MULTI-SCALE ADVANCEMENTS IN OPTICAL REMOTE SENSING

FOR SNOW AND WATER APPLICATIONS

by

James William Dillon

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Engineering

MONTANA STATE UNIVERSITY
Bozeman, Montana

July 2024

TABLE OF CONTENTS

TABLE OF CONTENTS CONTINUED

TABLE OF CONTENTS CONTINUED

TABLE OF CONTENTS CONTINUED

LIST OF TABLES

LIST OF TABLES CONTINUED

LIST OF FIGURES

Figure Page

LIST OF FIGURES CONTINUED

LIST OF FIGURES CONTINUED

LIST OF FIGURES CONTINUED

LIST OF FIGURES CONTINUED

LIST OF FIGURES CONTINUED

LIST OF FIGURES CONTINUED

LIST OF FIGURES CONTINUED

LIST OF FIGURES CONTINUED

LIST OF FIGURES CONTINUED

ABSTRACT

Snow occupies a large portion of the Earth's surface and represents an imperative hydrological resource, a severe hazard, and a key component in our planetary energy budget. The high reflectivity and albedo of snow exerts control on Earth's radiative energy balance, with a direct effect on global climate. Further, snowmelt accumulates in headwater channels; the timing of this phenomenon is critical to ecological health, water quality, downstream water supply, and watershed connectivity. Last, snow presents a serious hazard to both human life and infrastructure, namely in the form of avalanches. Hundreds of lives have been lost to snow avalanches in the past decade. Therefore, accurate measurements of physical snowpack properties and runoff are critical. A burgeoning era of remote sensing is allowing for more continuous measurements with rapidly increasing spatial, spectral, and temporal resolutions. As troves of data become available, researchers are developing new methodologies to keep pace. In the work presented here, we join these efforts.

The broad objective of our work was to use near-infrared (NIR) lidar and hyperspectral imaging to create maps of physical snowpack properties and headwater channels. We hypothesize that underutilized instruments and spatial analysis concepts can address limitations in snowpack mapping techniques. Using a cold laboratory and the natural laboratory of Montana's mountains, we examined three primary concepts: snow grain size, avalanche hazard assessment, and forested headwater channels. Regarding grain size, this parameter controls snow albedo and can be estimated from reflectance, although numerous methods towards this end produce dramatically disparate results. We performed an intercomparison, contrasting models and retrieval techniques to find the optimal method. Further, we validated the novel use of lidar for this purpose, with significant advantages over traditional passive detection. Towards avalanche forecasting, we created a technique that leverages NIR texture to map surface hoar, responsible for over a third of large avalanches. Additionally, we synthesized lidar data with dynamics modeling to predict threats to roadways. Last, towards improved streamflow monitoring, we developed a new means of mapping streams hidden in densely forested regions via airborne lidar. The methods presented here will aid practitioners across multiple fields of the cryosphere.

CHAPTER ONE

INTRODUCTION

Overview

Snow occupies a large portion of the Earth's surface and represents an imperative hydrological resource, a severe hazard, and a key component in our planetary energy budget. The high reflectivity and albedo of snow exerts control on Earth's radiative energy balance, with a direct effect on global and regional climate. Further, seasonal snowmelt accumulates in headwater channels, generating continuous flows that either persist annually or exhibit temporal dynamics. The timing of these phenomena is critical to ecological health, water quality, and watershed connectivity. Last, snow presents a serious hazard to both human life and infrastructure, namely in the form of avalanches. Hundreds of lives have been lost to snow avalanches in the past decade. Despite this threat, few methods towards mapping specific avalanche hazards exist, and avalanche flow models are constrained by poor input data. Therefore, accurate measurements of physical snowpack properties and snowmelt runoff are integral to climate change, headwater hydrology, and avalanche forecasting. However, much of the Earth's snow cover exists in remote and isolated terrain, making measurements of snow and snowmelt difficult. Further, snowpack properties and runoff are highly variable in both space and time. The goal of the research presented here is to use near-infrared (NIR) lidar and hyperspectral imagery, while integrating spatial and radiometric data where appropriate, to create maps of physical snowpack properties and snowmelt-supplied headwater channels. Resulting data products aim to inform estimations of Earth's surface energy balance, hydrologic practitioners, and avalanche forecasters.

Background

Snow occupies a large portion of the Earth's surface and represents an imperative hydrological resource, a severe hazard, and a key component in our planetary energy budget. The high reflectivity and albedo of snow exerts control on Earth's radiative energy balance, with a direct effect on global and regional climate (Wiscombe and Warren, 1980). Further, seasonal snowmelt constitutes a large natural reservoir, supplying one-sixth of the global population and 75% of the western United States with fresh water (Barnett et al., 2005; Simpkins, 2018). Snowmelt runoff accumulates in headwater channels, generating continuous flows that either persist annually or exhibit temporal dynamics (Buttle et al., 2012; Leigh et al., 2016; McDonough et al., 2011). The timing of these phenomena is critical to ecological health, water quality, and watershed connectivity (Camporese et al., 2010; Godsey and Kirchner, 2014; Hooshyar et al., 2015). Last, snow presents a serious hazard to both human life and infrastructure, namely in the form of avalanches. Hundreds of lives have been lost to snow avalanches in the past decade. Despite this threat, few methods towards mapping specific avalanche hazards exist (Buhler et al., 2014; Horton and Jamieson, 2017), and avalanche flow models are often constrained by poor input data (Buhler et al., 2011; Christen et al., 2010b; Maggioni et al., 2012). Therefore, accurate measurements of physical snowpack properties and snowmelt runoff are integral to climate models, headwater hydrology, and avalanche forecasts. However, much of the Earth's snow cover exists in remote and isolated terrain, making measurements of snow and snowmelt difficult. Further, snowpack properties, stability, and runoff are highly variable in both space and time (Deems et al., 2015; Lutz and Birkeland, 2011; Reuter et al., 2016; Schweizer et al., 2008).

Rapid advancements in remote sensing technologies have accelerated measurements of physical snowpack properties on a broad scale. Satellite and aerial data are publicly available in troves. Sensors continue to boast improved signal-to-noise ratios, spatial and spectral resolution, etc., while decreasing in cost. Long-range terrestrial instruments simplify data collection, as do pre-programmable routes for uncrewed aerial systems. Researchers have rushed to keep pace with the burgeoning era of remote sensing in the snow and hydrological sciences, developing ample new methodologies to make optimal use of abundant data. However, significant shortcomings and knowledge gaps still remain. For example, NIR sensors and imagers have been used to map snow optical grain radius (r_{opt}) for decades (Nolin and Dozier, 2000), as NIR reflectance is sensitive to snow microstructure. Numerous methods have been developed in this regard, ranging in scale from laboratory (e.g., Donahue et al., 2021, 2022), to ground-based (e.g., Matzl and Schneebeli, 2006; Painter et al., 2007), to airborne and spaceborne platforms (e.g., Painter et al., 2012; Seidel et al., 2016). This crucial optical snowpack parameter dictates snow spectral albedo, and thus maps of r_{opt} are used to initialize climate models. However, variability in the techniques and models used for r_{opt} retrievals can produce dramatically disparate results, creating substantial uncertainty in r_{opt} and thus albedo estimates. Further, many studies pursuing this end utilize passive detectors which neglect the influence of complex topography on bidirectional reflectance geometry, a severe oversight (Yang et al., 2017).

A similar problem exists when mapping hydrological dynamics in snowmelt-supplied headwater channels. Like snow properties, NIR imagers can also be utilized to delineate water bodies and wetted areas, leveraging a low reflectance signature (McFeeters, 1996; Xu, 2006). Yet, this technique falls short when dense vegetation is encountered, as aerial passive sensors are scarcely able to determine land cover characteristics beneath the forest canopy (Assendelft and Meerveldt, 2019). Unfortunately, most headwater channels are located under substantial vegetation, rendering such procedures inadequate (Hooshyar et al., 2015; Nadeau and Rains, 2007). The paucity of adequate remotely sensed data is perhaps most pronounced in the realm of avalanche forecasting and modeling. For instance, there is a profound disconnect between traditional snow optics and forecasters; despite numerous methods for mapping r_{opt}, no techniques exist to remotely map snow surface structures of concern for weak layer formation (Buhler et al., 2014; Horton and Jamieson, 2017). Furthermore, though modern avalanche simulators are quite sophisticated, inputs are often relatively rudimentary, with heavy reliance on manual point measurements of snow and coarse land cover maps that curb accuracy. Additional means of remotely mapping snowpack properties and runoff are therefore vital to hydrology and cryosphere science.

Underutilized instruments and spatial analysis concepts can address limitations in current snowpack mapping techniques. The use of Light Detection and Ranging (lidar) in the cryosphere, for example, has almost exclusively been as a spatial data tool for high-resolution snow depth mapping (Deems et al., 2013). However, given its profound spatial and radiometric capacity, lidar presents itself as an ideal instrument for measuring NIR reflectance and subsequent extraction of r_{opt}, for delineation of snowmelt channels under canopy cover, and for initialization of avalanche dynamics models. As an active remote sensing system, lidar holds several advantages relative to passive remote sensing (Yang et al., 2017), like the ability to penetrate canopy cover. Further, when using lidar, there are far fewer potentially confounding factors to account for, such as varying strength of illumination, solar zenith angle, relative azimuth between source and sensor, atmospheric conditions, diffuse radiation, and varying polarization state of incident irradiance. Additionally, lidar produces a coincident, high-resolution elevation data product (a Digital Surface Model, DSM) which can be used to account for spatial variation in incident and viewing zenith angles across complex topography. Just as lidar has been overlooked as a remote sensing instrument in the cryosphere, so too has NIR texture been overlooked as a valuable optical parameter. Traditionally snow studies have focused on the magnitude of NIR reflectance rather than localized reflectance variability, hence texture. The only optical texture analysis of snow to date (Champollion et al., 2013) found that texture can be used to delineate surface hoar, a prevalent avalanche weak layer once buried.

The goal of the research presented here is to use NIR lidar and hyperspectral imagery, while integrating spatial and radiometric data where appropriate, to create maps of physical snowpack properties and snowmelt-supplied headwater channels (Figure 1). Resulting data products aim to inform estimations of Earth's surface energy balance, hydrologic practitioners, and avalanche forecasters. Research will span from controlled laboratory environments, leveraging Montana State University's Subzero Research Laboratory (MSU's SRL), to terrestrial lidar data collected in the natural laboratory of Montana's well-instrumented mountain ranges, and finally to public aerial lidar acquisitions supplemented by USGS hydrological data. Statistical analyses are performed to evaluate the accuracy and uncertainty of subsequent data products.

