized by profile curvature and plan curvature. The aggregated mean value of elevation differences reveals that convex ridged terrain (positive plan curvature and negative profile curvature) typically exhibits the negative skewness (Figure 5.1 a,g). A concave bowl-shaped terrain with negative profile curvature does not have much negative skewness, even at high slopes (Figure 5.1 b), thereby indicating that the observed linear relationship between bias and slope (Figure 4.2 b) is mainly contributed by plan curvature. Moreover, the canopy height contributes significantly to positive bias when the canopy is short upon concave terrain, but it cannot equalize for all the observed negative skewness (Figure 5.1 c). Therefore, it can be inferred that plan curvature plays a significant role in influencing skewness. Based on these findings, this study does not recommend any bias correction based solely on slope and/or canopy height. Instead, the bias correction method used in this study can remove the negative skewness over different features (Figure 5.1 d,e,f).

The mechanism responsible for the observed negative skewness remains unclear. In a recent study, Moudrý et al. (2022) argued that increasing atmospheric scattering effects lead to underestimated terrain height due to the resulting increase in photon path length. To assess these effects, the authors suggested utilizing the cloud confidence flag (*cloud_flag_atm*). However, Figure A.5 showed that terrain plan curvature exhibited a strong correlation to negative skewness even in clear sky conditions. Thus, the negative skewness is yet an open question, and researchers must consider both factors mentioned in this study when assessing terrain height.

If the DTM1 is deemed to be an accurate representation of the surface, this study then can attribute the

5.2. The Uncertainties: ICESat-2, DEM and Bias Correction

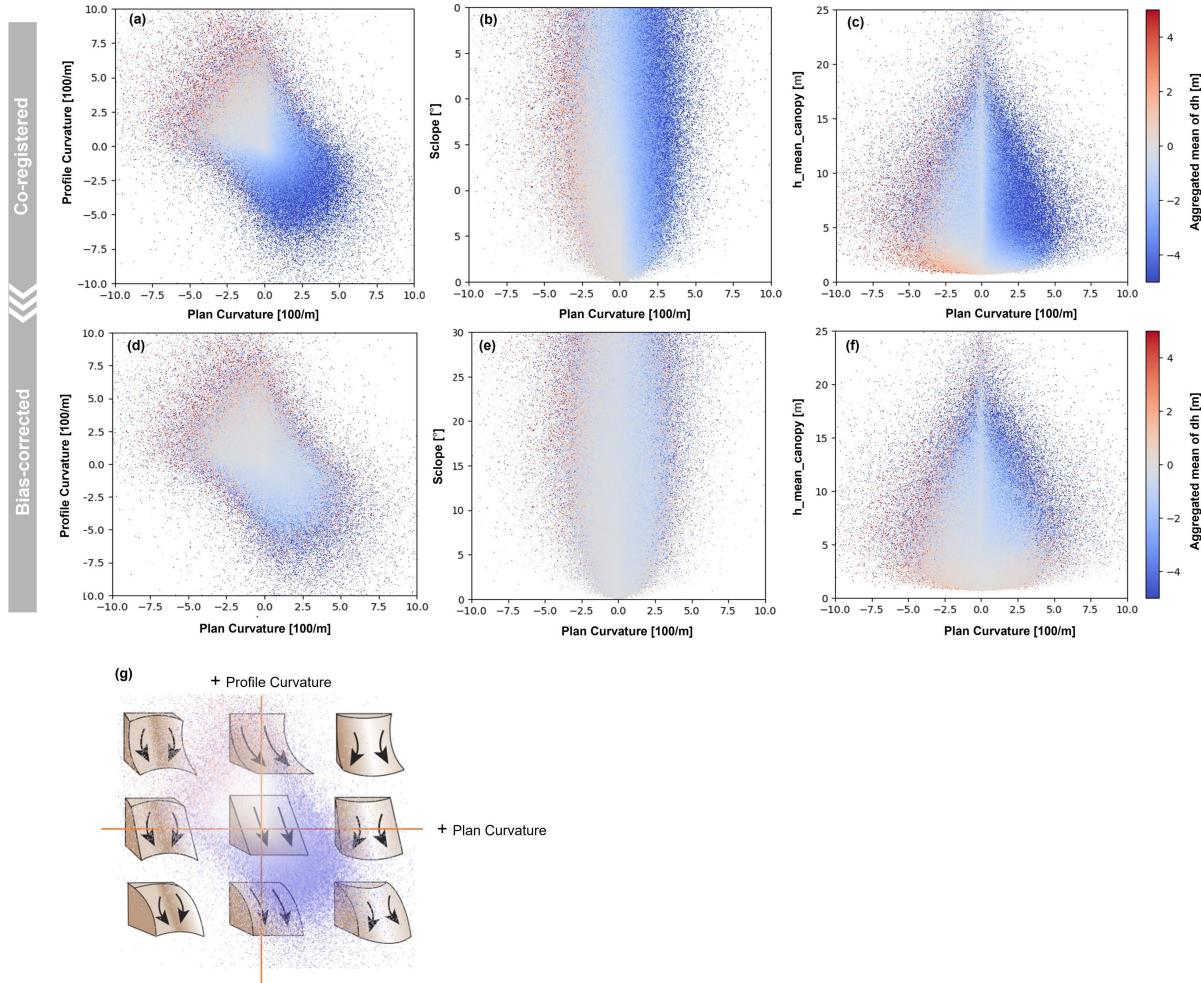


Figure 5.1.: Skewness Variations in Convex and Concave Terrains. The elevation difference is between ICESat-2 snow free segments and DTM1. The color points indicate the aggregated mean value of the elevation difference. Plots (a), (b) and (c) depict the bias pattern along plan curvature, profile curvature, slope and canopy height before the bias correction. A triangle window (a) indicating a bias free condition, the rest of area either give negative bias (blue) or positive bias (red). After the bias correction, most of negative bias are removed (d, e, f). Plot (g) shows a schematic of curvature combinations. The plot reveals that there is a negative correlation between profile curvature and plan curvature. For instance, the negative profile curvature always indicates positive profile curvatures Peckham (2011). Plot (g) is modified from Kimerling et al. (2016).

5. Discussion

bias to ATL08. In such scenario, it is plausible that ATL08's ground finding algorithm employed has a tendency to capture lower signals, resulting in an underestimation of convex surfaces while accurately capturing concave surfaces. The two possible components contribute:

- (1) Naturally, there is overestimation over concave, but underestimation of convex due to the limitation of coarse resolution (or footprint, see examples in Figure 5.2 a, A.4). However, this should be symmetrical bias and not as significant as what we observed in terms of negative skewness.
- (2) Trees present in concave and convex may lead to skewness. As observed overestimation on concave terrain (Figure 5.1 c, with a reddish tint at the bottom left) may have occurred due to the misclassification of short canopy as terrain surface. This misclassification is less likely to occur in convex terrain or areas with taller trees because trees are discernible from their surroundings due to their height.

However, both misclassification and resolution limitations cannot fully account for such a significant negative skewness. On the other hand, if ATL08 provides an accurate surface representation instead, it is possible that DTM1 overestimates convex surface.

5.2.2. Uncertainties in DEMs: vegetation, resolution, snowpack

The presence of surface objects, such as buildings and vegetation, also introduce negative bias into the residuals (see Figure 4.5). Thus, it is crucial to assess whether bias correction can eliminate the impact of surface objects. Comparing COP30 with FAB DEMs can help us investigate this issue. To facilitate the comparison, this study uses DTM1 as a reference surface for evaluating both DEMs before and after bias correction. This will provide insights into the effectiveness of bias correction methods in mitigating the impact of surface objects.

The Figure 5.2 illustrates the differences between FAB, GLO30, and DTM1 in terms of plan curvature and canopy height. The scatter points indicates a symmetrically spread of trees across convex and concave terrain. FAB has removed the forest, leading to a positive bias over concave areas and a negative bias over convex areas due to resolution limitations (Figure 5.2 a). On the other hand, GLO30 has a higher surface than DTM1 because of the presence of trees, which masks the difference except when canopy height cancels with negative bias over convex terrain (Figure fig@:dem_dtm1 b). The bias correction process involves using ICESat-2 snow-free segments to correct DEMs. This results in corrected DEMs with similar patterns (Figures 5.2 c, d and 5.1 c). After applying the same bias correction on DTM1, FAB remains higher over concave terrain and lower over convex terrain compared to DTM1 (Figure 5.2 e). Meanwhile, GLO30 retains vegetation height primarily for concave terrain (Figure 5.2 f).

Both GLO30 and FAB exhibit residual errors from their initial state that can potentially impact the determination of snow depth, as demonstrated in Figure 4.7. Moreover, as shown in Figure 5.2 a, b to e, f,

5.2. The Uncertainties: ICESat-2, DEM and Bias Correction

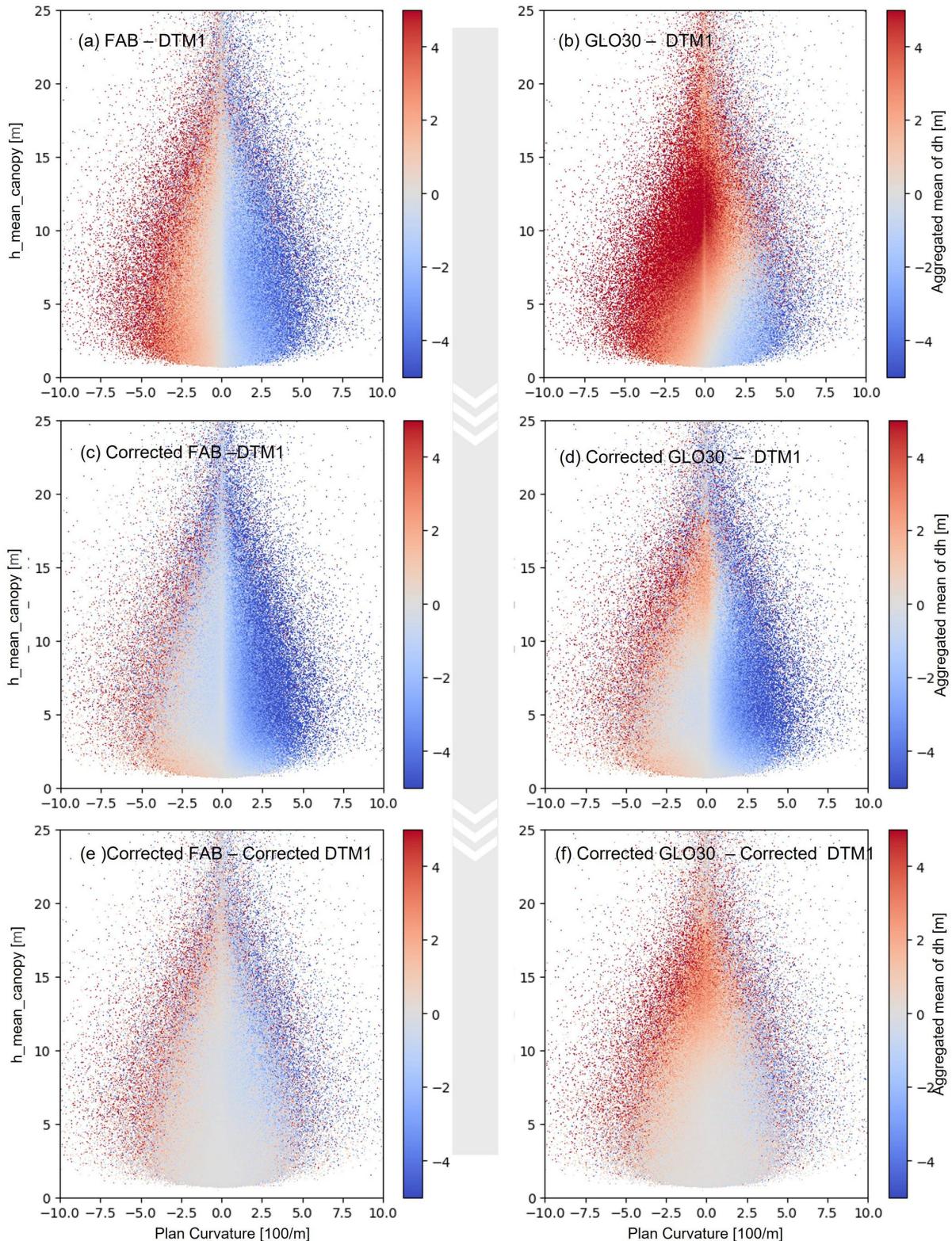


Figure 5.2.: Comparison of FAB, GLO30, and DTM1 in terms of Plan Curvature and Canopy Height. The color points indicate the aggregated mean value of the elevation difference of DEMs. From a,b to c,d to e,f, this figure show how bias correction affects the dh .

5. Discussion

the bias correction has led to the creation of a bias-free window over flat terrain with short vegetation cover, thereby enhancing the consistency of results obtained from different DEMs.

DTM1 and DTM10 are generated using lidar technology, which involves the crucial step of removing non-ground objects such as buildings and vegetation (X. Liu, 2008). It's also important to check whether these DEMs contain any snowpacks. Since the training elevation data from snow-free segments is strictly snow free, ideally, it will correct DEM to bare ground conditions. However, this study lacks a validation dataset to confirm that the DEM has been corrected to bare ground conditions..