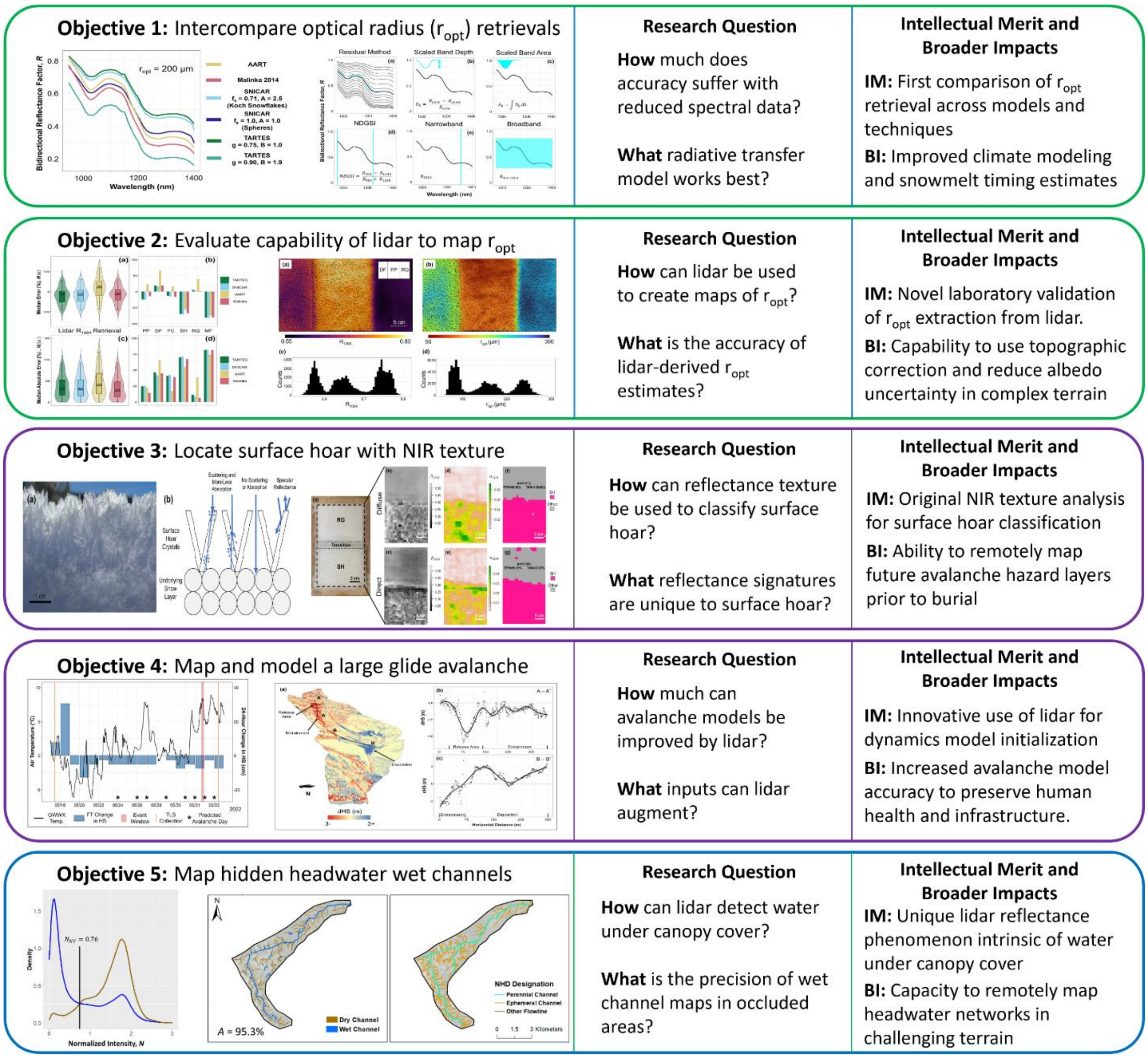

Figure 1: Summary of project research questions and objectives. Green, purple, and blue outlines indicate objectives related to surface energy/climate, avalanche, and hydrological goals, respectively.

Objectives

The following five objectives will be accomplished to achieve the overall goal of this proposal.

Objective 1: Intercompare retrievals of optical grain size between various radiative transfer models and retrieval techniques, ranging from broadband to hyperspectral, to determine optimal methods for accurate r_{opt} estimates.

Task 1: Using naturally collected specimens and the SRL's snowmaker, produce snow samples with varying microstructure while controlling for potentially confounding factors, such as light-absorbing impurities.

Task 2: Perform rigorous quantification of physical snowpack properties, including r_{opt}, density, and grain habit, using X-Ray Computed Microtomography (micro-CT) and microscopy.

Task 3: Acquire reflectance measurements using a NIR hyperspectral imager with a laboratory benchtop apparatus.

Task 4: Leverage measured spectra using six different retrieval techniques, in conjunction with four popular radiative transfer models, to produce r_{opt} retrievals for each sample.

Task 5: Execute an accuracy assessment and radiative transfer model/retrieval technique comparison to determine optimal methods for r_{opt} estimates relative to micro-CT measurements.

Objective 2: Demonstrate the novel capacity of lidar to create maps of r_{opt} from NIR reflectance measurements across a range of incidence angles, using radiative transfer inversion and leveraging a controlled subzero laboratory environment.

Task 1: Produce and characterize snow samples as with Objective 1.

Task 2: Scan samples at eight different incidence angles using a commercial terrestrial lidar unit.

Task 3: Compute reflectance-based estimates of r_{opt} via radiative transfer inversion using multiple models.

Task 4: Evaluate the accuracy of lidar r_{opt} retrievals across incidence angles and models, comparing to passive measurements from NIR hyperspectral imaging (Objective 1) and micro-CT benchmarks.

Task 5: Create statistical models to better relate lidar reflectance to grain size, if radiative transfer models are inaccurate.

Objective 3: Create classified maps delineating the extent of surface hoar from other snow surface structures on a pixelwise basis, leveraging NIR texture metrics.

Task 1: Repeat Tasks 1 and 2 from Objective 1 to create and characterize snow samples with a wide variety of well-defined microstructures.

Task 2: Scan each sample at a nadir geometry with both lidar and a compact hyperspectral imager, examining direct and diffuse illumination scenarios.

Task 3: Perform a moving window analysis to convert lidar and hyperspectral maps of NIR reflectance magnitude to pixelwise measurement of localized variance, or texture.

Task 4: Statistically analyze texture distributions between samples and, using empirical probability density functions, determine optimal texture thresholds to delineate surface hoar from other surface microstructures.

Task 5: Use critical thresholds in a classification tree workflow to map surface hoar extent on a per-pixel basis, then assess accuracy and uncertainty.

Objective 4: Utilize lidar-derived snowpack maps to document a large glide avalanche and to optimize an avalanche dynamics model. Use the tuned model to predict a critical release depth necessary for the avalanche to reach roads and infrastructure.

Task 1: Collect terrestrial lidar data surrounding a large glide avalanche event, leveraging partnerships with the U.S. Geological Survey.

Task 2: Acquire meteorological data from automated weather stations and SNOTEL sites to investigate conditions leading up to avalanche release.

Task 3: Using lidar-derived maps of snow depth and change in snow depth, create visualizations of the avalanche and perform a mass balance analysis.

Task 4: Re-create the event in the avalanche dynamics model RAMMS, using lidar maps and mass/volume calculations during initialization.

Task 5: After tuning the model, incrementally increase the depth in the release area to determine the critical scenario necessary for road impact.

Objective 5: Combine spatial and radiometric data from aerial lidar to map headwater channels under dense vegetation cover using overlapping lidar and stream gage data.

Task 1: Locate study sites with high-spatial resolution lidar acquisitions coincident with continuous USGS stream gage data to determine the site's hydrological condition.

Task 2: Quantify the statistical significance of lidar reflectance reductions between wetted headwater streams versus dry ground, particularly when encountering dense vegetation cover.

Task 3: Compute optimal lidar reflectance thresholds to map wet headwater channels in vegetated, mountainous environments using empirical probability density functions.

Task 4: Generate a Geographical Information System (GIS) workflow to delineate occluded drainage networks and further segregate wet channels from dry segments.

Task 5: Assess the accuracy of classified wet channel data products as compared to the National Hydrography Dataset (NHD).

References

Assendelft, R. S., & van Meerveld, H. J. (2019). A low-cost, multi-sensor system to monitor temporary stream dynamics in mountainous headwater catchments. *Sensors*, *19*(21), 4645.

Barnett, T.P., J.C. Adam and D.P. Lettenmaier 2005. Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, 438(7066): 303-309.

Bühler, Y., Christen, M., Kowalski, J., and Bartelt, P.: Sensitivity of snow avalanche simulations to digital elevation model quality and resolution, Annals of Glaciology, 52, 72–80, 2011

Bühler, Y., Meier, L., & Ginzler, C. (2014). Potential of operational high spatial resolution near-infrared remote sensing instruments for snow surface type mapping. *IEEE Geoscience and Remote Sensing Letters*, *12*(4), 821-825.

Buttle, J. M., Boon, S., Peters, D. L., Spence, C., Van Meerveld, H. J., & Whitfield, P. H. (2012). An overview of temporary stream hydrology in Canada. *Canadian Water Resources Journal/Revue Canadienne Des Ressources Hydriques*, *37*(4), 279-310.

Camporese, M., Paniconi, C., Putti, M., & Orlandini, S. (2010). Surface-subsurface flow modeling with path-based runoff routing, boundary condition-based coupling, and assimilation of multisource observation data. *Water Resources Research*, *46*(2).

Champollion, N., Picard, G., Arnaud, L., Lefebvre, E., & Fily, M. (2013). Hoar crystal development and disappearance at Dome C, Antarctica: observation by near-infrared photography and passive microwave satellite. *The Cryosphere*, *7*(4), 1247-1262.

Christen, M., Kowalski, J., and Bartelt, P.: RAMMS: Numerical simulation of dense snow avalanches in three-dimensional terrain, Cold Regions Science and Technology, 63, 1–14, 2010b.

Deems, J. S., Painter, T. H., & Finnegan, D. C. (2013). Lidar measurement of snow depth: a review. *Journal of Glaciology*, *59*(215), 467-479.

Deems, J. S., Gadomski, P. J., Vellone, D., Evanczyk, R., LeWinter, A. L., Birkeland, K. W., and Finnegan, D. C.: Mapping starting zone snow depth with a ground-based lidar to assist avalanche control and forecasting, Cold Regions Science and Technology, 120, 197–204, 2015.

Donahue, C., Skiles, S. M., & Hammonds, K. (2021). In situ effective snow grain size mapping using a compact hyperspectral imager. *Journal of Glaciology*, *67*(261), 49-57.

Donahue, C., Skiles, S. M., & Hammonds, K. (2022). Mapping liquid water content in snow at the millimeter scale: an intercomparison of mixed-phase optical property models using hyperspectral imaging and in situ measurements. *The Cryosphere*, *16*(1), 43-59.

Godsey, S. E., & Kirchner, J. W. (2014). Dynamic, discontinuous stream networks: hydrologically driven variations in active drainage density, flowing channels and stream order. *Hydrological Processes*, *28*(23), 5791-5803.

Hooshyar, M., Kim, S., Wang, D., & Medeiros, S. C. (2015). Wet channel network extraction by integrating LiDAR intensity and elevation data. *Water Resources Research*, *51*(12), 10029-10046.

Horton, S., & Jamieson, B. (2017). Spectral measurements of surface hoar crystals. *Journal of Glaciology*, *63*(239), 477-486.

Leigh, C., Boulton, A. J., Courtwright, J. L., Fritz, K., May, C. L., Walker, R. H., & Datry, T. (2016). Ecological research and management of intermittent rivers: an historical review and future directions. *Freshwater Biology*, *61*(8), 1181-1199.

Lutz, E. R., & Birkeland, K. W. (2011). Spatial patterns of surface hoar properties and incoming radiation on an inclined forest opening. *Journal of Glaciology*, *57*(202), 355-366.

Maggioni, M., Freppaz, M., Christen, M., Bartelt, P., and Zanini, E.: Back-calculation of small avalanches with the 2D avalanche dynamics model RAMMS: four events artificially triggered at the Seehore test site in Aosta Valley (NW-Italy), in: Proceedings of the International Snow Science Workshop, 16–21 September 2012, Anchorage, Alaska, pp. 591–598, 2012.

Matzl, M., & Schneebeli, M. (2006). Measuring specific surface area of snow by near-infrared photography. *Journal of Glaciology*, *52*(179), 558-564.

McDonough, O. T., Hosen, J. D., & Palmer, M. A. (2011). Temporary streams: the hydrology, geography, and ecology of non-perennially flowing waters. *River ecosystems: dynamics, management and conservation*, 259-290.

McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International journal of remote sensing*, *17*(7), 1425-1432.

Nadeau, T. L., & Rains, M. C. (2007). Hydrological connectivity between headwater streams and downstream waters: how science can inform policy 1. *JAWRA Journal of the American Water Resources Association*, *43*(1), 118-133.

Nolin, A. W., & Dozier, J. (2000). A hyperspectral method for remotely sensing the grain size of snow. *Remote sensing of Environment*, *74*(2), 207-216.

Reuter, B., Richter, B., & Schweizer, J. (2016). Snow instability patterns at the scale of a small basin. *Journal of Geophysical Research: Earth Surface*, *121*(2), 257-282.

Schweizer, J., Kronholm, K., Jamieson, J. B., & Birkeland, K. W. (2008). Review of spatial variability of snowpack properties and its importance for avalanche formation. *Cold Regions Science and Technology*, *51*(2-3), 253-272.

Seidel, F. C., Rittger, K., Skiles, S. M., Molotch, N. P., & Painter, T. H. (2016). Case study of spatial and temporal variability of snow cover, grain size, albedo and radiative forcing in the Sierra Nevada and Rocky Mountain snowpack derived from imaging spectroscopy. *The Cryosphere*, *10*(3), 1229-1244.

Simpkins, G. 2018. Snow-related water woes. *Nature Climate Change*, 8(11): 945-945.

Wiscombe, W. J., & Warren, S. G. (1980). A model for the spectral albedo of snow. I: Pure snow. *Journal of Atmospheric Sciences*, *37*(12), 2712-2733.

Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International journal of remote sensing*, *27*(14), 3025-3033.

Yang, Y., Marshak, A., Han, M., Palm, S. P., & Harding, D. J. (2017). Snow grain size retrieval over the polar ice sheets with the Ice, Cloud, and land Elevation Satellite (ICESat) observations. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *188*, 159-164.

CHAPTER TWO

OPTICAL GRAIN SIZE RETRIEVAL: A LABORATORY COMPARISON OF RADIATIVE TRANSFER MODELS, SHAPE PARAMETERS, AND TECHNIQUES

Contribution of Authors and Co-Authors

Manuscript in Chapter 2

Author: James Dillon

Contributions: Conceptualized the study, collected laboratory data, analyzed results, and drafted the manuscript.

Co-Author: Christopher Donahue

Contributions: Provided guidance and advised during conceptualization and analysis.

Co-Author: Evan Schehrer

Contributions: Assisted with laboratory sample preparation and data collection.

Co-Author: Kevin Hammonds

Contributions: Acquired funding, was responsible for project administration, and supervised throughout.

Manuscript Information

James Dillon, Christopher Donahue, Evan Schehrer, Kevin Hammonds

Status of Manuscript:
X Prepared for submission to a peer-reviewed journal
☐ Officially submitted to a peer-reviewed journal
☐ Accepted by a peer-reviewed journal
☐ Published in a peer-reviewed journal

Optical grain size retrieval: a laboratory comparison of radiative transfer models, shape parameters, and techniques

James W. Dillon[1], Christopher P. Donahue[2,3], Evan N. Schehrer[1], Kevin D. Hammonds[1]
[1]Department of Civil Engineering, Montana State University, Bozeman, MT, USA
[2]Department of Geography, Earth, and Environmental Sciences, University of Northern British Columbia, Prince George, British Columbia, Canada
[3]Hakai Institute, Campbell River, Quadra Island, British Columbia V9W 0B7, Canada

Abstract. The near-infrared (NIR) albedo of snow is controlled by optical snow grain size (r_{opt}). Therefore, characterizing the spatial and temporal variability in r_{opt} at the snow surface is critical for understanding melt timing and magnitude for water availability, and Earth's energy budget towards future climates. While numerous studies have demonstrated estimates of r_{opt} via optical instruments that span scales from handheld-devices to satellites, they leverage differing retrieval techniques, radiative transfer models, and modeled snow grain shapes. Variation in these factors cause tremendous uncertainty in r_{opt} retrievals, yet a thorough comparison has yet to be conducted. To address this knowledge gap, we conducted a laboratory bidirectional reflectance study using NIR hyperspectral imaging (NIR-HSI) under direct zenith illumination and nadir observation. Towards enhanced r_{opt} retrieval accuracy, we sought to determine 1) the optimal modeled snow grain shape, 2) the best-performing radiative transfer model, and 3) variability associated with retrieval techniques, spanning broadband, narrowband, multispectral, and hyperspectral approaches. Shape optimization proved to be vital and interplay between shape variables was key. Regarding retrieval techniques, the residual method performed best, although multispectral and single-band retrievals were comparable to their hyperspectral counterparts at times, likely due to idealized laboratory conditions and high instrument SNR. Following shape-optimization, the SNICAR and TARTES models produced the best results (median absolute error of 15.6 – 17.4%, depending on technique), outperforming the AART model and the model of Malinka (2014). Our results show that the appropriate combination of retrieval technique, radiative transfer model, and shape parameters are imperative for accurate r_{opt} retrievals. As NIR-HSI units and other NIR detectors become more economical, and as their spatial and temporal resolution become more robust, the findings presented here may provide guidance for improved r_{opt} (and thus snow albedo) mapping.

1 Introduction and background

Snow, the most reflective natural surface on Earth, occupies large portions of Earth's surface and plays a critical role for climate and hydrology (Dumont et al., 2021). Snow has a high (up to 90%) albedo, defined as the ratio of reflected solar radiation at the snow surface to that of incoming solar radiation, a significant contributor to the Earth's overall surface energy balance. Furthermore, snow albedo is sensitive to snow microstructure, and this interaction is responsible for numerous climatic feedback loops (Flanner et al., 2012). In terms of hydrology, snow albedo drives the timing and magnitude of snowmelt in mountainous regions which is imperative for water forecasting (Marks and Dozier, 1992). Thus, accurate measurements and modeled estimates of snow albedo, particularly with regards to spatiotemporal variation, are key to understanding future climate, melt rates, and water availability.

The optical properties of ice are well understood (Perovich and Govoni, 1992; Picard et al., 2016; Warren and Brandt, 2008; Warren, 1982,1984), which has led to the development of numerous snow radiative transfer models used to predict the reflectance or albedo of snow based on optical conditions and physical snowpack parameters (Flanner and Zender, 2005; Kokhanovsky and Zege, 2004; Libois et al., 2013; Malinka, 2014; Malinka et al., 2016; Stamnes et al., 1988). In the visible wavelengths, snow is highly reflective, and albedo is primarily driven by impurities near the snow surface (Skiles et al., 2012, 2018). In the near-infrared (NIR) wavelengths ice is absorptive and the primary driver of NIR albedo is the path length of ice, often characterized by the optical grain size, expressed here as a radius r_{opt}. In fact, optical grain size is the primary parameter controlling broadband albedo of clean snow (Wiscombe and Warren, 1980). Therefore, characterizing the spatial and temporal variability in r_{opt} at the snow surface is critical for accurately estimating albedo from remote sensing instruments. By considering the snowpack as a collection of identical, spherical ice particles (Grenfell and Warren, 1999) the optical grain size can then be related to the physical snow microstructure through the ice surface per unit mass (Legagneux et al., 2002), or specific surface area, SSA (Eq. 1).

$$r_{opt} = \frac{3}{\rho_{ice} * SSA} \tag{1}$$

There is an inverse relationship between NIR albedo and optical grain size; as grain size increases, the albedo decreases due to increased absorption. This relationship is the basis from which snow reflectance measurements can be used to retrieve estimates of r_{opt}. A common practice is to simulate snow reflectance for a wide range of r_{opt} values using a radiative transfer model and to populate a lookup table that can then be used compared to measured reflectance. Over the last several decades, numerous methods have been developed to relate modeled to measured spectra that have leveraged single band all the way to full hyperspectral retrieval methods. These efforts range from in situ (e.g., Donahue et al., 2021, 2022b; Gallet et al., 2009; Gergely et al., 2014; Matzl and Schneebeli, 2006; Painter et al., 2007) to airborne platforms (e.g., Donahue et al., 2023; Nolin and Dozier, 2000; Seidel et al., 2016; Skiles et al., 2023) to spaceborne sensors (e.g., Bair et al., 2020; Bohn et al., 2021; Painter et al., 2009, 2012).