5.2.3. Assumptions behind bias correction: snow-on vs snow-free

It is important to note that the bias correction model is trained on a snow-free surface bias correction, and only corrects the DEMs to their “ICESat-2 ATL08 snow-free condition,” which is not equivalent to the “true surface” but rather still encompasses bias patterns of ICESat-2 ATL08 snow-free segments.

The workflow presented in this study assumes that there is a consistent bias pattern in ICESat-2 elevation measurements, regardless of the presence or absence of snow. However, when transitioning from a snow-free surface to a snow-covered surface, the bias of ICESat-2 may fluctuate due to changes in surface roughness, reflectivity, and canopy height caused by snow cover. Generally, more photons are available, leading to lower uncertainties. This is reflected in features such as the number of photons ($n_{te_photons}$), signal-to-noise-ratio (snr), and $h_{te_uncertainties}$ (Figure 5.3). However, currently there is no such fieldwork or studies to compare the bias pattern of snow-off segments and snow-on segments.

This study trained the bias correction model using snow-off segments, which may result in an ‘over-estimation’ of predicted bias correction for snow-on segments. This is especially true in areas where the DEM provides accurate surface elevation and the main bias is from ICESat-2. The problem can be solved if the bias correction model can differentiate between biases introduced by ICESat-2 and those caused by the DEMs themselves.

5.3. ICESat-2 Derived Snow Depth: Cutting-out

Table 5.1.: Impact of Cutting-out Negative Values on Snow Depth Statistics

Data	Cutting-out	Mean	Median	Std	Snow depth < cutting-out
ERA5 Land	-	0.88 m	0.74 m	0.65 m	-
ICESat-2 - DTM1	-1 m	0.62 m	0.42 m	1.02 m	7.93%

5.4. Statistical Downscaling: Noise and Randomization

Data	Cutting-out	Mean	Median	Std	Snow depth < cutting-out
ICESat-2 - DTM10	-0.5 m	0.72 m	0.47 m	0.98 m	14.19 %
	-0.2 m	0.81 m	0.54 m	0.97 m	21.36 %
	0	0.92 m	0.62 m	0.97 m	29.54 %
	-1 m	0.62 m	0.42 m	1.02 m	7.75%
	-0.5 m	0.72 m	0.48 m	0.99 m	13.95 %
	-0.2 m	0.81 m	0.54 m	0.98 m	21.05 %
	0	0.92 m	0.62 m	0.98 m	29.21 %
	-1 m	0.84 m	0.64 m	1.12 m	11.75%
	-0.5 m	0.95 m	0.70 m	1.08 m	17.27%
	-0.2 m	1.04 m	0.77 m	1.06 m	22.80 %
ICESat-2 - GLO30	0	1.13 m	0.83 m	1.05 m	28.26 %
	-1 m	0.81 m	0.58 m	1.20 m	12.52 %
	-0.5 m	0.94 m	0.66 m	1.16 m	19.19%
	-0.2 m	1.06 m	0.74 m	1.14 m	25.80 %
ICESat-2 - FAB	0	1.16 m	0.82 m	1.14 m	31.96%

The bias correction applied in this study made the distribution more symmetrical, but it did not eliminate negative values (Figure 5.4). About 30% of the data fell below 0 m, and after removing these negative values, the median and mean shifted by 0.2-0.3 m on four DEMs (Table 5.1). In practice, Melvold & Skaugen (2013) used a threshold of -1 m as a cutting-out value in NVE's ALS snow survey. This study selected a cutting value of -0.5 m for GLO30 and FAB and -0.1 m for DTM1 and DTM10. However, there are no rules for determining the best cutting-out values yet. The distribution of negative values is related to survey techniques, DEM qualities, topographic factors, and snowpack thickness. This study noticed different areas and different method may require different cutting-out values.

As the cutting-out of negative snow depth can significantly impact snow depth statistics and results of aggregation, particularly in coastal regions where the snowpack is thin, researchers must exercise caution when selecting thresholds for cutting-out negative values.

5.4. Statistical Downscaling: Noise and Randomization

The downscaling of snow depth has several limitations from data sources. Firstly, the variables used are insufficient to fully express the snow depth distribution, i.e., a lack of variability in the input fea-

5. Discussion

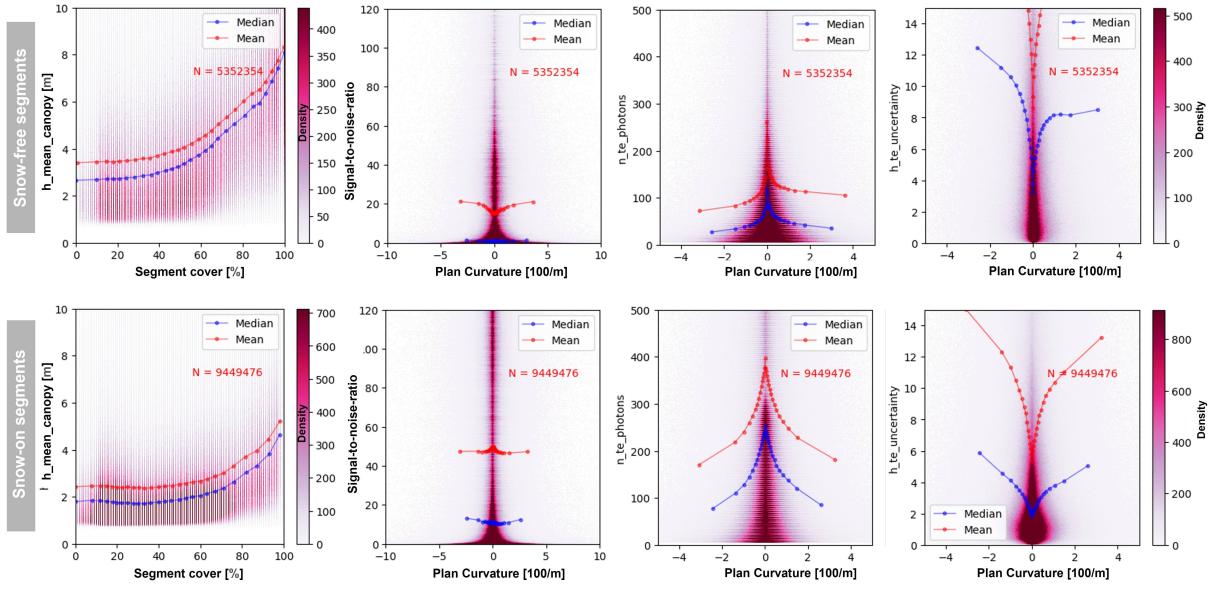


Figure 5.3.: Features Related to Snow-on and Snow-free Surface. The surface features exhibit significant differences in terms of canopy height, photon count, signal-to-noise ratio, and terrain surface uncertainties. For example, the canopy height of a snow-free segment is much greater than that of a snow-covered segment under the same percentage of vegetation cover. The density plot has been normalized based on the number of data points.

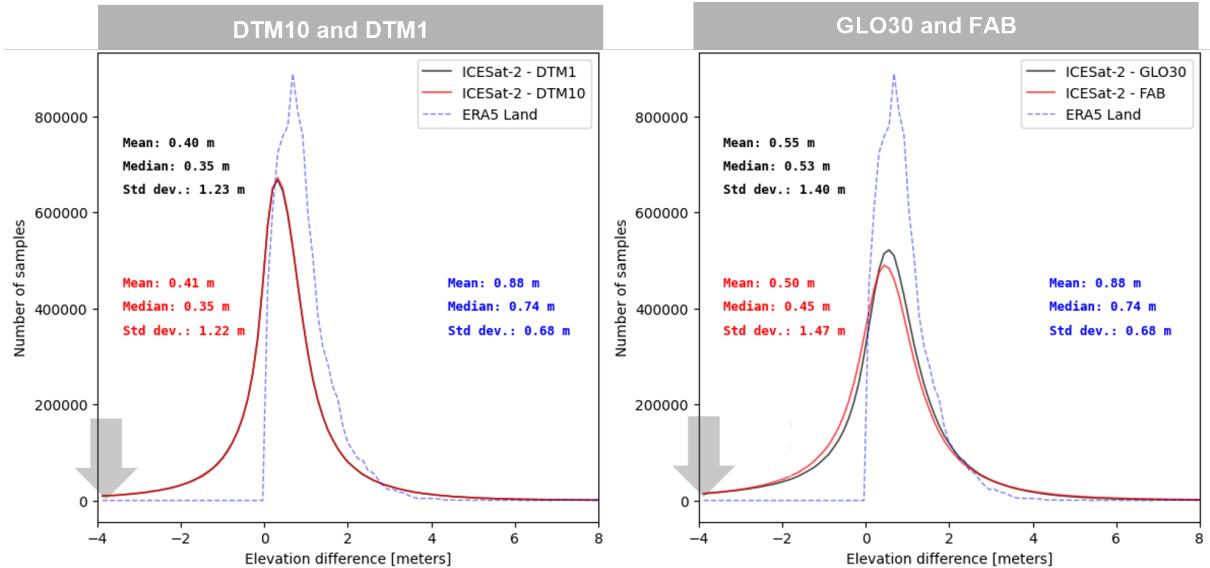


Figure 5.4.: The Negative Snow Depth and Cutting-out Value. Values below the threshold are excluded from the statistics, resulting different statistics. Histograms depict the statistics median, mean with a cutting-out thresholds of -4 m (arrow). When the arrow move toward zero, Table 5.1 show the difference.

5.4. Statistical Downscaling: Noise and Randomization

tures. For instance, the model ignores the seasonal changes in vegetation, and the wind variables are coarse in resolution (9 km), which does not fully capture the variability of wind with respect to vegetation, terrain aspect, and surface roughness. Additionally, potential errors may be introduced by the ERA5 data as its consistency during accumulation and melting periods may vary. Therefore, the bias pattern could differ but the regression model does not intentionally correct for this inconsistency. Lastly, a historical training dataset of 4 years is insufficient since climate studies require a 10 to 30 year dataset to represent climate normal (Arguez & Vose, 2011).

When downscaling, reproducing snow that is thicker than 5 meters was difficult and require additional offset adjustments. From a perspective of training dataset, if the training snow depth and predictors has unpredictable noise, it can cause the regression model to have a harder time finding meaningful patterns in the data, resulting in a model that is less able to capture the full range of variability in the target variable. This can lead to predictions that are more tightly clustered around the mean or median value of the target variable, showing lower overall variance. From a mathematical standpoint, a deterministic regression solely predicts the variance that can be explained by the predictor, disregarding the random processes (unexplained variability). Consequently, the predicted values are less variable or ‘flatter’ than the source data.

To improve results, there are several possible ways can be explored. Firstly, it is important to note that snow depth follows a gamma distribution as shown in Gisnås et al. (2016). Consequently, instead of assuming that $x - \mu$ conforms to a normal distribution in regression, where x denotes sample values of snow depth and μ denotes the true mean of the population (the snow depth from ERA5 Land), it is preferable to provide prior knowledge about the distribution of local variability of snow. Additionally, quality control should be implemented on ICESat-2 derived snow depth data by incorporating supplementary datasets. Lastly, integrating third-party data sources can enhance the feature engineering of vegetation parameters.

In summary, improving downscaling methods for snow depth requires addressing limitations such as inadequate and inaccurate predictor variables while considering potential errors from ERA5 data and historical training datasets’ short length. Possible solutions include using gamma distributions for accurate representation of local variability in snow depths; incorporating additional imagery data for quality control of ICESat-2 derived snow depth and improving feature engineering.

6. Concluding Remarks and Future Outlooks

Seasonal snow accumulation is difficult to observe and model accurately. To obtain accurate snow depth measurements through elevation differencing, repeated observations are necessary under both snow-on and snow-free conditions. Using DEMs as a reference requires careful consideration of their quality and efficient co-registration. Overall, This study provides a comprehensive workflow for accurately measuring snow depth using ICESat-2 data and demonstrates the potential of machine learning techniques for correcting DEM bias and statistical downscaling.

Four different DEMs were utilized as a reference ground. including DTM1, DTM10, Copernicus GLO30, and FABDEM in which forest and building it removed from GLO30. The data were co-registered to ICESat-2 ATL08 snow-free segments and a XGBoost model was applied to remove elevation bias. The study retrieved 9.2 million snow depth measurements with an estimated overall precision NMAD of 0.55 m, 0.55 m, 0.90 m, and 1.04 m for each respective DEM. The ICESat2-DTM1 derived snow depth showed excellent correlation (R-Squared 0.94) with lidar-based snow depth validation from Spring 2022 (67 points). Bias correction is essential for elevation differencing, e.g. GLO30 and FAB demonstrated an improvement after bias correction in their R-square values from 0.55 to 0.76 and 0.50 to 0.74 respectively. While downscaling was successful in reproducing snow depth at a hill-slope scale level with Spearman correlation coefficient of 0.79 to 0.81 and R-square values 0.66 to 0.68 in the validation area, it proved challenging to reproduce the snowpack over a thickness of five meters.