Although many studies have demonstrated success at estimating r_{opt}, these differing methods can produce disparate retrievals. This is a salient point, as incorrect ropt estimates can result in substantial differences in predicted snow albedo, which can dramatically influence earth system and climate models (Räisänen et al., 2017; Robledano et al., 2023). To begin, the single band, combination of bands, or spectral features used to match reflectance measurements to simulations (hereafter "retrieval technique") plays a role in retrieval variability. Depending on the instrument, collected data may be broadband, narrowband, multispectral, or hyperspectral. For broadband platforms, only a simulated average reflectance over the sensor bandwidth can be evaluated, while narrowband measurements are matched to the sensor's central wavelength. When using multispectral instruments, a normalized index, such as the Normalized Difference Grain Size Index, or NDGSI (Painter et al., 2012) is often used. Hyperspectral sensors collect continuous spectral measurements and allow for a variety of retrieval techniques, such as measuring the depth or area of normalized ice absorption features (Clark and Roush, 1984; Nolin and Dozier, 2000), or even best-match fitting the entire spectrum, known as the residual method (Donahue et al., 2022).

In addition to differing retrieval techniques, there are several snow radiative transfer models that have been developed to simulate spectra, and the variability between these models also plays a key role in retrieval uncertainty. The longstanding benchmark are strict numerical codes that solve the radiative transfer equations, such as the DIScrete-Ordinate Radiative Transfer model, or DISORT (Stamnes et al., 1988). However, for many practical applications, faster and simpler approximations are often preferred. For instance, the SNow, ICe, and Aerosol Radiative – Adding-Doubling (SNICAR-AD) model (Flanner et al., 2021) is a frequently employed two-stream approximation, hence one that rapidly integrates across all viewing zenith and azimuth angles to produce albedo estimates, that has demonstrated excellent agreement with DISORT (Dang et al., 2019). Despite being an albedo model, SNICAR – AD and others are frequently compared against measured bidirectional reflectance for r_{opt} retrieval at nadir viewing angles, where albedo and reflectance are nearly identical (Dumont et al., 2010). The Approximate Asymptotic Radiative Transfer (AART) snow model is a bidirectional reflectance simulation based on an asymptotic approximation to the radiative transfer equation and geometric optics (Kokhanovsky and Zege, 2004). More recently, Malinka (2014) leveraged this asymptotic theory in a bidirectional reflectance model based on a random binary mixture of two immiscible materials (air and ice), in which optical characteristics change in a stochastic manner between discrete values. Libois et al. (2013) combined a two-stream and asymptotic approximation scheme to create the Two-stream Analytical Radiative TransfEr in Snow (TARTES) albedo model with advanced inclusion of snow grain shape dependence.

Beyond the retrieval technique and radiative transfer model, another factor of relevance is the matter of modeled snow grain shape. Modeled shape, and subsequently how the single scattering grain properties are calculated, is perhaps the biggest difference between the aforementioned models, and thus the greatest cause of r_{opt} retrieval uncertainty between models, and even within a given model. Historically, snow has been modeled as a collection of spheres of equivalent size (Grenfell and Warren, 1999), and single scattering properties determined from Mie calculations. This was originally true of SNICAR – AD, although the model has since been expanded to reflect the address the prevailing belief that the spherical assumption is an oversimplification. While SNICAR – AD still calculates the single scattering albedo of snow using a spherical assumption and Mie calculations, the influence of grain shape on scattering asymmetry (specifically the asymmetry parameter, *g*) is now considered via parameterizations from He et al. (2017). The grain shape can be varied based on the combination of two parameters; the shape factor (f_s) and aspect ratio (*A*). Within the model there are four selectable combinations of these parameters that represent idealized shapes: spheres, spheroids, hexagonal plates, and Koch snowflakes. The shape factor is defined as the ratio of the specific-projected-area-defined effective diameter of a nonspherical grain to that of a spherical grain with the same volume, representing the effect of nonsphericity (He et al., 2017). Altering the combination of f_s and *A* amounts to varying the value of *g*.

The other models examined here (AART, Malinka 2014, and TARTES) leverage geometric optics to calculate single scattering properties. Both AART and the model of Malinka (2014) have fixed "shapes": fractals and a random mixture, respectively. The TARTES model, however, is tunable, and accounts for the influence of shape on both scattering asymmetry and absorption. Shapes in TARTES are also dependent on a two parameter combination: the absorption enhancement parameter, *B* (which is related to single scattering albedo), and the asymmetry parameter, *g*. When creating TARTES, Libois et al. (2013) called for a systemic determination of *B* and *g* in both the field and laboratory using independent measurements of SSA. Although the topic of modeled shape has received greater attention in recent years (e.g., He et al., 2017; Libois et al., 2013, 2014; Robledano et al., 2023), additional experiments in a controlled laboratory environment would be beneficial to the snow optics community. In summary, we note that model and shape selection result in profound differences in simulated spectra (Fig. 1), and this, in conjunction with differing retrieval techniques, can all significantly introduce error and uncertainty into r_{opt} retrievals (Fig. 2). Despite this variability, a thorough intercomparison of retrieval techniques and models has yet to be conducted.

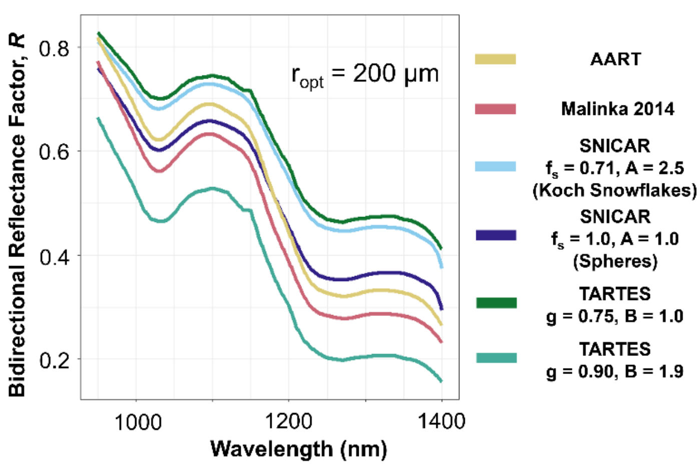

Figure 1: Modeled snow spectra with a constant r_{opt} value of 200 microns, demonstrating substantial variability between radiative transfer models and shape parameter inputs.

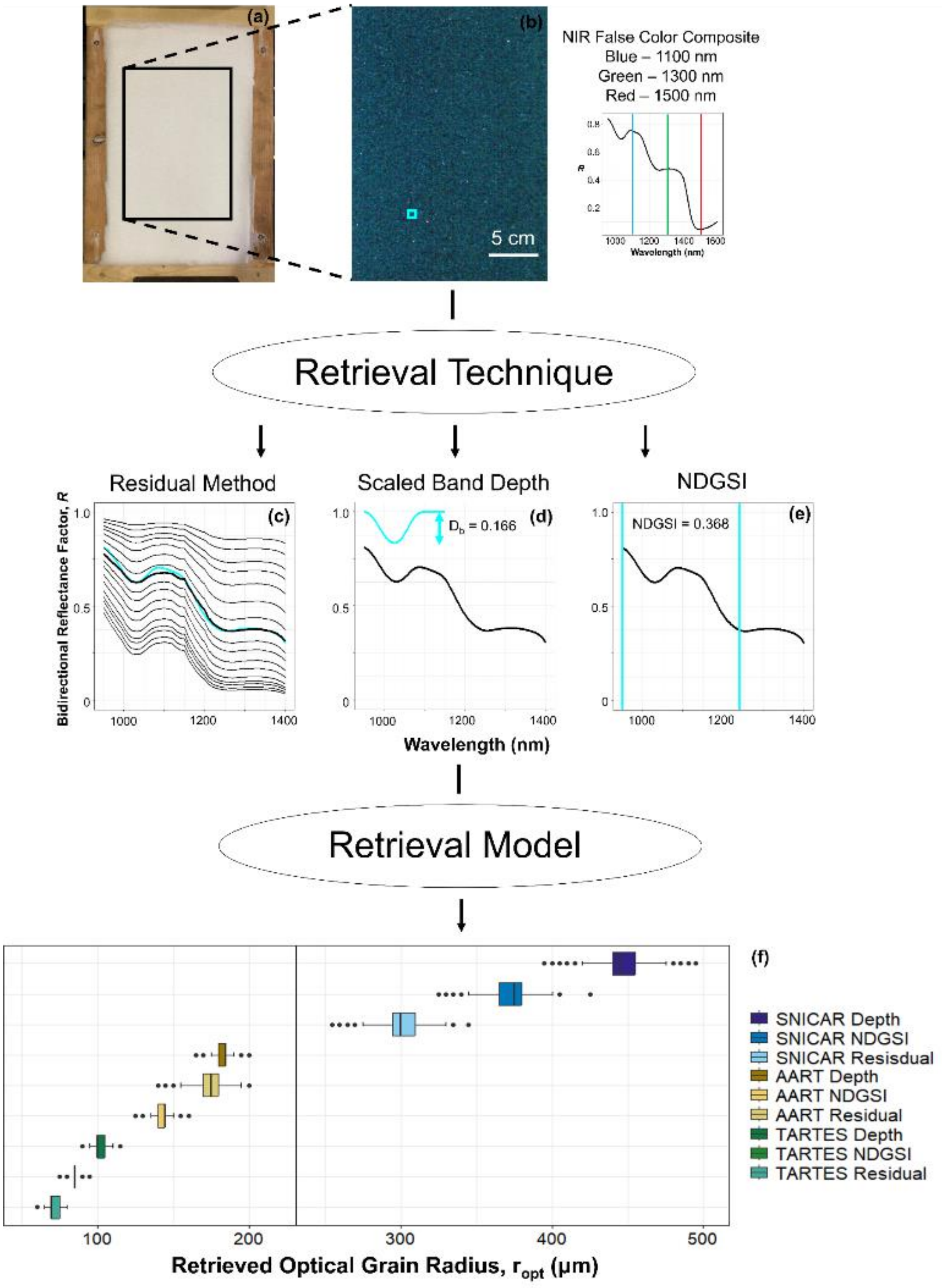

Figure 2: A demonstration of ropt retrieval variability across differing retrieval techniques and radiative transfer models. A visible photograph of a snow sample is shown in (a), as compared to a NIR false color composite (FCC) image in (b), produced from hyperspectral imaging. The data from the highlighted pixel (enlarged for clarity) is then evaluated using the residual (c), scaled band depth (d), and NDGSI (e) retrieval techniques, to be discussed in greater detail in Sect. 2.4. The combination of different retrieval techniques and radiative transfer models leads to dramatic differences in ropt retrievals (f). The black vertical reference line in (f) represents the reference micro-CT ropt measurement. All methods are discussed further in Sect. 2. Data shown are from Sample 18. Regarding shape, Koch snowflakes are used for SNICAR and values of 1.9 and 0.875 for B and g, respectively, in TARTES.