In addition to the main results presented above, this study also yielded several noteworthy findings and understandings:

- (A) This study proposes a novel algorithm, Gradient Descent Co-registration, to address the challenge of DEM co-registration. The method has greatly improved efficiency and opened up new possibilities for handling large-scale datasets at a fine resolution. The algorithm has been published as a tool (Appendix A.5) for the cryosphere and remote sensing community.
- (B) A bias correction workflow has been demonstrated to be a critical step for the elevation differencing workflow of snow depth. The results showed that using the global DEM GLO30 had better performance than previous studies, with an overall NMAD of 1.45 m. Bias correction further improved NMAD to below 1 m in mainland Norway. However, bias correction methods may not fully eliminate biases when accurate vegetation information is lacking. When using FABDEM, which employs a random forest algorithm to eliminate forests from GLO30 and integrates var-

6. Concluding Remarks and Future Outlooks

ious data sources, remaining biases were still evident in vegetated regions. Nonetheless, flat areas with sparse vegetation like Hardangervidda and Finnmark may be more manageable for retrieving snow depth using GLO30 and FAB. So that the approach still has the potential to be employed globally.

- (C) Selecting appropriate cutting-out thresholds is crucial for the elevation differencing method. By using dynamic thresholds under specific conditions, results on thin snowpack can be improved.
- (D) The negative skewness of ICESat-2 ATL08 has been observed in many studies. This study highlights the critical role played by curvature and vegetation, which is more important than slope (Enderlin et al., 2022; Tian & Shan, 2021) and atmospheric scattering (Moudrý et al., 2022). The bias correction proposed in this study can resolve this issue.
- (E) ICESat-2 offers continuous global measurements of snow surface profile throughout the season at footprint size of ~17 m, providing valuable observations for subgrid snow distribution. Statistical downscaling is a promising technique to utilize the ICESat-2 derived snow depth. It extracts empirical snow distribution at hillslope scales. Moreover, with an additional bias correction step, the further study can work on climate model outputs, projecting snow distribution in future scenarios.
- (F) This study found a systematic difference between the ICESat-2 derived snow depth and ERA5 Land snow depth. Comparing two dataset's correctness is challenging because of differences in snow depth representation and scales. The future study can look further into the differences.
- (G) The downscaled snow depth data derived from ICESat-2 has numerous potential applications for future research. One example is the impact of snow depth on the thermal regime of permafrost. Accurate representation of snow depth at the hill-slope scale can lead to improvements in discharge modeling, more precise estimation of snow sublimation in wind-exposed mountain ridges, and avoidance of unrealistic snow cover in hydrological models. Additionally, it can improve the representation of vegetation responses to snow conditions, facilitate better risk assessment of avalanches, and aid in various other studies related to snow and its interactions with the environment.
- (H) Machine learning is a powerful tool that has been demonstrated to be effective in correcting DEM bias and statistical downscaling in this study. In contrast to traditional statistics, machine learning is capable of handling large datasets and intricate relationships between variables. However, the elevation differencing approach may not be dependable when the snowpack is thin. In these cases, satellite imagery can be an invaluable source of spatial variability time series of snow cover, as it provides prior knowledge of conditions. Consequently, a scalable ensemble of machine learning models, involving collaboration between experts from various fields, including remote sensing, meteorology, hydrology, and machine learning, is essential for successfully measuring and modeling snow depth.

References

- Abdalla, S., Abdeh Kolahchi, A., Ablain, M., Adusumilli, S., Aich Bhowmick, S., Alou-Font, E., Amarouche, L., Andersen, O. B., Antich, H., Aouf, L., Arbic, B., Armitage, T., Arnault, S., Artana, C., Aulicino, G., Ayoub, N., Badulin, S., Baker, S., Banks, C., ... Zlotnicki, V. (2021). Altimetry for the future: Building on 25 years of progress. *Advances in Space Research*, 68(2), 319–363. <https://doi.org/10.1016/j.asr.2021.01.022>
- Alonso-González, E., Aalstad, K., Baba, M. W., Revuelto, J., López-Moreno, J. I., Fiddes, J., Essery, R., & Gascoin, S. (2022). The Multiple Snow Data Assimilation System (MuSA v1.0). *Geoscientific Model Development*, 15(24), 9127–9155. <https://doi.org/10.5194/gmd-15-9127-2022>
- AMAP. (2017). *Snow, Water, Ice and Permafrost in the Arctic (SWIPA) 2017*.
- Amazon. (2023). *What Is Amazon SageMaker? - Amazon SageMaker*. <https://docs.aws.amazon.com/sagemaker/latest/dg/whatis.html>
- Anderson, E. A. (1968). Development and testing of snow pack energy balance equations. *Water Resources Research*, 4(1), 19–37. <https://doi.org/10.1029/WR004I001P00019>
- Arguez, A., & Vose, R. S. (2011). The Definition of the Standard WMO Climate Normal: The Key to Deriving Alternative Climate Normals. *Bulletin of the American Meteorological Society*, 92(6), 699–704. <https://doi.org/10.1175/2010BAMS2955.1>
- Barnett, T. P., Adam, J. C., & Lettenmaier, D. P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, 438(7066), 303–309. <https://doi.org/10.1038/nature04141>
- Bennett, K. E., Miller, G., Busey, R., Chen, M., Lathrop, E. R., Dann, J. B., Nutt, M., Crumley, R., Dillard, S. L., Dafflon, B., Kumar, J., Bolton, W. R., Wilson, C. J., Iversen, C. M., & Wullschleger, S. D. (2022). Spatial patterns of snow distribution in the sub-Arctic. *The Cryosphere*, 16(8), 3269–3293. <https://doi.org/10.5194/tc-16-3269-2022>
- Besl, P. J., & McKay, N. D. (1992). A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2), 239–256. <https://doi.org/10.1109/34.121791>
- Bormann, K. J., Brown, R. D., Derksen, C., & Painter, T. H. (2018). Estimating snow-cover trends from space. *Nature Climate Change*, 8(11), 924–928. <https://doi.org/10.1038/s41558-018-0318-3>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Brown, R. D., & Mote, P. W. (2009). The Response of Northern Hemisphere Snow Cover to a Changing Climate*. *Journal of Climate*, 22(8), 2124–2145. <https://doi.org/10.1175/2008JCLI2665.1>

6. Concluding Remarks and Future Outlooks

- Buchhorn, M., Smets, B., Bertels, L., Roo, B. D., Lesiv, M., Tseddbazar, N.-E., Herold, M., & Fritz, S. (2020). *Copernicus Global Land Service: Land Cover 100m: Collection 3: Epoch 2019: Globe*. Zenodo. <https://zenodo.org/record/3939050>
- Callaghan, T. V., & Johansson, M. (2021). Snow, ice, and the biosphere. In *Snow and Ice-Related Hazards, Risks, and Disasters* (pp. 137–164). Elsevier. <https://linkinghub.elsevier.com/retrieve/pii/B9780128171295000123>
- Callaghan, T. V., Johansson, M., Brown, R. D., Groisman, P. Ya., Labba, N., Radionov, V., Bradley, R. S., Blangy, S., Bulygina, O. N., Christensen, T. R., Colman, J. E., Essery, R. L. H., Forbes, B. C., Forchhammer, M. C., Golubev, V. N., Honrath, R. E., Juday, G. P., Meshcherskaya, A. V., Phoenix, G. K., ... Wood, E. F. (2011). Multiple Effects of Changes in Arctic Snow Cover. *AMBIO*, 40(S1), 32–45. <https://doi.org/10.1007/s13280-011-0213-x>
- Carrer, M., Dibona, R., Prendin, A. L., & Brunetti, M. (2023). Recent waning snowpack in the Alps is unprecedented in the last six centuries. *Nature Climate Change*, 1–6. <https://doi.org/10.1038/s41558-022-01575-3>
- Carroll, M. L., Townshend, J. R., DiMiceli, C. M., Noojipady, P., & Sohlberg, R. A. (2009). A new global raster water mask at 250 m resolution. *International Journal of Digital Earth*, 2(4), 291–308. <https://doi.org/10.1080/17538940902951401>
- Chang, A. T. C., Foster, J. L., & Hall, D. K. (1987). Nimbus-7 SMMR Derived Global Snow Cover Parameters. *Annals of Glaciology*, 9, 39–44. <https://doi.org/10.3189/S0260305500200736>
- Chen, T., & Guestrin, C. (2016). XGBoost. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Chen, W., Yao, T., Zhang, G., Li, F., Zheng, G., Zhou, Y., & Xu, F. (2022). Towards ice-thickness inversion: An evaluation of global digital elevation models (DEMs) in the glacierized Tibetan Plateau. *The Cryosphere*, 16(1), 197–218. <https://doi.org/10.5194/tc-16-197-2022>
- Dee, D. P., Balmaseda, M., Balsamo, G., Engelen, R., Simmons, A. J., & Thépaut, J.-N. (2014). Toward a Consistent Reanalysis of the Climate System. *Bulletin of the American Meteorological Society*, 95(8), 1235–1248. <https://doi.org/10.1175/BAMS-D-13-00043.1>
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., ... Vitart, F. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656), 553–597. <https://doi.org/10.1002/qj.828>
- Deems, J. S., Painter, T. H., & Finnegan, D. C. (2013). Lidar measurement of snow depth: A review. *Journal of Glaciology*, 59(215), 467–479. <https://doi.org/10.3189/2013JoG12J154>
- Deschamps-Berger, C., Gascoin, S., Berthier, E., Deems, J., Gutmann, E., Dehecq, A., Shean, D., & Dumont, M. (2020). Snow depth mapping from stereo satellite imagery in mountainous terrain: Evaluation using airborne laser-scanning data. *The Cryosphere*, 14(9), 2925–2940. <https://doi.org/10.5194/tc-14-2925-2020>

- Deschamps-Berger, C., Gascoin, S., Shean, D., Besso, H., Guiot, A., & López-Moreno, J. I. (2022). *Evaluation of snow depth retrievals from ICESat-2 using airborne laser-scanning data* [Preprint]. Snow/Remote Sensing. <https://tc.copernicus.org/preprints/tc-2022-191/>
- 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(13), 4094–4134. <https://doi.org/10.1080/01431161.2011.640964>
- DVORNIKOV, Y. (2015). GIS and field data based modelling of snow water equivalent in shrub tundra. *Fennia – International Journal of Geography*. <https://doi.org/10.11143/46363>
- Enderlin, E. M., Elkin, C. M., Gendreau, M., Marshall, H. P., O’Neil, S., McNeil, C., Florentine, C., & Sass, L. (2022). Uncertainty of ICESat-2 ATL06- and ATL08-derived snow depths for glacierized and vegetated mountain regions. *Remote Sensing of Environment*, 283, 113307. <https://doi.org/10.1016/j.rse.2022.113307>
- Estilow, T. W., Young, A. H., & Robinson, D. A. (2015). A long-term Northern Hemisphere snow cover extent data record for climate studies and monitoring. *Earth System Science Data*, 7(1), 137–142. <https://doi.org/10.5194/essd-7-137-2015>
- European Space Agency, & Sinergise. (2021). *Copernicus Global Digital Elevation Model* [Data set]. <https://doi.org/10.5069/G9028PQB>
- Fiddes, J., Aalstad, K., & Lehning, M. (2022). TopoCLIM: Rapid topography-based downscaling of regional climate model output in complex terrain v1.1. *Geoscientific Model Development*, 15(4), 1753–1768. <https://doi.org/10.5194/gmd-15-1753-2022>
- Fiddes, J., & Gruber, S. (2014). TopoSCALE v.1.0: Downscaling gridded climate data in complex terrain. *Geoscientific Model Development*, 7(1), 387–405. <https://doi.org/10.5194/gmd-7-387-2014>
- Fowler, H. J., Blenkinsop, S., & Tebaldi, C. (2007). Linking climate change modelling to impacts studies: Recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology*, 27(12), 1547–1578. <https://doi.org/10.1002/joc.1556>
- Fox-Kemper, B., Hewitt, H. T., Xiao, C., Aðalgeirsdóttir, G., Drijfhout, S. S., Edwards, T. L., Golledge, N. R., Hemer, M., Kopp, R. E., Krinner, G., Mix, A., Notz, D., Nowicki, S., Nurhati, I. S., Ruiz, L., Sallée, J.-B., Slanger, A. B. A., & Yu, Y. (2021). Chapter 9: Ocean, cryosphere, and sea level change. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, 1211–1361. <https://doi.org/10.1017/9781009157896.011>
- Freudiger, D., Kohn, I., Seibert, J., Stahl, K., & Weiler, M. (2017). Snow redistribution for the hydrological modeling of alpine catchments. *Wiley Interdisciplinary Reviews: Water*, 4(5), e1232. <https://doi.org/10.1002/wat2.1232>
- Girod, L., Nuth, C., Kääb, A., McNabb, R., & Galland, O. (2017). MMASTER: Improved ASTER DEMs for Elevation Change Monitoring. *Remote Sensing*, 9(7), 704. <https://doi.org/10.3390/rs9070704>
- Giroto, M., Musselman, K. N., & Essery, R. L. H. (2020). Data Assimilation Improves Estimates of Climate-Sensitive Seasonal Snow. *Current Climate Change Reports*, 6(3), 81–94.

6. Concluding Remarks and Future Outlooks

- <https://doi.org/10.1007/s40641-020-00159-7>
- Gisnås, K., Westermann, S., Schuler, T. V., Melvold, K., & Etzelmüller, B. (2016). Small-scale variation of snow in a regional permafrost model. *The Cryosphere*, 10(3), 1201–1215. <https://doi.org/10.5194/tc-10-1201-2016>
- Gruber, A., Wessel, B., Huber, M., & Roth, A. (2012). Operational TanDEM-X DEM calibration and first validation results. *ISPRS Journal of Photogrammetry and Remote Sensing*, 73, 39–49. <https://doi.org/10.1016/j.isprsjprs.2012.06.002>
- Gruen, A., & Akca, D. (2005). Least squares 3D surface and curve matching. *ISPRS Journal of Photogrammetry and Remote Sensing*, 59(3), 151–174. <https://doi.org/10.1016/j.isprsjprs.2005.02.006>
- Grünwald, T., Stötter, J., Pomeroy, J. W., Dadic, R., Moreno Baños, I., Marturià, J., Spross, M., Hopkinson, C., Burlando, P., & Lehning, M. (2013). Statistical modelling of the snow depth distribution in open alpine terrain. *Hydrology and Earth System Sciences*, 17(8), 3005–3021. <https://doi.org/10.5194/hess-17-3005-2013>
- Gulev, S. K., Thorne, P. W., Ahn, J., Dentener, F. J., Domingues, C. M., Gerland, S., Gong, D., Kaufman, D. S., Nnamchi, H. C., Quaas, J., Rivera, J. A., Sathyendranath, S., Smith, S. L., Trewin, B., von Schuckmann, K., & Vose, R. S. (2021). *Chapter 2 Changing state of the climate system* (V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, & B. Zhou, Eds.; pp. 287–422). Cambridge University Press. <https://doi.org/10.1017/9781009157896.004>.
- Hall, D. K., Riggs, G. A., & Salomonson, V. V. (2006). MODIS Snow and Sea Ice Products. *Earth Science Satellite Remote Sensing*, 154–181. https://doi.org/10.1007/978-3-540-37293-6_9
- Harpold, A. A., & Molotch, N. P. (2015). Sensitivity of soil water availability to changing snowmelt timing in the western U.S. *Geophysical Research Letters*, 42(19), 8011–8020. <https://doi.org/10.1002/2015GL065855>
- Hawker, L., Uhe, P., Paulo, L., Sosa, J., Savage, J., Sampson, C., & Neal, J. (2022). A 30 m global map of elevation with forests and buildings removed. *Environmental Research Letters*, 17(2), 024016. <https://doi.org/10.1088/1748-9326/ac4d4f>
- Henderson, G. R., Peings, Y., Furtado, J. C., & Kushner, P. J. (2018). Snow–atmosphere coupling in the Northern Hemisphere. *Nature Climate Change*, 8(11), 954–963. <https://doi.org/10.1038/s41558-018-0295-6>
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., ... Thépaut, J. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049. <https://doi.org/10.1002/qj.3803>
- Hobeichi, S., Nishant, N., Shao, Y., Abramowitz, G., Pitman, A., Sherwood, S., Bishop, C., & Green, S.

- (2023). Using Machine Learning to Cut the Cost of Dynamical Downscaling. *Earth's Future*, 11(3), e2022EF003291. <https://doi.org/10.1029/2022EF003291>
- Hock, R., Rasul, G., Adler, C., Cáceres, B., Gruber, S., Hirabayashi, Y., Jackson, M., Kääb, A., Kang, S., Kutuzov, S., Noetzli, J., Rixen, C., & Zhang, Y. (2019). Chapter 2: High mountain areas. In *The Ocean and Cryosphere in a Changing Climate: Special Report of the Intergovernmental Panel on Climate Change* (1st ed., pp. 131–202). Cambridge University Press. <https://doi.org/10.1017/9781009157964.004>.
- 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(4), 398–406. <https://doi.org/10.1016/j.isprsjprs.2009.02.003>
- Hu, X., Hao, X., Wang, J., Huang, G., Li, H., & Yang, Q. (2022). Can the Depth of Seasonal Snow be Estimated From ICESat-2 Products: A Case Investigation in Altay, Northwest China. *IEEE Geoscience and Remote Sensing Letters*, 19, 1–5. <https://doi.org/10.1109/LGRS.2021.3078805>
- Hugonnet, R., Brun, F., Berthier, E., Dehecq, A., Mannerfelt, E. S., Eckert, N., & Farinotti, D. (2022). Uncertainty analysis of digital elevation models by spatial inference from stable terrain. 17. <https://doi.org/10.1109/JSTARS.2022.3188922>.
- Hugonnet, R., McNabb, R., Berthier, E., Menounos, B., Nuth, C., Girod, L., Farinotti, D., Huss, M., Dussaillant, I., Brun, F., & Kääb, A. (2021). Accelerated global glacier mass loss in the early twenty-first century. *Nature*, 592(7856), 726–731. <https://doi.org/10.1038/s41586-021-03436-z>
- Immerzeel, W. W., Lutz, A. F., Andrade, M., Bahl, A., Biemans, H., Bolch, T., Hyde, S., Brumby, S., Davies, B. J., Elmore, A. C., Emmer, A., Feng, M., Fernández, A., Haritashya, U., Kargel, J. S., Koppes, M., Kraaijenbrink, P. D. A., Kulkarni, A. V., Mayewski, P. A., ... Baillie, J. E. M. (2020). Importance and vulnerability of the world's water towers. *Nature*, 577(7790), 364–369. <https://doi.org/10.1038/s41586-019-1822-y>
- Kaspi, Y., & Schneider, T. (2011). Winter cold of eastern continental boundaries induced by warm ocean waters. *Nature*, 471(7340), 621–624. <https://doi.org/10.1038/nature09924>
- Keil, P., Mauritsen, T., Jungclaus, J., Hedemann, C., Olonscheck, D., & Ghosh, R. (2020). Multiple drivers of the North Atlantic warming hole. *Nature Climate Change*, 10(7), 667–671. <https://doi.org/10.1038/s41558-020-0819-8>
- Ketzler, G., Römer, W., & Beylich, A. A. (2021). The Climate of Norway. In A. A. Beylich (Ed.), *Landscapes and Landforms of Norway* (pp. 7–29). Springer International Publishing. https://doi.org/10.1007/978-3-030-52563-7_2
- Kimerling, A. J., Buckley, A. R., Muehrcke, P. C., & Muehrcke, J. O. (2016). *Map Use: Reading, Analysis, Interpretation*. Esri Press Academic. <https://books.google.com?id=22HPjwEACAAJ>
- Kingsford, C., & Salzberg, S. L. (2008). What are decision trees? *Nature Biotechnology*, 26(9), 1011–1013. <https://doi.org/10.1038/nbt0908-1011>
- Klehmet, K., Geyer, B., & Rockel, B. (2013). A regional climate model hindcast for Siberia: Analysis of snow water equivalent. *The Cryosphere*, 7(4), 1017–1034. <https://doi.org/10.5194/tc-7-1017-2013>

6. Concluding Remarks and Future Outlooks

2013

- Kraaijenbrink, P. D. A., Stigter, E. E., Yao, T., & Immerzeel, W. W. (2021). Climate change decisive for Asia's snow meltwater supply. *Nature Climate Change*, 11(7), 591–597. <https://doi.org/10.1038/s41558-021-01074-x>
- Kumar, P., Saharwardi, Md. S., Banerjee, A., Azam, Mohd. F., Dubey, A. K., & Murtugudde, R. (2019). Snowfall Variability Dictates Glacier Mass Balance Variability in Himalaya-Karakoram. *Scientific Reports*, 9(1), 1–9. <https://doi.org/10.1038/s41598-019-54553-9>
- Kunkel, K. E., Robinson, D. A., Champion, S., Yin, X., Estilow, T., & Frankson, R. M. (2016). Trends and Extremes in Northern Hemisphere Snow Characteristics. *Current Climate Change Reports*, 2(2), 65–73. <https://doi.org/10.1007/s40641-016-0036-8>
- Kwok, R., & Cunningham, G. F. (2008). ICESat over Arctic sea ice: Estimation of snow depth and ice thickness. *Journal of Geophysical Research*, 113(C8). <https://doi.org/10.1029/2008JC004753>
- Laloyaux, P., Balmaseda, M., Dee, D., Mogensen, K., & Janssen, P. (2016). A coupled data assimilation system for climate reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 142(694), 65–78. <https://doi.org/10.1002/qj.2629>
- Li, G., Wang, Z. S., & Huang, N. (2018). A Snow Distribution Model Based on Snowfall and Snow Drifting Simulations in Mountain Area. *Journal of Geophysical Research: Atmospheres*, 123(14), 7193–7203. <https://doi.org/10.1029/2018JD028434>
- Li, H., Deng, Q., & Wang, L. (2017). Automatic Co-Registration of Digital Elevation Models Based on Centroids of Subwatersheds. *IEEE Transactions on Geoscience and Remote Sensing*, 55(11), 6639–6650. <https://doi.org/10.1109/TGRS.2017.2731048>
- Li, H., Zhao, J., Yan, B., Yue, L., & Wang, L. (2022). Global DEMs vary from one to another: An evaluation of newly released Copernicus, NASA and AW3D30 DEM on selected terrains of China using ICESat-2 altimetry data. *International Journal of Digital Earth*, 15(1), 1149–1168. <https://doi.org/10.1080/17538947.2022.2094002>
- Li, T., Hu, Y., Liu, B., Jiang, L., Wang, H., & Shen, X. (2022). Co-registration and residual correction of digital elevation models: A comparative study. *The Cryosphere Discussions*, 1–22. <https://doi.org/10.5194/tc-2022-205>
- Lievens, H., Brangers, I., Marshall, H.-P., Jonas, T., Olefs, M., & De Lannoy, G. (2022). Sentinel-1 snow depth retrieval at sub-kilometer resolution over the European Alps. *The Cryosphere*, 16(1), 159–177. <https://doi.org/10.5194/tc-16-159-2022>
- Lievens, H., Demuzere, M., Marshall, H.-P., Reichle, R. H., Brucker, L., Brangers, I., de Rosnay, P., Dumont, M., Girotto, M., Immerzeel, W. W., Jonas, T., Kim, E. J., Koch, I., Marty, C., Saloranta, T., Schöber, J., & De Lannoy, G. J. M. (2019). Snow depth variability in the Northern Hemisphere mountains observed from space. *Nature Communications*, 10(1), 4629. <https://doi.org/10.1038/s41467-019-12566-y>
- Liu, A. (2021). Performance evaluation of GEDI and ICESat-2 laser altimeter data for terrain and canopy height retrievals. *Remote Sensing of Environment*, 16.

- Liu, X. (2008). Airborne LiDAR for DEM generation: Some critical issues. *Progress in Physical Geography: Earth and Environment*, 32(1), 31–49. <https://doi.org/10.1177/0309133308089496>
- Livneh, B., & Badger, A. M. (2020). Drought less predictable under declining future snowpack. *Nature Climate Change*, 10(5), 452–458. <https://doi.org/10.1038/s41558-020-0754-8>
- LuoJus, K., Pullainen, J., Takala, M., Lemmetyinen, J., Mortimer, C., Derksen, C., Mudryk, L., Moisander, M., Hiltunen, M., Smolander, T., Ikonen, J., Cohen, J., Salminen, M., Norberg, J., Veijola, K., & Venäläinen, P. (2021). GlobSnow v3.0 Northern Hemisphere snow water equivalent dataset. *Scientific Data*, 8(1), 163. <https://doi.org/10.1038/s41597-021-00939-2>
- Lussana, C. (2018). *seNorge_2018 daily mean temperature 1957-2017* [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.2023997>
- Macander, M. J., Swingley, C. S., Joly, K., & Reynolds, M. K. (2015). Landsat-based snow persistence map for northwest Alaska. *Remote Sensing of Environment*, 163, 23–31. <https://doi.org/10.1016/j.rse.2015.02.028>
- Magruder, L., Neuenschwander, A., & Klotz, B. (2021). Digital terrain model elevation corrections using space-based imagery and ICESat-2 laser altimetry. *Remote Sensing of Environment*, 264, 112621. <https://doi.org/10.1016/j.rse.2021.112621>
- Maraun, D., & Widmann, M. (2018). *Statistical Downscaling and Bias Correction for Climate Research*. 10.1017/9781107588783
- Markus, T., Neumann, T., Martino, A., Abdalati, W., Brunt, K., Csatho, B., Farrell, S., Fricker, H., Gardner, A., Harding, D., Jasinski, M., Kwok, R., Magruder, L., Lubin, D., Luthcke, S., Morison, J., Nelson, R., Neuenschwander, A., Palm, S., ... Zwally, J. (2017). The Ice, Cloud, and land Elevation Satellite-2 (ICESat-2): Science requirements, concept, and implementation. *Remote Sensing of Environment*, 190, 260–273. <https://doi.org/10.1016/j.rse.2016.12.029>
- Matiu, M., Crespi, A., Bertoldi, G., Carmagnola, C. M., Marty, C., Morin, S., Schöner, W., Cat Berro, D., Chiogna, G., De Gregorio, L., Kotlarski, S., Majone, B., Resch, G., Terzaghi, S., Valt, M., Beozzo, W., Cianfarra, P., Gouttevin, I., Marcolini, G., ... Weilguni, V. (2021). Observed snow depth trends in the European Alps: 1971 to 2019. *The Cryosphere*, 15(3), 1343–1382. <https://doi.org/10.5194/tc-15-1343-2021>
- Maune, D. (2007). *Digital elevation model technologies and applications: The DEM users manual*.
- Melvold, K., & Skaugen, T. (2013). Multiscale spatial variability of lidar-derived and modeled snow depth on Hardangervidda, Norway. *Annals of Glaciology*, 54(62), 273–281. <https://doi.org/10.3189/2013AoG62A161>
- Melvold, K., & Skaugen, T. (2019). *Snow depth on Hardangervidda, South Norway 2008-2009* [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.2572732>
- Meredith, M., Sommerkorn, M., Cassotta, S., Derksen, C., Ekaykin, A., Hollowed, A., Kofinas, G., Mackintosh, A., Melbourne-Thomas, J., Muelbert, M., Ottersen, G., Pritchard, H., & A. G., S. (2019). Chapter 3: Polar Regions. IPCC Special Report on the Ocean and Cryosphere in a Changing Climate. In *The Ocean and Cryosphere in a Changing Climate: Special Report of the Intergovernmental Panel on Cli-*