To address these uncertainties, we conducted a laboratory reflectance study to assess r_{opt} retrievals across three factors: retrieval technique, radiative transfer model, and simulated snow grain shape. In an effort to provide future r_{opt} mapping efforts with additional guidance, we sought to address the following questions:

i. Which retrieval technique works best, and to what extent does hyperspectral data improve upon multispectral, narrowband, and broadband retrieval alternatives?
ii. Which radiative transfer model works best?
iii. What combination of optical snow grain shape parameters is the most effective?
iv. How do retrieval technique, radiative transfer model, and simulated snow grain shape interplay regarding r_{opt} retrieval accuracy?

Thus, we created snow samples across a wide range of grain habits and physical properties, each characterized using microscopy and X-ray computed microtomography (micro-CT). We obtained optical data using NIR hyperspectral imaging (NIR-HSI), and determined subsequent r_{opt} retrievals using numerous retrieval techniques, radiative transfer models, and shape parameter combinations. We analyzed resulting values statistically against reference r_{opt} measurements from micro-CT. This represents one of few extensive datasets combining NIR bidirectional reflectance measurements with micro-CT characterization of snow microstructure.

2 Methodology

We aimed to prepare laboratory snow samples with a wide variety of well-defined grain habits and microstructures, acquire optical measurements, and use radiative transfer modeling to perform and intercompare r_{opt} retrievals. Section 2.1 describes snow sample preparation and physical characterization, Sect. 2.2 outlines the acquisition of NIR-HSI data, Sect. 2.3 covers radiative transfer modeling, and Sect. 2.4 discusses retrieval techniques and statistical analyses. The flowchart in Fig. 3 illustrates the entirety of our retrieval comparison process.

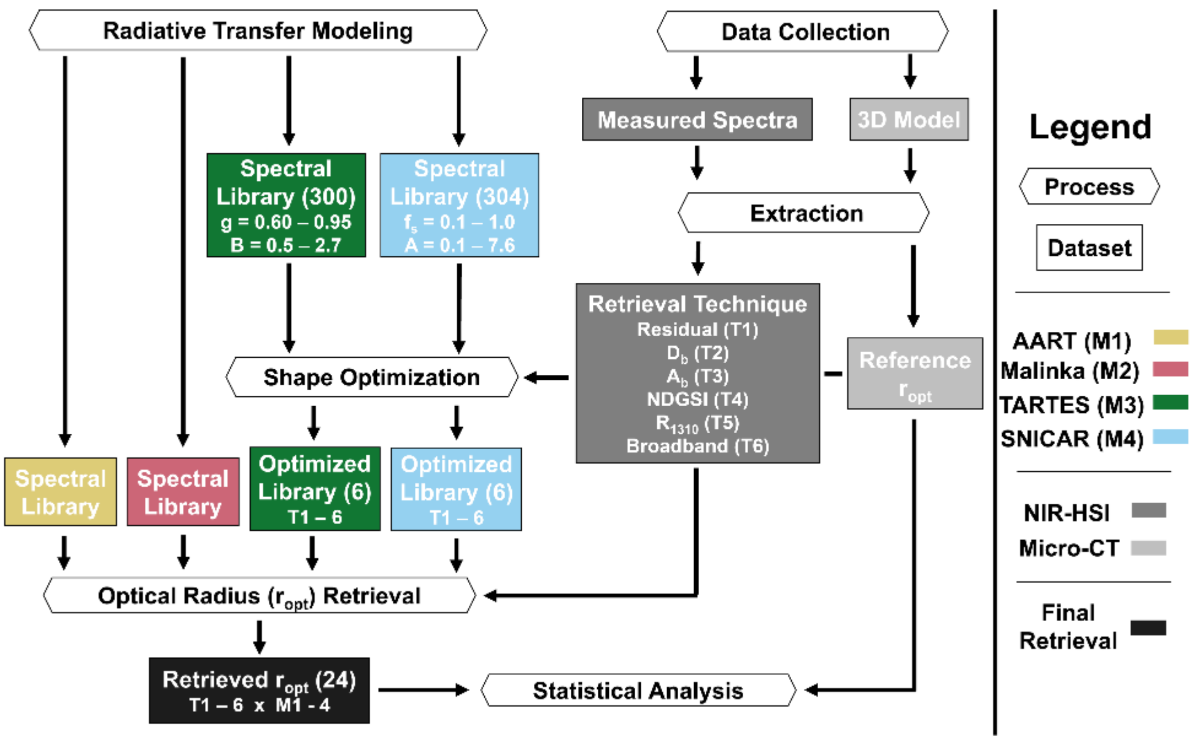

Figure 3: Flowchart of the r_{opt} retrieval and comparison process. Reflectance data from NIR-HSI are paired with four radiative transfer models (M1 – M4) to produce a variety of r_{opt} retrievals. NIR-HSI data are acquired at nadir, and six different sets of data (T1 – T6) are extracted for use in retrieval. For the radiative transfer models that can consider shape, the optimal shape parameter combinations are determined for each retrieval technique. Grain size retrievals for each combination of model and retrieval technique are compared on a samplewise basis to r_{opt} equivalent measurements from micro-CT.

2.1 Sample preparation and physical characterization

The samples used here, and thus the methods for sample preparation and physical characterization, are identical to those from Dillon et al. (2024a). Sample creation and characterization are briefly summarized here, but we refer the reader to the aforementioned publication for a full description.

2.1.1 Sample preparation

We utilized Montana State University's Subzero Research Laboratory (SRL) for sample preparation and assessment. The snow used in these experiments was a combination of laboratory-grown crystals produced in the SRL's snowmaking apparatus and natural undisturbed snow that we collected from the surrounding area. We kept all samples in a cold room at -30° C and allowed them to equilibrate for at least 24 hours prior to evaluation. We prepared forty-one snow samples from twelve batches of differing snow grains. From the bulk batches, we sieved snow grains through various mesh sizes to further promote disparate microstructures (Table 1). The exception to this was surface hoar, which we grew following the methods used by Stanton et al. (2016). Sample grain habits included precipitation particles (PP), decomposing and fragmented precipitation particles (DF), rounded grains (RG), melt forms (MF), faceted crystals (FC), depth hoar (DH), and surface hoar (SH) (Fierz et al., 2009). We prepared snow samples to be microstructurally homogeneous, both laterally across the sample and vertically over sample depth of 3.8 cm.

Table 1: Physical snow sample characteristics organized by primary grain habit and listed in order of decreasing surface area-to-volume ratio therein. Adapted from Dillon et al. (2024a).

Sample #	Batch ID	Primary Grain Habit	Secondary Grain Habit(s)	Micro-CT SSA (kg m^{-2})	Micro-CT r_{opt} (μm)	Micro-CT ρ (kg m^{-3})	Sieve Size (mm)		Notes
							Passed	Caught	
1	A	PP	PPrm, DF	35.85	91.3	176	2.38	1.18	
2	A	PP	PPrm, DF	31.60	103.5	217	2.38	-	
3	A	PP	PPrm, DF	28.69	114.0	211	1.18	0.42	
4	B	PP	PPgp	34.67	94.4	160	2.38	1.18	
5	C	PP	DF	36.10	90.6	94	-	-	In situ fresh PP
6	C	PP	DF	22.40	146.1	286	2.38	1.18	
7	C	PP	DF	22.30	146.7	280	0.85	0.42	
8	C	PP	DF	21.94	149.1	275	2.38	-	
9	C	PP	DF	20.05	163.1	303	1.18	0.85	
10	D	DF	RG	29.92	109.3	293	2.38	1.18	
11	D	DF	RG	28.19	116.1	323	0.85	0.42	
12	D	DF	RG	27.38	119.5	351	1.18	0.85	
13	D	DF	RG	22.65	144.4	365	2.38	-	
14	E	DF	DFbk, RGwp	17.74	184.4	374	0.85	-	
15	F	DF	PP	15.69	208.5	322	2.38	-	
16	F	DF	PP	15.04	217.5	312	2.38	1.18	
17	F	DF	PP	14.89	219.8	309	1.18	0.85	
18	F	DF	PP	14.17	230.9	382	0.85	-	
19	G	FC	DH	16.00	204.5	407	1.18	0.42	
20	G	FC	DH	12.34	265.0	448	2.38	1.18	
21	G	FC	DH	11.19	292.4	417	6.3	3.35	
22	G	FC	DH	10.96	298.5	472	6.3	-	
23	G	FC	DH	10.76	304.0	404	3.35	2.38	
24	H	SH	RG	15.83	206.6	213	6.3	-	Re-sieved SH grains
25	H	SH	RG	11.80	277.3	65	-	-	In situ SH atop RGs
26	H	SH	RG	8.18	400.0	94	-	-	Smaller than S25
27	I	RG	DF	14.75	221.7	381	2.38	1.18	
28	I	RG	DF	14.26	229.4	419	1.18	0.85	
29	I	RG	DF	13.93	234.9	431	2.38	-	
30	I	RG	DF	13.56	241.4	489	0.85	0.42	
31	I	RG	DF	13.52	241.9	452	-	-	S29 melt-refreeze
32	J	RG	DF	15.01	218.0	394	0.85	-	
33	J	RG	DF	14.67	223.0	355	0.42	-	
34	J	RG	DF	11.62	281.4	460	1.18	-	
35	K	RG	DF	12.14	269.5	404	1.18	0.85	
36	K	RG	DF	11.82	276.8	428	0.85	0.42	
37	L	MF	RG	5.41	604.8	582	2.38	0.42	
38	L	MF	RG	4.02	813.0	545	6.3	-	
39	L	MF	RG	3.41	958.5	512	3.15	2.38	
40	L	MF	RG	3.14	1041.7	467	-	-	Refrozen in situ
41	L	MF	RG	2.58	1265.8	433	6.3	3.15	

2.1.2 Physical characterization

We thoroughly characterized the physical properties of each sample, as summarized in Table 1. First, we performed microscopy on representative grains from each batch prior to sieving (Fig. 4), and classified grain habits using a crystal card and lens following Fierz et al. (2009). After sieving and sample preparation, we collected micro-CT data from each sample using a Bruker SkyScan 1173 housed in a -10° C chamber within the SRL, generally following the protocol outlined by Donahue et al. (2021). To prepare samples, we used a cylindrical holder with 3 cm diameter x 4 cm length, which allowed for a voxel size of 14.5 µm. The voxel size of 14.5 µm was the finest spatial resolution achievable with the relatively large cylindrical micro-CT sample holder used in this study. The larger micro-CT sample holder was chosen to provide sufficient surface area for larger-grained samples (e.g., surface hoar) to be encapsulated and transported to micro-CT for measurement. We intuitively note that this will result in neglection or mischaracterization of fine dendrites smaller than this size. Consequentially, it is possible that the SSA of samples with a PP primary grain habit (particularly Samples 1 – 5) was underestimated by micro-CT. After scanning, we performed thresholding of grey-scale images into ice and air phases by visual inspection (e.g., Fig. 4). Reconstructions via the marching cubes method (Lorensen and Cline, 1987) allowed us to determine the volume and surface area in 3D, which we used to calculate the SSA, and thus r_{opt} via Eq. 1, and density of each sample. We used these micro-CT r_{opt} values as benchmarks for comparison to retrievals.