6. Concluding Remarks and Future Outlooks

- mate Change*. Cambridge University Press. <https://www.cambridge.org/core/product/identifier/9781009157964/type/book>
- Mesa-Mingorance, J. L., & Ariza-López, F. J. (2020). Accuracy Assessment of Digital Elevation Models (DEMs): A Critical Review of Practices of the Past Three Decades. *Remote Sensing*, 12(16), 2630. <https://doi.org/10.3390/rs12162630>
- Mo, H. M., Dai, L. Y., Fan, F., Che, T., & Hong, H. P. (2016). Extreme snow hazard and ground snow load for China. *Natural Hazards*, 84(3), 2095–2120. <https://doi.org/10.1007/s11069-016-2536-1>
- Morin, P., Porter, C., Cloutier, M., Howat, I., Noh, M.-J., Willis, M., Bates, B., Williamson, C., & Peterman, K. (2016). *ArcticDEM; A Publicly Available, High Resolution Elevation Model of the Arctic*. EPSC2016-8396. <https://ui.adsabs.harvard.edu/abs/2016EGUGA..18.8396M>
- Mott, R., Vionnet, V., & Grunewald, T. (2018). The Seasonal Snow Cover Dynamics: Review on Wind-Driven Coupling Processes. *Frontiers in Earth Science*, 6. <https://doi.org/10.3389/feart.2018.00197>
- Moudrý, V., Gdulová, K., Gábor, L., Šárovcová, E., Barták, V., Leroy, F., Špatenková, O., Rocchini, D., & Prošek, J. (2022). Effects of environmental conditions on ICESat-2 terrain and canopy heights retrievals in Central European mountains. *Remote Sensing of Environment*, 279, 113112. <https://doi.org/10.1016/j.rse.2022.113112>
- Mudryk, L. R., Derksen, C., Kushner, P. J., & Brown, R. (2015). Characterization of Northern Hemisphere Snow Water Equivalent Datasets, 1981–2010. *Journal of Climate*, 28(20), 8037–8051. <https://doi.org/10.1175/JCLI-D-15-0229.1>
- Mudryk, L., Santolaria-Otín, M., Krinner, G., Ménégoz, M., Derksen, C., Brutel-Vuilmet, C., Brady, M., & Essery, R. (2020). Historical Northern Hemisphere snow cover trends and projected changes in the CMIP6 multi-model ensemble. *The Cryosphere*, 14(7), 2495–2514. <https://doi.org/10.5194/tc-14-2495-2020>
- Muñoz Sabater, J. (2021a). *ERA5-Land hourly data from 1950 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. 10.24381/cds.e2161bac
- Muñoz Sabater, J. (2021b). *ERA5-Land monthly data from 1950 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. 10.24381/cds.68d2bb30
- Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., & Thépaut, J.-N. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, 13(9), 4349–4383. <https://doi.org/10.5194/essd-13-4349-2021>
- Musselman, K. N., Addor, N., Vano, J. A., & Molotch, N. P. (2021). Winter melt trends portend widespread declines in snow water resources. *Nature Climate Change*, 11(5), 418–424. <https://doi.org/10.1038/s41558-021-01014-9>
- Myers-Smith, I. H., Kerby, J. T., Phoenix, G. K., Bjerke, J. W., Epstein, H. E., Assmann, J. J., John, C., Andreu-Hayles, L., Angers-Blondin, S., Beck, P. S. A., Berner, L. T., Bhatt, U. S., Bjorkman, A. D.,

- Blok, D., Bryn, A., Christiansen, C. T., Cornelissen, J. H. C., Cunliffe, A. M., Elmendorf, S. C., ... Wipf, S. (2020). Complexity revealed in the greening of the Arctic. *Nature Climate Change*, 10(2), 106–117. <https://doi.org/10.1038/S41558-019-0688-1>
- Neuenschwander, A. L., & Sheridan, R. (2021). *ATLAS/ICESat-2 L3A Land and Vegetation Height, version 5* [Data set]. NASA National Snow and Ice Data Center Distributed Active Archive Center. <https://doi.org/10.5067/ATLAS/ATL08.005>
- Neuenschwander, A., & Magruder, L. (2016). The Potential Impact of Vertical Sampling Uncertainty on ICESat-2/ATLAS Terrain and Canopy Height Retrievals for Multiple Ecosystems. *Remote Sensing*, 8(12), 1039. <https://doi.org/10.3390/rs8121039>
- Neuenschwander, A., & Pitts, K. (2019). The ATL08 land and vegetation product for the ICESat-2 Mission. *Remote Sensing of Environment*, 221, 247–259. <https://doi.org/10.1016/j.rse.2018.11.005>
- Neuenschwander, A., Pitts, K., Jelley, B., Robbins, J., Markel, J., Popescu, S., Nelson, R., Harding, D., Klotz, B., & Sheridan, R. (2022). *Ice, Cloud, and Land Elevation Satellite 2 (ICESat-2) algorithm theoretical basis documents (ATBD) for Land-Vegetation along-track products (ATL08)*.
- Neumann, T., Brenner, A., Hancock, D., Robbins, J., Saba, J., Harbeck, K., Gibbons, A., Lee, J., Luthcke, S., & Rebeld, T. (2021). *Ice, Cloud, and Land Elevation Satellite - 2 (ICESat-2) Project Algorithm Theoretical Basis Document (ATBD) for Global Geolocated Photons ATL03*. NASA Goddard Space Flight Center. <https://doi.org/10.5067/D0VQW9KX3GTU>
- Nolin, A. W. (2010). Recent advances in remote sensing of seasonal snow. *Journal of Glaciology*, 56(200), 1141–1150. <https://doi.org/10.3189/002214311796406077>
- Nuth, C., & Kääb, A. (2011). Co-registration and bias corrections of satellite elevation data sets for quantifying glacier thickness change. *The Cryosphere*, 5(1), 271–290. <https://doi.org/10.5194/tc-5-271-2011>
- Oksanen, J., & Sarjakoski, T. (2006). Uncovering the statistical and spatial characteristics of fine toposcale DEM error. *International Journal of Geographical Information Science*, 20(4), 345–369. <https://doi.org/10.1080/13658810500433891>
- Olvmo, M., Holmer, B., Thorsson, S., Reese, H., & Lindberg, F. (2020). Sub-arctic palsa degradation and the role of climatic drivers in the largest coherent palsa mire complex in Sweden (Vissátvuopmi), 1955–2016. *Scientific Reports*, 10(1), 8937. <https://doi.org/10.1038/s41598-020-65719-1>
- Orsolini, Y., Wegmann, M., Dutra, E., Liu, B., Balsamo, G., Yang, K., de Rosnay, P., Zhu, C., Wang, W., Senan, R., & Arduini, G. (2019). Evaluation of snow depth and snow cover over the Tibetan Plateau in global reanalyses using in situ and satellite remote sensing observations. *The Cryosphere*, 13(8), 2221–2239. <https://doi.org/10.5194/tc-13-2221-2019>
- Palmer, S., McMillan, M., & Morlighem, M. (2015). Subglacial lake drainage detected beneath the Greenland ice sheet. *Nature Communications*, 6(1), 8408. <https://doi.org/10.1038/ncomms9408>
- Paul, F., Bolch, T., Kääb, A., Nagler, T., Nuth, C., Scharrer, K., Shepherd, A., Strozzi, T., Ticconi, F., Bambiri, R., Berthier, E., Bevan, S., Gourmelen, N., Heid, T., Jeong, S., Kunz, M., Lauknes, T. R., Luckman, A., Merryman Boncori, J. P., ... Van Niel, T. (2015). The glaciers climate change initiative: Methods

6. Concluding Remarks and Future Outlooks

- for creating glacier area, elevation change and velocity products. *Remote Sensing of Environment*, 162, 408–426. <https://doi.org/10.1016/j.rse.2013.07.043>
- Peckham, S. (2011, May 30). *Profile, Plan and Streamline Curvature: A Simple Derivation and Applications*.
- PENG, S., PIAO, S., CIAIS, P., FANG, J., & WANG, X. (2010). Change in winter snow depth and its impacts on vegetation in China. *Global Change Biology*, 16(11), no-no. <https://doi.org/10.1111/j.1365-2486.2010.02210.x>
- Perovich, D. K., Light, B., Eicken, H., Jones, K. F., Runciman, K., & Nghiem, S. V. (2007). Increasing solar heating of the Arctic Ocean and adjacent seas, 1979–2005: Attribution and role in the ice-albedo feedback. *Geophysical Research Letters*, 34(19). <https://doi.org/10.1029/2007GL031480>
- Pflug, J. M., Hughes, M., & Lundquist, J. D. (2021). Downscaling Snow Deposition Using Historic Snow Depth Patterns: Diagnosing Limitations From Snowfall Biases, Winter Snow Losses, and Interannual Snow Pattern Repeatability. *Water Resources Research*, 57(8), e2021WR029999. <https://doi.org/10.1029/2021WR029999>
- Pritchard, H. D., Ligtenberg, S. R. M., Fricker, H. A., Vaughan, D. G., van den Broeke, M. R., & Padman, L. (2012). Antarctic ice-sheet loss driven by basal melting of ice shelves. *Nature*, 484(7395), 502–505. <https://doi.org/10.1038/nature10968>
- Qin, Y., Abatzoglou, J. T., Siebert, S., Huning, L. S., AghaKouchak, A., Mankin, J. S., Hong, C., Tong, D., Davis, S. J., & Mueller, N. D. (2020). Agricultural risks from changing snowmelt. *Nature Climate Change*, 10(5), 459–465. <https://doi.org/10.1038/s41558-020-0746-8>
- Quante, L., Willner, S. N., Middelanis, R., & Levermann, A. (2021). Regions of intensification of extreme snowfall under future warming. *Scientific Reports*, 11(1), 16621. <https://doi.org/10.1038/s41598-021-95979-4>
- Rantanen, M., Karpechko, A. Yu., Lipponen, A., Nordling, K., Hyvärinen, O., Ruosteenoja, K., Vihma, T., & Laaksonen, A. (2022). The Arctic has warmed nearly four times faster than the globe since 1979. *Communications Earth & Environment*, 3(1, 1), 1–10. <https://doi.org/10.1038/s43247-022-00498-3>
- Rashmi, K. V., & Gilad-Bachrach, R. (2015, May 7). *DART: Dropouts meet Multiple Additive Regression Trees*. <https://doi.org/10.48550/arXiv.1505.01866>
- Rasmussen, R., Baker, B., Kochendorfer, J., Meyers, T., Landolt, S., Fischer, A. P., Black, J., Thériault, J. M., Kucera, P., Gochis, D., Smith, C., Nitu, R., Hall, M., Ikeda, K., & Gutmann, E. (2012). How Well Are We Measuring Snow: The NOAA/FAA/NCAR Winter Precipitation Test Bed. *Bulletin of the American Meteorological Society*, 93(6), 811–829. <https://doi.org/10.1175/BAMS-D-11-00052.1>
- Réveillet, M., Dumont, M., Gascoin, S., Lafaysse, M., Nabat, P., Ribes, A., Nheili, R., Tuzet, F., Ménégoz, M., Morin, S., Picard, G., & Ginoux, P. (2022). Black carbon and dust alter the response of mountain snow cover under climate change. *Nature Communications*, 13(1), 5279. <https://doi.org/10.1038/s41467-022-32501-y>
- Revuelto, J., López-Moreno, J. I., Azorin-Molina, C., & Vicente-Serrano, S. M. (2014). Topographic con-

- trol of snowpack distribution in a small catchment in the central Spanish Pyrenees: Intra- and inter-annual persistence. *The Cryosphere*, 8(5), 1989–2006. <https://doi.org/10.5194/tc-8-1989-2014>
- Rhoades, A. M., Hatchett, B. J., Risser, M. D., Collins, W. D., Bambach, N. E., Huning, L. S., McCrary, R., Siirila-Woodburn, E. R., Ullrich, P. A., Wehner, M. F., Zarzycki, C. M., & Jones, A. D. (2022). Asymmetric emergence of low-to-no snow in the midlatitudes of the American Cordillera. *Nature Climate Change*, 12(12), 1151–1159. <https://doi.org/10.1038/s41558-022-01518-y>
- Riihelä, A., Bright, R. M., & Anttila, K. (2021). Recent strengthening of snow and ice albedo feedback driven by Antarctic sea-ice loss. *Nature Geoscience*, 14(11), 832–836. <https://doi.org/10.1038/s41561-021-00841-x>
- Rosen, P. A., Hensley, S., Joughin, I. R., Li, F. K., Madsen, S. N., Rodriguez, E., & Goldstein, R. M. (2000). Synthetic aperture radar interferometry. *Proceedings of the IEEE*, 88(3), 333–382. <https://doi.org/10.1109/5.838084>
- Rott, H., & Nagler, T. (1993). *Capabilities of ERS-1 SAR for snow and glacier monitoring in Alpine areas*.
- Sachindra, D. A., Ahmed, K., Rashid, Md. M., Shahid, S., & Perera, B. J. C. (2018). Statistical downscaling of precipitation using machine learning techniques. *Atmospheric Research*, 212, 240–258. <https://doi.org/10.1016/j.atmosres.2018.05.022>
- Schutz, B. E., Zwally, H. J., Shuman, C. A., Hancock, D., & DiMarzio, J. P. (2005). Overview of the ICESat Mission. *Geophysical Research Letters*, 32(21). <https://doi.org/10.1029/2005GL024009>
- Schweizer, J., Bartelt, P., & van Herwijnen, A. (2021). Snow avalanches. In *Snow and Ice-Related Hazards, Risks, and Disasters* (pp. 377–416). Elsevier. <https://linkinghub.elsevier.com/retrieve/pii/B9780128171295000019>
- Schweizer, J., Bruce Jamieson, J., & Schneebeli, M. (2003). Snow avalanche formation. *Reviews of Geophysics*, 41(4). <https://doi.org/10.1029/2002RG000123>
- Seppälä, M. (1994). Snow depth controls palsu growth. *Permafrost and Periglacial Processes*, 5(4), 283–288. <https://doi.org/10.1002/ppp.3430050407>
- Seppälä, M. (2005). *The physical geography of Fennoscandia*. Oxford University Press.
- Siirila-Woodburn, E. R., Rhoades, A. M., Hatchett, B. J., Huning, L. S., Szinai, J., Tague, C., Nico, P. S., Feldman, D. R., Jones, A. D., Collins, W. D., & Kaatz, L. (2021). A low-to-no snow future and its impacts on water resources in the western United States. *Nature Reviews Earth & Environment*, 2(11), 800–819. <https://doi.org/10.1038/s43017-021-00219-y>
- Smith, B., Fricker, H. A., Holschuh, N., Gardner, A. S., Adusumilli, S., Brunt, K. M., Csatho, B., Harbeck, K., Huth, A., Neumann, T., Nilsson, J., & Siegfried, M. R. (2019). Land ice height-retrieval algorithm for NASA's ICESat-2 photon-counting laser altimeter. *Remote Sensing of Environment*, 233, 111352. <https://doi.org/10.1016/j.rse.2019.111352>
- Spall, J. C. (1998). Implementation of the simultaneous perturbation algorithm for stochastic optimization. *IEEE Transactions on Aerospace and Electronic Systems*, 34(3), 817–823. <https://doi.org/10.1109/7.705889>
- Stiles, W. H., & Ulaby, F. T. (1980). The active and passive microwave response to snow parameters: 1.

6. Concluding Remarks and Future Outlooks

- Wetness. *Journal of Geophysical Research*, 85(C2), 1037. <https://doi.org/10.1029/JC085iC02p0103>
- 7
- Sturm, M., Holmgren, J., McFadden, J. P., Liston, G. E., Chapin, F. S., & Racine, C. H. (2001). Snow-Shrub Interactions in Arctic Tundra: A Hypothesis with Climatic Implications. *Journal of Climate*, 14(3), 336–344. [https://doi.org/10.1175/1520-0442\(2001\)014%3C0336:SSIIAT%3E2.0.CO;2](https://doi.org/10.1175/1520-0442(2001)014%3C0336:SSIIAT%3E2.0.CO;2)
- Sturm, M., & Massom, R. (2016). Snow in the sea ice system: Friend or foe? In *Sea Ice: Third Edition* (pp. 65–109). <https://doi.org/10.1002/9781118778371.ch3>
- Sturm, M., & Wagner, A. M. (2010). Using repeated patterns in snow distribution modeling: An Arctic example. *Water Resources Research*, 46(12). <https://doi.org/10.1029/2010WR009434>
- Tian, X., & Shan, J. (2021). Comprehensive Evaluation of the ICESat-2 ATL08 Terrain Product. *IEEE Transactions on Geoscience and Remote Sensing*, 59(10), 8195–8209. <https://doi.org/10.1109/TGRS.2021.3051086>
- Treichler, D., & Kääb, A. (2017). Snow depth from ICESat laser altimetry – A test study in southern Norway. *Remote Sensing of Environment*, 191, 389–401. <https://doi.org/10.1016/j.rse.2017.01.022>
- Trujillo, E., Ramírez, J. A., & Elder, K. J. (2007). Topographic, meteorologic, and canopy controls on the scaling characteristics of the spatial distribution of snow depth fields. *Water Resources Research*, 43(7). <https://doi.org/10.1029/2006WR005317>
- Vacaflor, P., Lenzano, M. G., Vich, A., & Lenzano, L. (2022). Co-Registration Methods and Error Analysis for Four Decades (1979–2018) of Glacier Elevation Changes in the Southern Patagonian Icefield. *Remote Sensing*, 14(4), 820. <https://doi.org/10.3390/rs14040820>
- Wang, X. (2022). WRF-based dynamical downscaling over High Mountain Asia. <https://depositonce.tu-berlin.de/handle/11303/16159>
- Wang, X., Cheng, X., Gong, P., Huang, H., Li, Z., & Li, X. (2011). Earth science applications of ICESat/GLAS: A review. *International Journal of Remote Sensing*, 32(23), 8837–8864. <https://doi.org/10.1080/01431161.2010.547533>
- Weiss, A. (2001). *Topographic position and landforms analysis*. [Poster presentation]. ESRI user conference, San Diego, CA. http://www.jennessent.com/downloads/TPI-poster-TNC_18x22.pdf
- Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., & Reynolds, J. M. (2012). “Structure-from-Motion” photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*, 179, 300–314. <https://doi.org/10.1016/j.geomorph.2012.08.021>
- Wilby, R. L., & Wigley, T. M. L. (1997). Downscaling general circulation model output: A review of methods and limitations. *Progress in Physical Geography: Earth and Environment*, 21(4), 530–548. <https://doi.org/10.1177/030913339702100403>
- Wipf, S., Stoeckli, V., & Bebi, P. (2009). Winter climate change in alpine tundra: Plant responses to changes in snow depth and snowmelt timing. *Climatic Change*, 94(1-2), 105–121. <https://doi.org/10.1007/S10584-009-9546-X>
- Würzer, S., & Jonas, T. (2018). Spatio-temporal aspects of snowpack runoff formation during rain on snow. *Hydrological Processes*, 32(23), 3434–3445. <https://doi.org/10.1002/hyp.13240>

- Xdem*. (2021). Zenodo. <https://doi.org/10.5281/zenodo.4809698>
- Xiong, L., Li, S., Tang, G., & Strobl, J. (2022). Geomorphometry and terrain analysis: Data, methods, platforms and applications. *Earth-Science Reviews*, 233, 104191. <https://doi.org/10.1016/j.earsci.rev.2022.104191>
- Xu, W., Ma, L., Ma, M., Zhang, H., & Yuan, W. (2017). Spatial-Temporal Variability of Snow Cover and Depth in the Qinghai-Tibetan Plateau. *Journal of Climate*, 30(4), 1521–1533. <https://doi.org/10.1175/JCLI-D-15-0732.1>
- Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J. C., Sampson, C. C., Kanae, S., & Bates, P. D. (2017). A high-accuracy map of global terrain elevations. *Geophysical Research Letters*, 44(11), 5844–5853. <https://doi.org/10.1002/2017GL072874>
- Yueh, S. H., Dinardo, S. J., Akgiray, A., West, R., Cline, D. W., & Elder, K. (2009). Airborne Ku-Band Polarimetric Radar Remote Sensing of Terrestrial Snow Cover. *IEEE Transactions on Geoscience and Remote Sensing*, 47(10), 3347–3364. <https://doi.org/10.1109/TGRS.2009.2022945>
- Zevenbergen, L. W., & Thorne, C. R. (1987). Quantitative analysis of land surface topography. *Earth Surface Processes and Landforms*, 12(1), 47–56. <https://doi.org/10.1002/esp.3290120107>
- Zhu, P., Kim, T., Jin, Z., Lin, C., Wang, X., Ciais, P., Mueller, N. D., Aghakouchak, A., Huang, J., Mulla, D., & Makowski, D. (2022). The critical benefits of snowpack insulation and snowmelt for winter wheat productivity. *Nature Climate Change*, 12(5, 5), 485–490. <https://doi.org/10.1038/s41558-022-01327-3>

A. Appendix

A.1. Additional Figures

A. Appendix

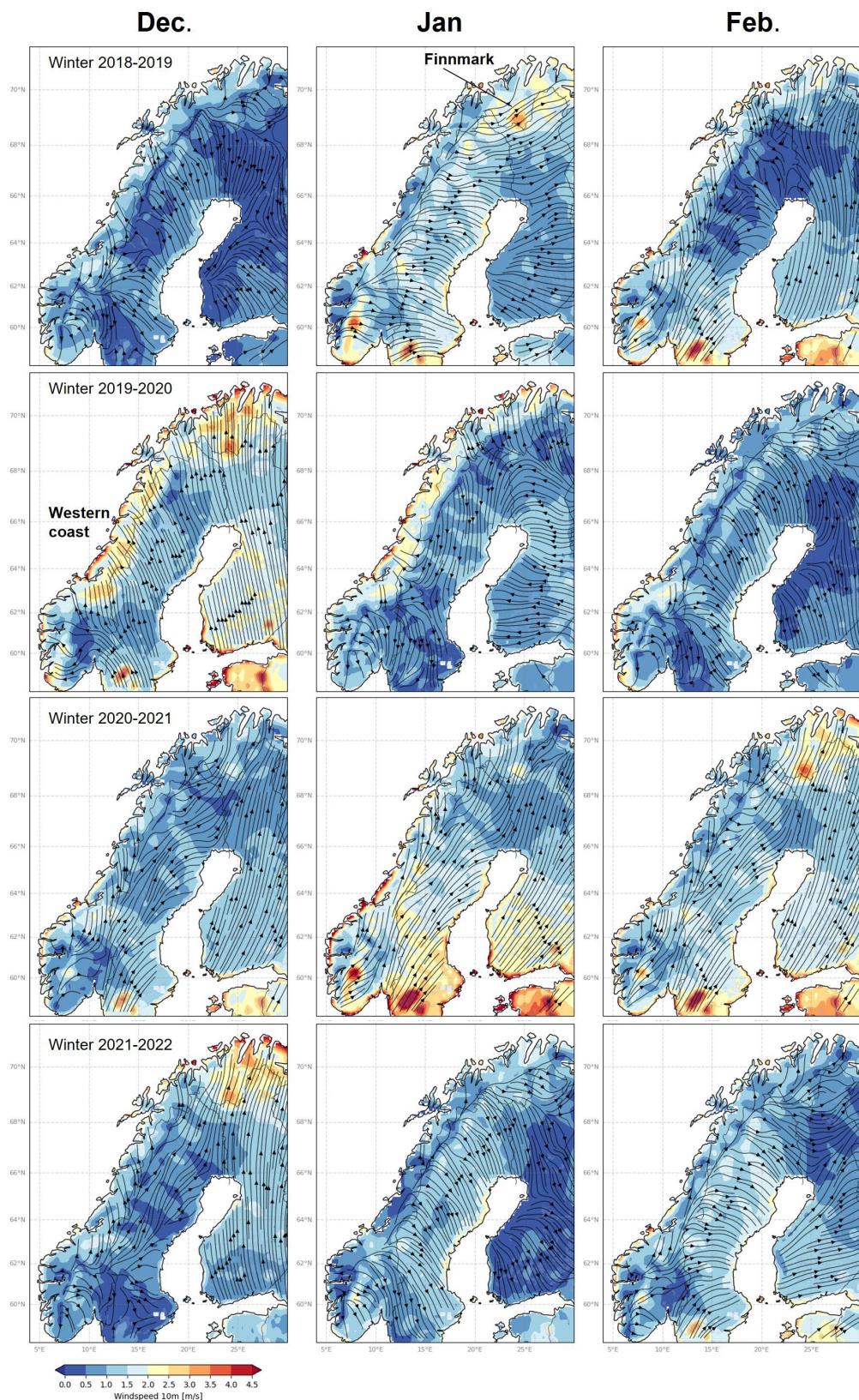


Figure A.1.: The Monthly Average of Wind Speed and Direction in Winter (2019-2022). Data source: ERA5-Land.
96