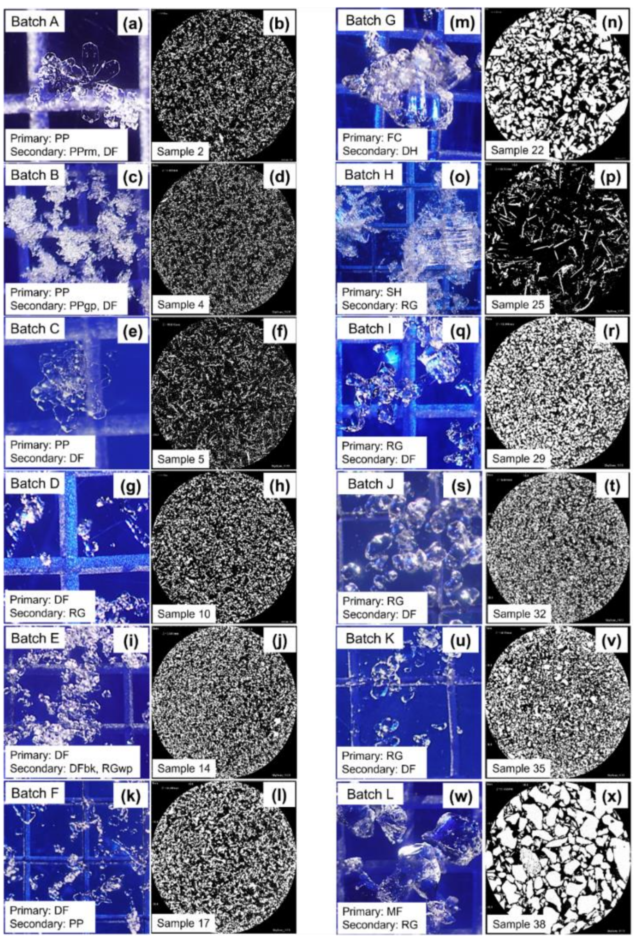

Figure 4: Microscopy images of grains from each initial batch (left columns) and binary micro-CT cross-sections from representative samples (right columns). In the microscopy images, the grid size on the underlying blue grain card is 2 mm. Originally published by *The Cryosphere* in Dillon et al. (2024a).

2.2 Hyperspectral imaging

We used a Resonon Inc. Pika NIR-640 near-infrared hyperspectral imager to map snow reflectance in the NIR (www.resonon.com). Donahue et al. (2021) provide a detailed description of the instrument. Briefly, the imager's spectral resolution ranges from 2.39 to 2.50 nm, and measures 336 bands across the NIR region from 891–1711 nm. It constructs a 2D image containing the full spectrum in each pixel by collecting the image line by line, known commonly as a "push broom" scanner. We used a Resonon benchtop linear scanning stage to move the sample beneath the sensor. For more details on the benchtop apparatus, see Donahue et al. (2022).

We positioned the hyperspectral imager above the linear translating stage that held the samples. The lens of the imager is surrounded by a set of four halogen lamps that produce direct illumination (Fig. 5a). The halogen lamps and lens of the imager are at a height of 38 and 47 cm above the snow surface, respectively. We used a large Spectralon white diffuse reflectance panel to perform calibration, resulting in a reflectance factor (R) measured for each band in every individual pixel of the image. The spectralon panel is 30.5 x 30.5 cm, thus larger in both dimensions than our optical ROI (Fig. 5b). We built a sample holder with the same external dimensions as our snow sample holders, but specifically made to hold the spectralon panel, both centered on the ROI and at the same distance from the illumination source as the snow surfaces. For each snow sample scan, we also conducted a reference scan with the spectralon panel. This allowed for pixel-by-pixel calibration of the entire optical ROI, thus accounting for any heterogeneous illumination. We made these reference measurements for each sample. We acquired all optical data immediately prior to micro-CT analysis at a constant temperature of -10° C.

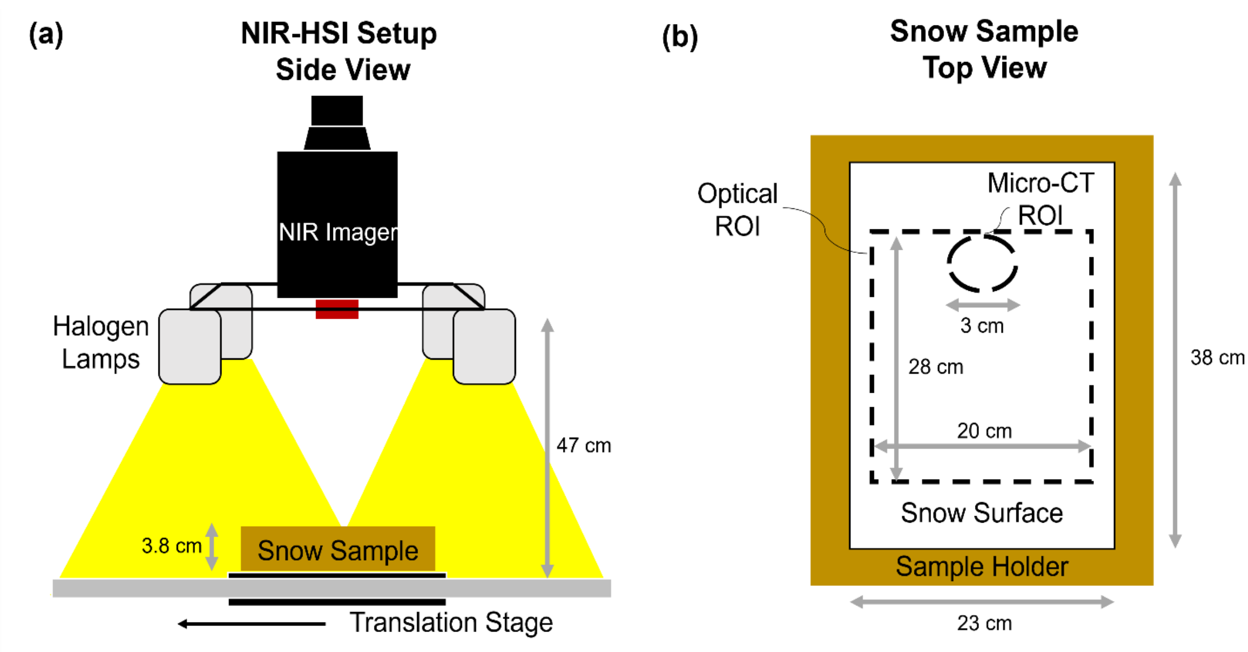

Figure 5: Laboratory data collection schematic for hyperspectral imaging (a). Data regions-of-interest (ROIs) within the snow sample are illustrated in (b). Adapted from Dillon et al. (2024a).

Initial processing took place in Resonon's proprietary Spectronon software, and analyses thereafter performed in Rstudio. We began by truncating the reflectance data from each dataset to a central region-of-interest (ROI) encapsulating the micro-CT ROI (Fig. 5b), in an effort to reduce edge effects. Resulting NIR-HSI ROIs contained 224,000 pixels with a spatial resolution of 0.5 mm. Reflectance images were produced from 188 of the 336 available bands, ranging from 951 – 1403 nm, trimmed to reduce noise at the lower end of the imager spectral range and to exclude regions where snow and ice are scarcely reflective. Examples of measured spectra are presented in Fig. 6.