A.1. Additional Figures

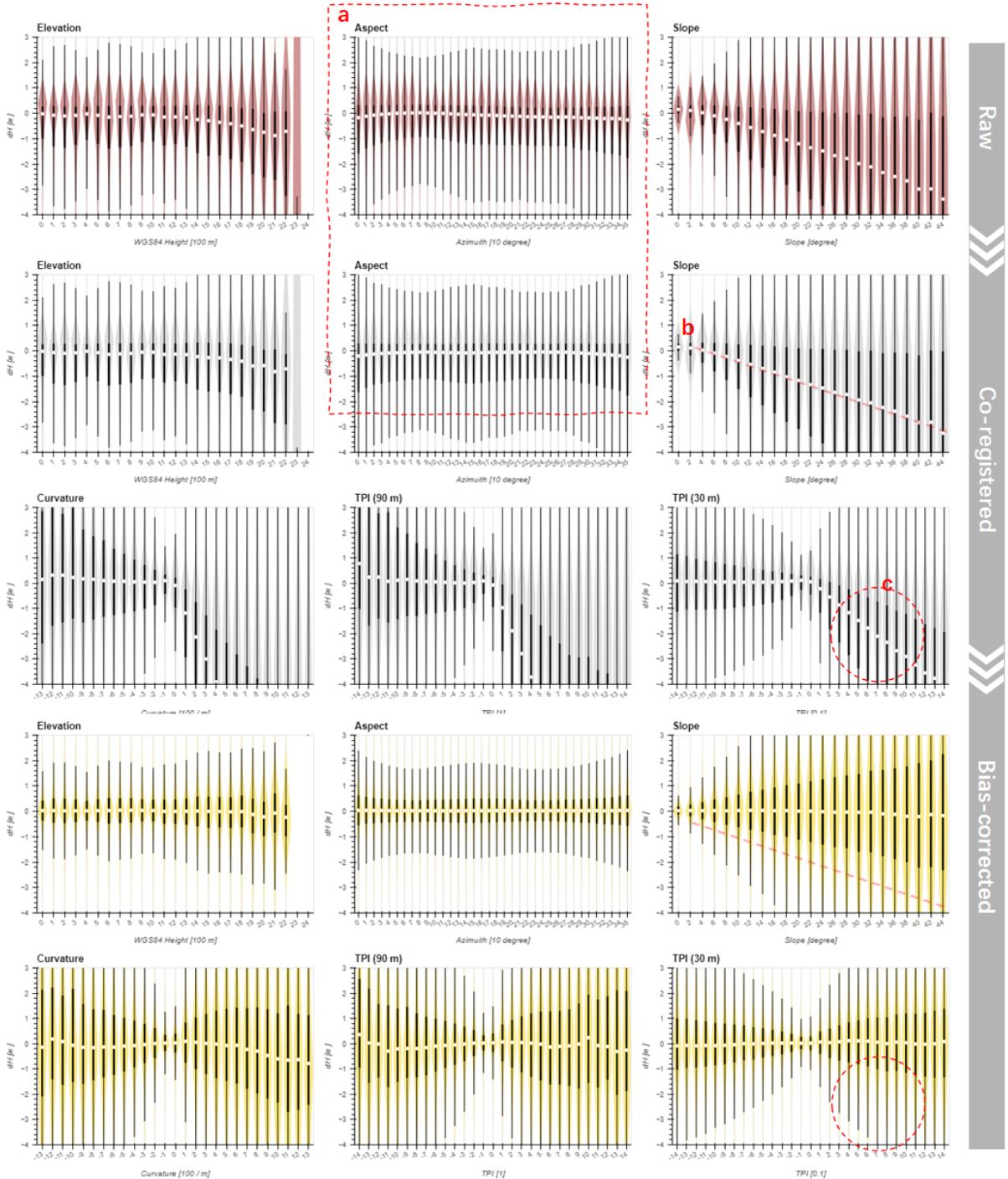


Figure A.2.: Statistic binning Analysis of DEM Errors: Co-registration and Bias correction (ICESat-2 - DTM1). The dataset was divided into categories at three stages, and for each category, a violin plot displays the distribution of data. The median of each dataset is represented by a white dot at the center of the violin. The thick bar illustrates the 25% and 75% quantiles, respectively. If the violin is skewed to one side, it indicates a deviation from normal distribution. The aspect-dependent bias is identified and eliminated by co-registration (a); The negative bias are widespread after co-registration (a, b, c).

A. Appendix

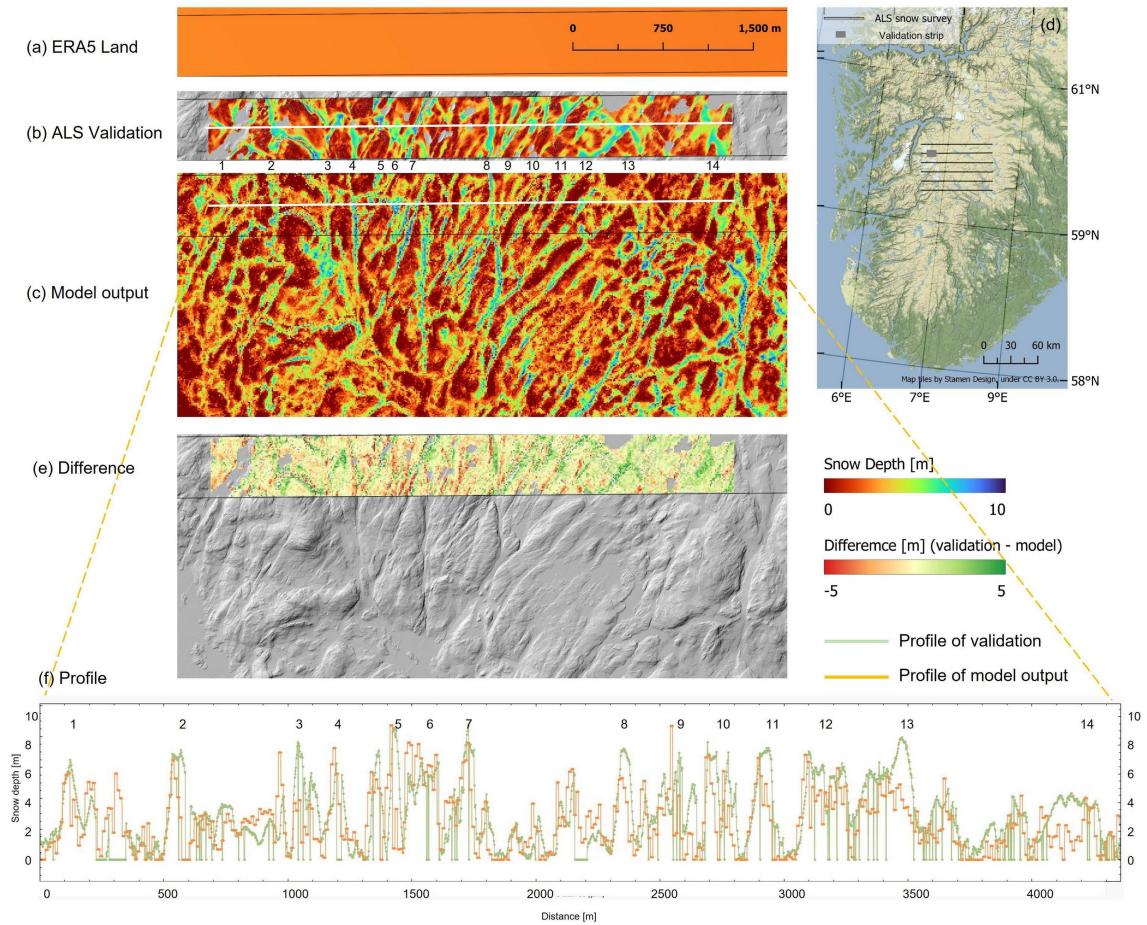


Figure A.3.: The Validation of Snow Depth Downscaling (2009). The snow depth from ERA5 Land that needs to be downscaled shows no variability at this scale (a). The NVE's ALS snow survey conducted in April 2009 provided validation strip, which is located in the northwest of the survey area (d) and shows significant variations across depressions and hilltops. The same strip is also displayed in Melvold & Skaugen (2013)'s Figure 3 (b). The downscaled snow depth from the ICESat 2 - DTM1 model reproduced most of the variability (c). However, there were differences between the model and validation, which are shown in (e), while the profile difference can be seen in (f). The depressions are numbered 1 to 14 from west to east. Note: The validated snow depth is at a resolution of 2 meters and excludes the water surface. The model is trained in a region 100 km wider in all directions than the validation area. And the output is at a resolution of 10 meters.

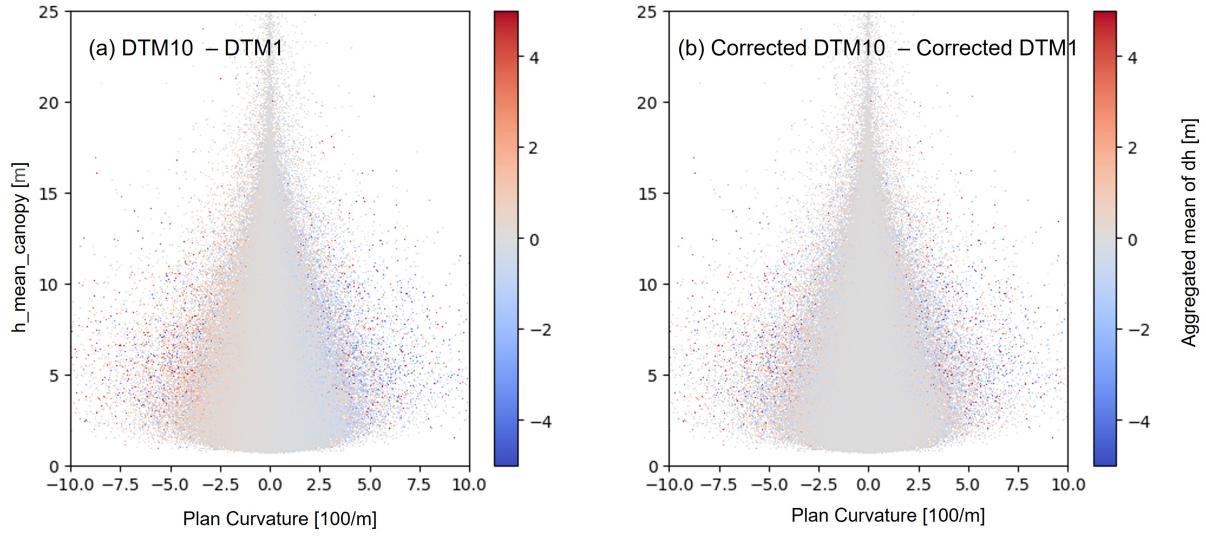


Figure A.4.: Comparison of DTM1 and DTM10 in terms of Plan Curvature and Canopy Height. The resolution limitation of DTM10 produces a positive bias over concave terrain and a negative bias over convex terrain (a). The bias patterns does not keep after bias correction (b).

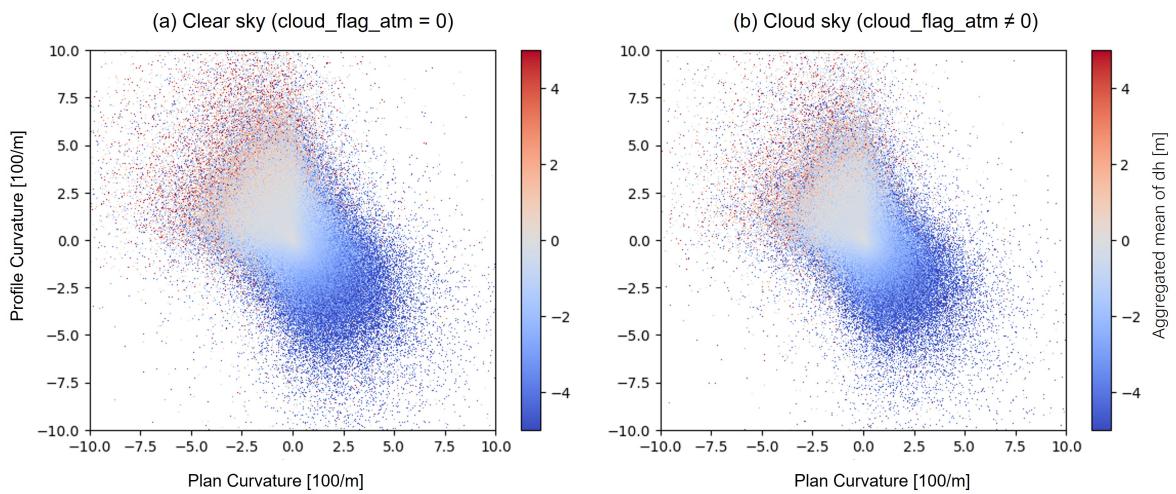


Figure A.5.: Comparison of Clear Sky and Cloudy Sky in Terms of Plan Curvature and Profile Curvature. The presence of cloud result in terrain height underestimation as noted by Moudrý et al. (2022). However, this phenomenon still exhibits in convex terrain under clear sky condition (a, b).