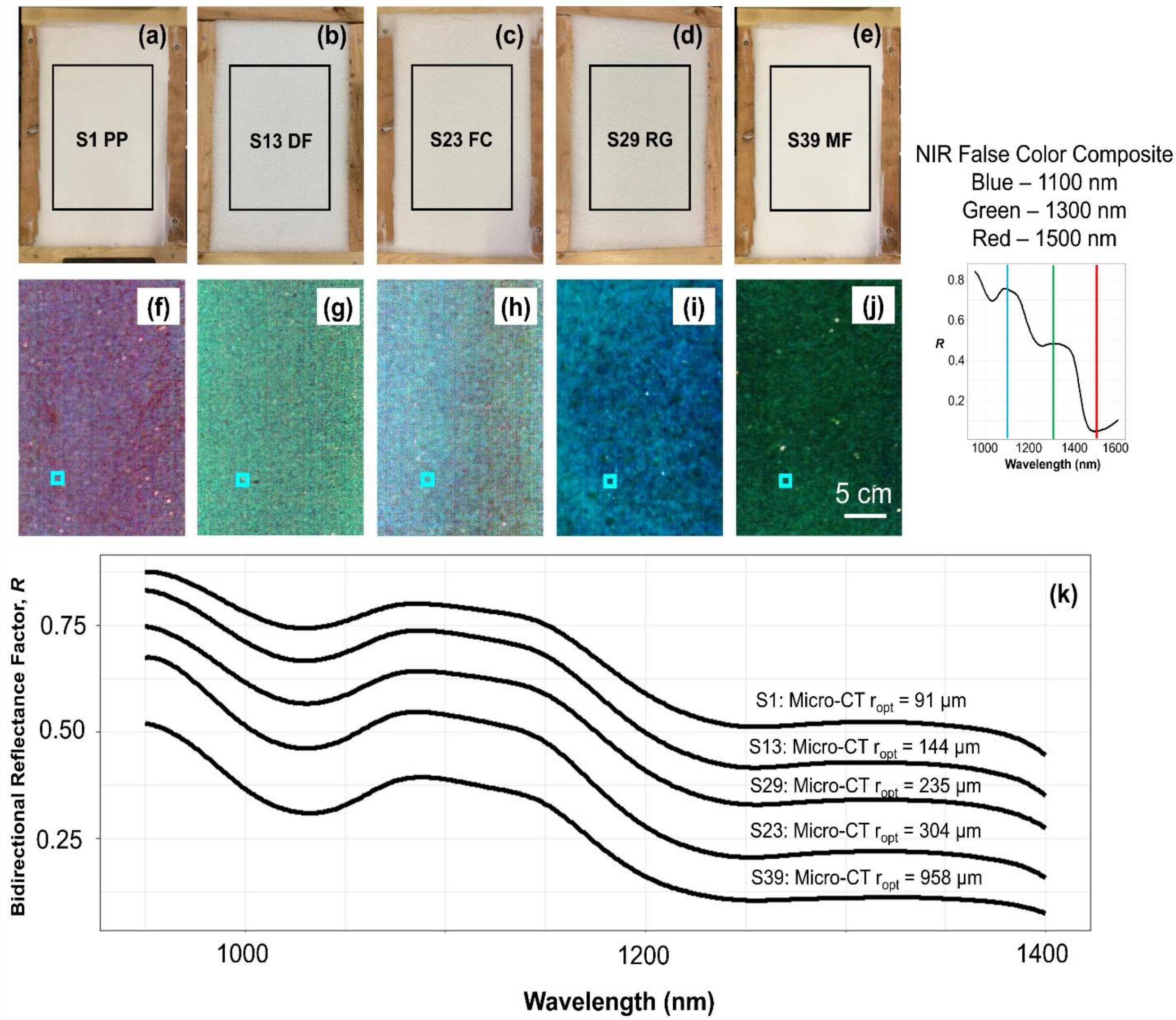

Figure 6: Visible photographs of several snow samples with differing microstructure (a – e) contrasted with their NIR false color composite (FCC) counterparts (f – j). Example spectra from the (enlarged) pixel in each FCC image are shown in (k), illustrating the well-known relationship between grain size and reflectance in the NIR spectral range.

2.3 Radiative transfer modeling

To model snow reflectance, we utilized the four commonly used snow radiative transfer models described in Sect. 1: TARTES, SNICAR, AART, and the model of Malinka (2014). As discussed, TARTES and SNICAR each have two tunable shape parameters which can dramatically vary the modeled spectra and subsequent grain size retrievals. To further investigate the influence of modeled snow grain shape, we produced numerous spectral libraries for both TARTES and SNICAR using modulated combinations of shape parameters. For TARTES, we evaluated absorption enhancement parameter, B, values from 0.8 – 2.7 at increments of 0.1, and asymmetry factor, g, values from 0.60 – 0.95 at increments of 0.025. These ranges spanned all reasonable values based on previous literature (e.g., Libois et al., 2013, 2014; Robledano et al., 2023). Similarly, for SNICAR we varied the shape parameter (f_s) from 0.1 – 1.0 at 0.05 increments and Aspect Factor (A) from 0.1 – 7.6 with steps of 0.5, again spanning all reasonable values (e.g., He et al., 2017) and nearly the full range selectable values in the model. Thus, in total we produced 300 spectral libraries for TARTES and 304 for SNICAR, all at nadir, with each constituting a different combination of shape parameters (Fig. 3). For AART and Malinka, where the modeled snow grain shape was fixed, we generated a single table. All spectral libraries ranged from 950 – 1400 nm at 1 nm resolution, and across r_{opt} values of 30 – 1500 μm at 5 μm increments.

2.4 Retrieval techniques

The goal of an optical grain size retrieval is to match a measured spectrum to a modeled spectrum and obtain the quantitative property. Therefore, to begin, all NIR-HSI data were resampled from the native spectral resolution of ~2.5 nm to 1 nm resolution to match the modeled spectral libraries using spline interpolation. Next, we evaluated six commonly used retrieval techniques (Fig. 7): three hyperspectral, one multispectral, one using a single reflectance value to emulate a narrowband retrieval, and one pseudo-broadband retrieval. Following Donahue et al. (2022), the residual method (Fig. 7a) leverages the entire spectrum and minimizes the residual between the measured and modeled spectrums on a band-by-band basis. Alternative hyperspectral approaches are to assess spectral features related to the prominent ice absorption feature centered at 1030 nm. The scaled band depth, D_b (Fig. 7b), and scaled band area, A_b (Fig. 7c), approaches evaluate the continuum-removed and normalized 1030 nm absorption feature (Clark, 1984; Nolin and Dozier, 2000). Here, the absorption feature is defined as a range from 950 nm (fixed due to the range of the NIR-HSI instrument) to the local maxima around 1100 nm.

For a multispectral retrieval we used the NDGSI (Fig. 7d), which quantifies the relative difference between two reflectance values in the NIR range. For our single wavelength retrieval (Fig. 7e), we selected 1310 nm, a relevant selection given its use in the IceCube (Zuanon and A2 Photonic Sensors, 2013) and DUFISSS (Gallet et al., 2009) instruments. We also evaluated narrowband accuracy at 1064 nm for better comparison with NIR lidar in future publications (Appendix A). Last, to emulate a broadband retrieval, we calculated the average reflectance across the entire measured spectrum (Fig. 7f).

For each retrieval technique, we matched the extracted data from NIR-HSI measurements to the modeled spectra with the closest value, and "retrieved" the corresponding grain size. We repeated this process for all pixels in each sample across all spectral libraries. For the radiative transfer models that can consider shape (TARTES and SNICAR), we identified the optimal shape parameter pairing for each retrieval technique. We calculated samplewise medians of retrieved r_{opt} and compared them to each other and across retrieval technique/model combinations, as well as to reference micro-CT values.

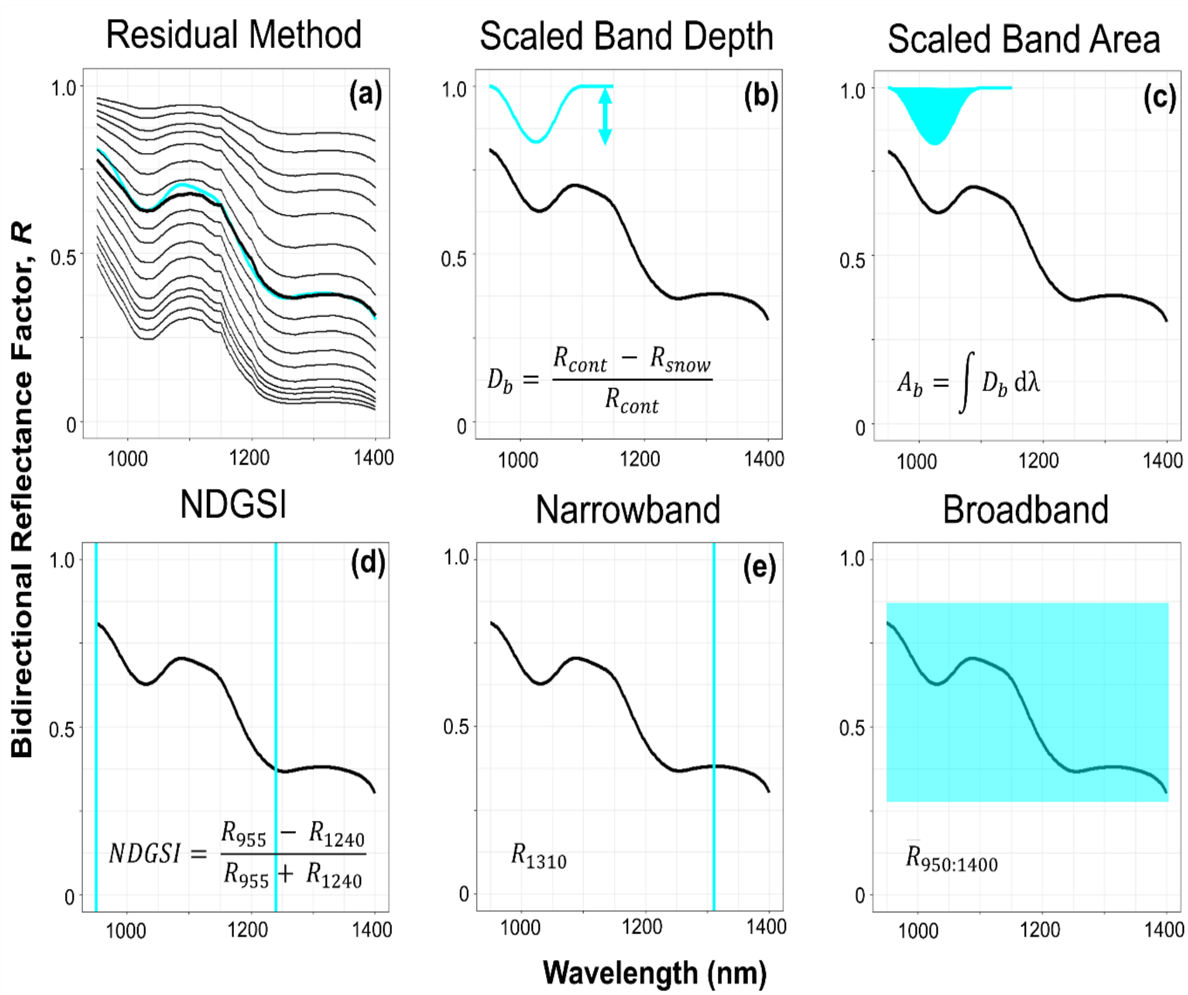

Figure 7: Examples and definitions of the six retrieval techniques evaluated.

3 Results

3.1 Shape parameter optimization

3.1.1 TARTES

Beginning with the TARTES spectral library, we calculated samplewise median values of retrieved r_{opt} for each absorption enhancement/asymmetry (*B/g*) parameter combination. To visualize the influence of shape parameters, we extracted the median absolute error (relative to micro-CT r_{opt}) across all samples for each technique and colored the heat map in Fig. 8 by these error values. The optimal shape parameter combinations yielded median absolute error values of 15.5 – 17.2%, varying slightly by retrieval technique, with hyper- and multispectral techniques generally outperforming narrow- and broadband. However, the substantial dependence of median absolute error values on shape parameters highlights the importance of selecting an optimal shape parameter combination.