A. Appendix

A.2. Technical Details of ICESat-2 ATL08

A.2.1. Processing and QC Steps of ICESat-2 ATL08 Land products

This is a set of QC instructions for computing segment parameters for the land products. The full version please refer to A. Neuenschwander et al. (2022):

1. The first step is to determine the classed photons for each 100-meter segment, classifying them as ground, canopy, or top of canopy. If there are fewer than 50 classed photons in a segment, do not calculate land or canopy products. If there are 50 or more classed photons in a segment, extract the ground photons to create the land products.
2. Next, if the number of ground photons is greater than 5% of the total number of classed photons within the segment, compute statistics on the ground photons including mean, median, min, max, standard deviation, mode, and skew. These heights will be reported on the product as *h_te_mean*, *h_te_median*, *h_te_min*, *h_te_max*, *h_te_mode*, and *h_te_skew*.
3. Compute the standard deviation of the ground photons about the interpolated terrain surface, FINALGROUND. This value is reported as *h_te_std*.
4. Compute the residuals of the ground photon Z heights about the interpolated terrain surface, FINALGROUND. The product is the root sum of squares of the ground photon residuals combined with the sigma_atlas_land. This parameter is reported as *h_te_uncertainty*.
5. Compute a linear fit on the ground photons and report the slope. This parameter is *terrain_slope*.
6. Calculate a best fit terrain elevation at the mid-point location of the 100-meter segment. Calculate each terrain photon's distance along-track into the 100-meter segment using the corresponding ATL03 20-meter products *segment_length* and *dist_ph_along*, and determine the mid-segment distance (expected to be $50\text{ m} \pm 0.5\text{ m}$). Use the mid-segment distance to linearly interpolate a mid-segment time (delta_time). Use the mid-segment time to linearly interpolate other mid-segment parameters: interpolated terrain surface as *h_te_interp*, *latitude* and *longitude*.
7. Calculate a linear fit as well as 3rd and 4th order polynomial fits to the terrain photons in the segment. Create a slope-adjusted and weighted mid-segment variable, weighted Z, from the linear fit. Use *terrain_slope* to apply a slope correction to each terrain photon by subtracting the terrain photon heights from the linear fit. Determine the mid-segment location of the linear fit, and add that height to the slope-corrected terrain photons. Apply a linear weighting to each photon based on its distance to the mid-segment location: $1/\sqrt{(\text{photon distance along} - \text{mid-segment distance})^2}$. Calculate the weighted mid-segment terrain height, using weighted Z as sum (each adjusted terrain height * its weight) / sum(all weights).
8. Finally, determine which of the three fits is best by calculating the mean and standard deviation

A.2. Technical Details of ICESat-2 ATL08

of the fit errors. If one of the fits has both the smallest mean and standard deviations, use that fit. Else, use the fit with the smallest standard deviation. If more than one fit has the same mean and standard deviation, use the linear fit. This is *h_te_best_fit*.

A. Appendix

A.3. Recommendation For Winter Canopy Height Correction

Vegetation height has a significant impact on the quality of DEM and ground sampling of ICESat-2. It shows high feature importance in the regression model of bias correction (Figure 4.6). However, the bias caused by vegetation height is unbalance across different terrains (refer to Sections 5.2.1 and 5.2.2). It has been observed that if we directly use canopy height from ICESat-2, ignoring the presence of snow cover. It could lead to an underestimation of the correction strength. Therefore, it is crucial to accurately accurately canopy height over snow-on surface.

Based on canopy information from ICESat-2 snow-free segments, the solution used in this study is presented here:

1. Training a binary regression model to determine the presence or absence of the tree.
2. Training a canopy height regression model to estimate the canopy height.
3. Applying the estimated canopy height (2) for snow-on segments according to the result from (1).

A.4. The Consistency of ICESat-2 ATL08 Snow-free Segments

The stability and consistency over different conditions and periods are crucial for a reliable dataset. The data binning can be used to quantify errors and determine if there are any systematic biases. As shown in Figure A.6 the elevation difference between ATL08 snow-free and DTM1 are dividing into smaller groups, where the following factors have been considered under control:

- The background noise is a significant factor affecting bias. The lidar sensor experiences interference from skylight, resulting in better performance at night. The median of the differences between night observations and DTM1 is stable at 0 m, while day observations have a negative median ranging from -0.05 m to -0.1 m, both tending to give negative estimations, indicating by Q1 and Q3 (Figure A.6 a)
- The quantity of photons is crucial as it affects ground-finding algorithms' performance. The Figure fig@quantile_metrics in previous section show that the least 5% of *n_photons* quantiles has biggest NMAD and negative mean error, comparing to conditions with sufficient photons.
- The snow free segments are selected under 'snow free condition' as classified by NOAA daily snow mask products. The *brightness_flag* reveals the presence of small snowpack which is wrongly classified into 'snow free condition'. Figure A.6 c show that the 'snow pack' contributes the positive bias (~ 0.2 m to 0.3 m).
- The pair 1,2,3 exhibit a slightly negative bias of around 0.1 m (Figure A.6 b)
- The interpolation algorithms such as *h_te_interp* or *h_te_bestfit* can affect accuracy. Some studies show that *t_te_best_fit* has slightly higher accuracy and smaller variation than *h_te_interp* (A. Neuenschwander & Pitts, 2019; Tian & Shan, 2021). This study test on *h_te_bestfit* it does make the results better than *h_te_interp* (not shown).
- The strong or weak beams play an important role in signal-to-noise ratios over low reflectance objects (such as sea) or complex geometry (such as dense vegetation) (A. Neuenschwander & Pitts, 2019). However, there is no noticeable bias between strong and weak beams on this snow-free dataset of this study (Figure A.6 a,b,c).
- Moreover, the year 2018 to 2022 appears to be consistency (Figure A.6 a,b,c).

The article discusses the importance of stability and consistency in datasets and how data binning can be used to quantify errors and detect systematic biases. The study focuses on elevation differences between ATL08 snow-free and DTM1, considering factors such as background noise, quantity of photons, snow-free segments, beam strength, and interpolation algorithms. The study finds that the year 2018 to 2022 shows consistency in the dataset.

A. Appendix

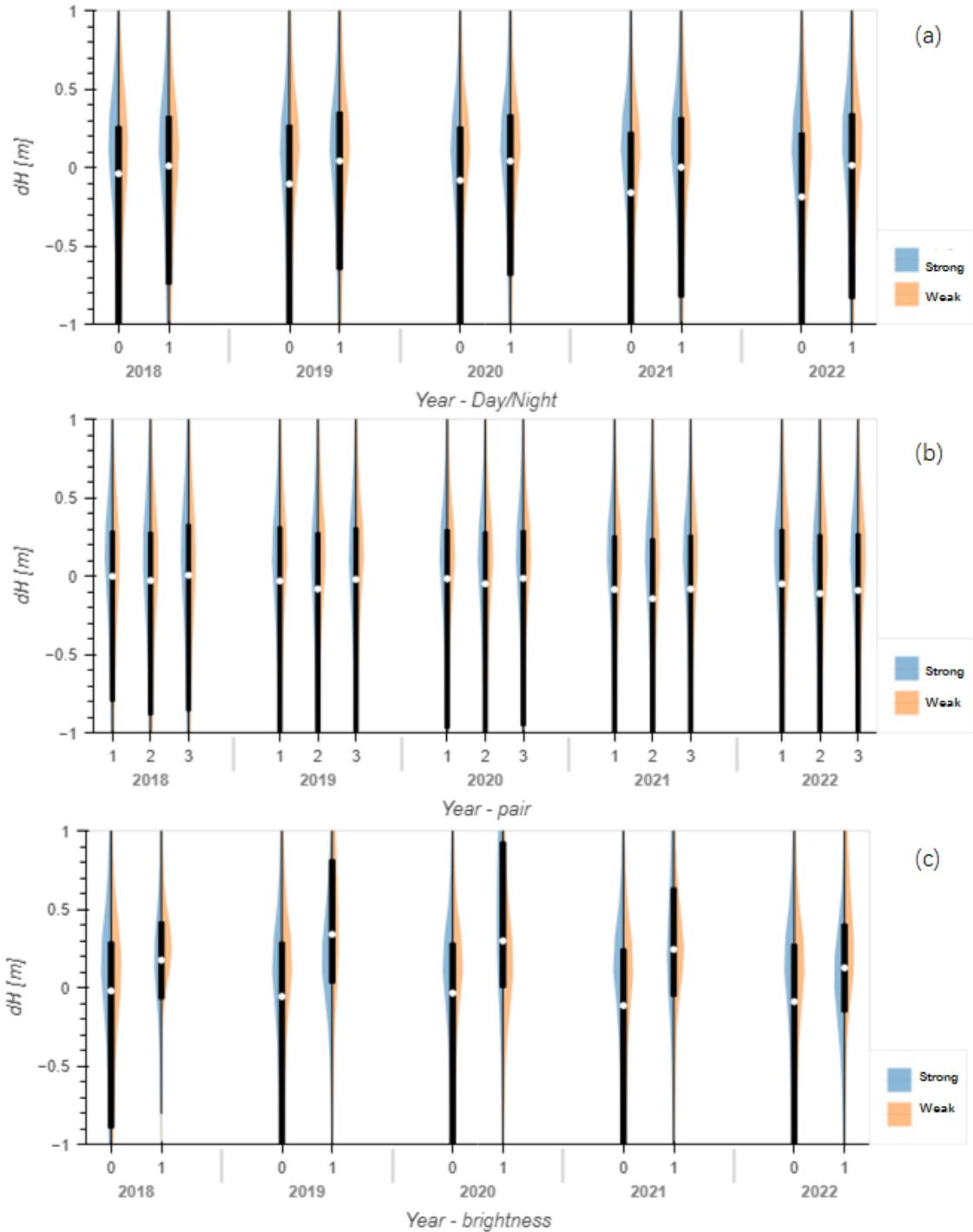


Figure A.6.: ICESat-2 Consistency: Year, Day/night, Strong/Weak, Beams and Pairs. The x-axis is grouped by year and further sub-categorized by day/night (0/1), pair (1/2/3), and brightness_flag (0/1). The plot is split based on the strong and weak beam. The median is represented by a white dot, while the thick bar's edge indicates the 25% (Q1) and 75% quantiles (Q3), respectively.

A.5. Code Repository

A.5.1. Offset adjustment: spatial aggregation, linear adjustment and quantile mapping

Listing A.1: The Offset Adjustment

```

1  import numpy as np
2  from scipy.stats import gamma
3
4  def offset_adjustment_aggregate(df,dems=['dtm1','dtm10','cop30','fab']):
5
6      df['E_bins'] = pd.cut(df['E'], bins=np.arange(np.floor(df['E'].min()), np.ceil(df['E'])
7          ].max()) + 10000, 10000)
8
9      df['N_bins'] = pd.cut(df['N'], bins=np.arange(np.floor(df['N'].min()), np.ceil(df['N']
10          ].max()) + 10000, 10000))
11
12     df_grouped = df.groupby(['E_bins', 'N_bins'])
13
14
15     # Compute the mean of the grouped values
16
17     for dem in dems:
18
19         df[f'bias_{dem}'] = df_grouped[f'df_predict_{dem}'].transform(lambda x: x.median())
20
21         df[f'sd_predict_{dem}_'] = df[f'sd_predict_{dem}'] - df[f'bias_{dem}']
22
23     return df
24
25
26 def offset_adjustment_linear(df,dems=['dtm1','dtm10','cop30','fab']):
27
28     # parameters need to change accoding to observations
29
30     slope = [0.48,0.49,0.49,0.48]
31
32     intercept = [0.74,0.75,0.77,0.77]
33
34     # linear conversion excepth values beyond min and max
35
36     for dem,s,i in zip(dems,slope,intercept):
37
38         df[f'sd_predict_{dem}_'] = (df[f'sd_predict_{dem}'] - i) / s
39
40         df.loc[df[f'sd_predict_{dem}_'] < 0, f'sd_predict_{dem}_'] = 0
41
42         df.loc[df[f'sd_predict_{dem}_'] > 10, f'sd_predict_{dem}_'] = 10
43
44     return df
45
46
47 def offset_adjustment_qm(df,dems=['dtm1','dtm10','cop30','fab']):
48
49     ...
50

```

A. Appendix

```
51     requires a target distribution
52
53     ...
54
55     # input
56
57     target_q = np.percentile(df['sd_nve_10'], np.linspace(0, 100, 101))
58
59     # Compute the mean of the grouped values
60
61     for dem in dems:
62
63         # Compute the quantiles of the original data and the target distribution
64
65         x_q = np.percentile(df[f'sd_predict_{dem}'], np.linspace(0, 100, 101))
66
67         x_mapped = np.interp(df[f'sd_predict_{dem}'], x_q, target_q)
68
69         df[f'sd_predict_{dem}_'] = x_mapped
70
71     return df
72
73 # example
74 pts_nve = offset_adjustment_linear(pts_nve)
```

A.5.2. Gradient descent co-registration

The gradient descent co-registration algorithm used in this study was implemented in Python and can be accessed on GitHub at <https://github.com/GlacioHack/xdem>. A brief overview, an application case, a discussion of the algorithm is provided in Section 3.4, 4.1, 5.1, but for a detailed and actual code implementation, please refer to the online repository.

A.5.3. XSnow

XSnow (<https://github.com/liuh886/xsnow>) is the production repository in this study. It contains the following function / scripts:

1. Download data from ERA5 Land Reanalysis
2. Download data from ICESat-2
3. Process the data (3) into dataframe formats.
4. Extract, couple the dataframe (3) with climate dataset (1)
5. Pipeline the processes of DEM co-registration
6. Bias correction regression model
7. Snow depth downscaling regression model
8. Various plot functions for this study.