We can see that, for a given technique, a variety of shape parameter combinations produce reasonable error (i.e., yellow tiles). It appears that interplay between the two shape parameters is an important consideration, and thus the best selection for one shape parameter depends on the value of the other (and, to a lesser extent, on the retrieval technique). Within the heat maps, an interesting, yet predictable, pattern emerges in an inverse relationship between *B* and *g*. As individual grains become more absorptive (via an increase in *B*) accurate results are still achieved by reducing the extent to which grains preferentially scatter forward, hence a decrease in *g*, resulting in a larger portion of the (unabsorbed) light escaping the snowpack. While our results for the D_b retrieval technique are in good agreement with Robledano et al. (2023), for most retrieval techniques our optimal *B*/*g* combinations were closer to the idealized shapes of hexagonal plates, cubes, cuboids, and fractals (discussed further in Sect. 4). Across all retrieval techniques, the median optimal values of *B* and *g* were 1.7 and 0.775, respectively.

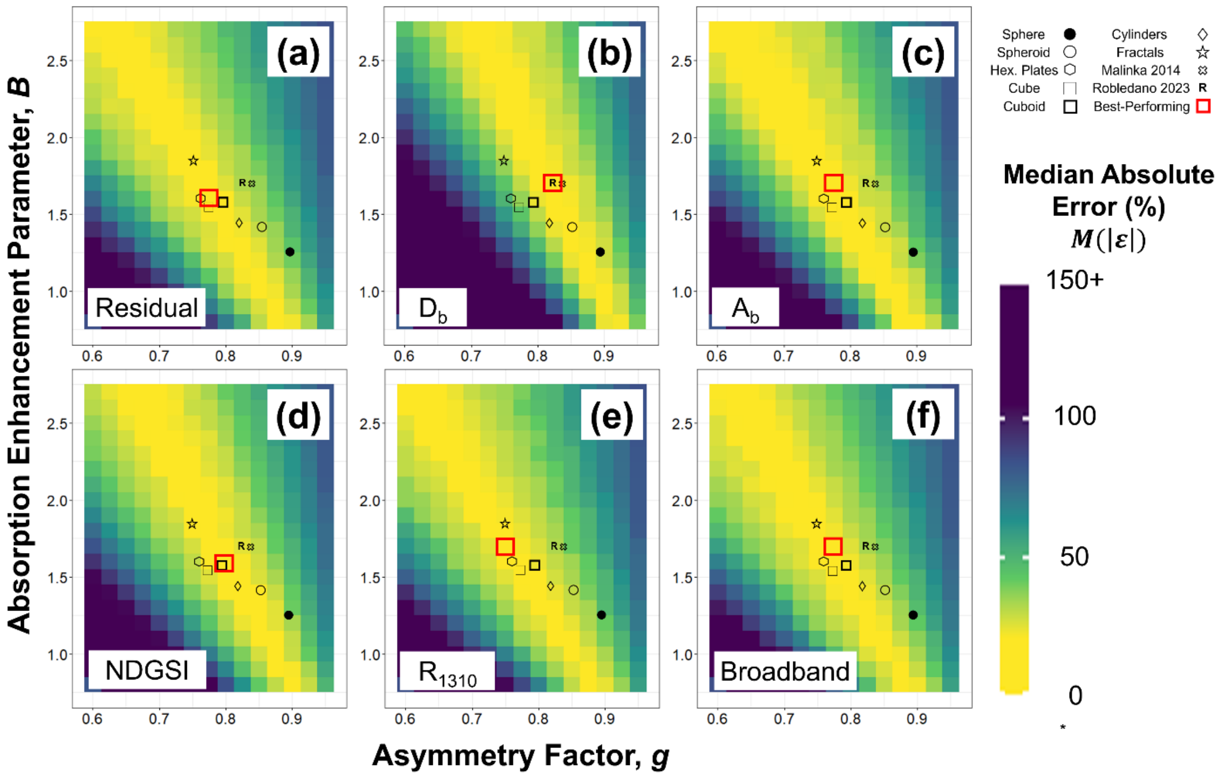

Figure 8: Heat maps depicting median r_{opt} absolute retrieval error for TARTES as a function of shape parameters, across retrieval techniques (a – f). The best-performing combination tile for each technique is boxed in red, while the optimal combination from Robledano et al. (2023) is marked with an "R", as well as other idealized shapes evaluated in their work.

3.1.2 SNICAR

We performed the same heat map optimization analysis on shape parameter combinations in SNICAR (Fig. 9). The optimal shape parameter combinations yielded median absolute error values of 16.5 – 17.7%, values very comparable to TARTES. Again, significant shape dependence and patterns of optimal accuracy are apparent in the heat maps. To reiterate a key point from Sect. 1, unlike TARTES where the effect of shape on both absorption and asymmetry is considered, in SNICAR a spherical assumption is built into the single scattering albedo (and thus *B*). Therefore, altering the combination of shape factor, f_s, and aspect ratio, *A*, is essentially akin to modulating *g*, while the value of *B* stays fixed at that of a sphere (hence 1.25; Fig. 8). However, as we can see in Fig. 8, even for spherical values of *B*, there are corresponding values of *g* that fall within the "stripe" of optimal accuracy in TARTES, and thus it is perhaps unsurprising that certain combinations of f_s/*A* can yield similar retrieval accuracy in SNICAR. The optimal combination was often somewhere between the idealized shapes of spheres, spheroids, hexagonal plates and fractals (Fig. 9) as described by He et al. (2017). Across all retrieval techniques, the median optimal values of f_s and *A* were 0.95 and 2.1, respectively, essentially amounting to an elongated spheroid.

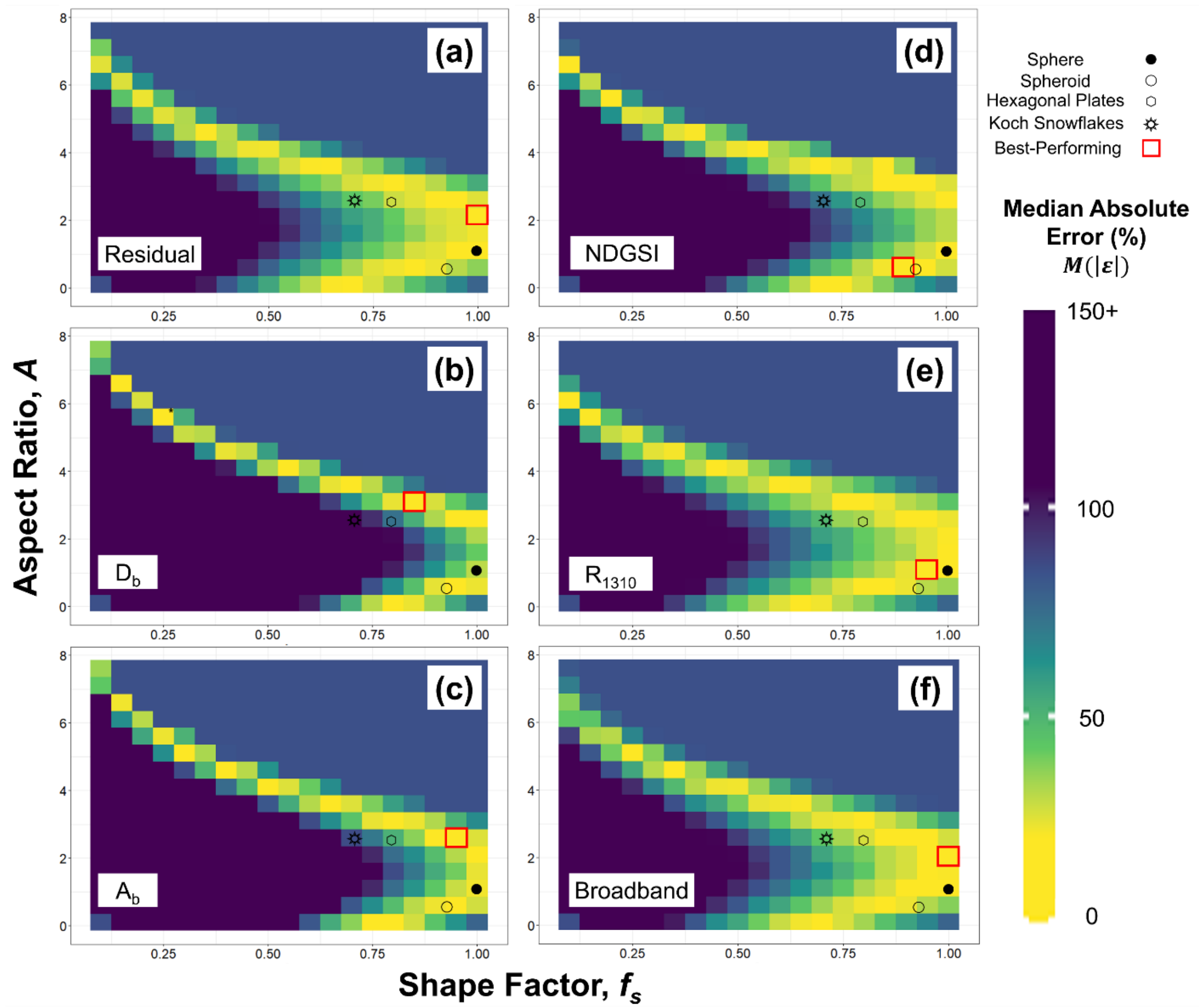

Figure 9: Heat maps depicting median r_{opt} absolute retrieval error for SNICAR as a function of shape parameters, across retrieval techniques (a – f). The best-performing combination tile for each technique is boxed in red, while the locations of idealized shapes from He et al. (2017) are denoted as well.

We can further observe the importance of modeled snow grain shape by comparing samplewise retrievals from SNICAR across the four pre-selected shapes using the residual method (Fig. 10). The modeled shape strongly influences both overall error and variance, with optimized shape parameters (Fig. 10e) outperforming all pre-selected shapes. Even for the optimized case, we can see that it is difficult to correctly retrieve r_{opt} for different measured grain habits (particularly SH, FC, and MF) simultaneously. As expected, some shape parameter combinations fit observed grain habits better than others. Optimized parameters for each model/retrieval technique are used hereafter in Sect. 3.2.

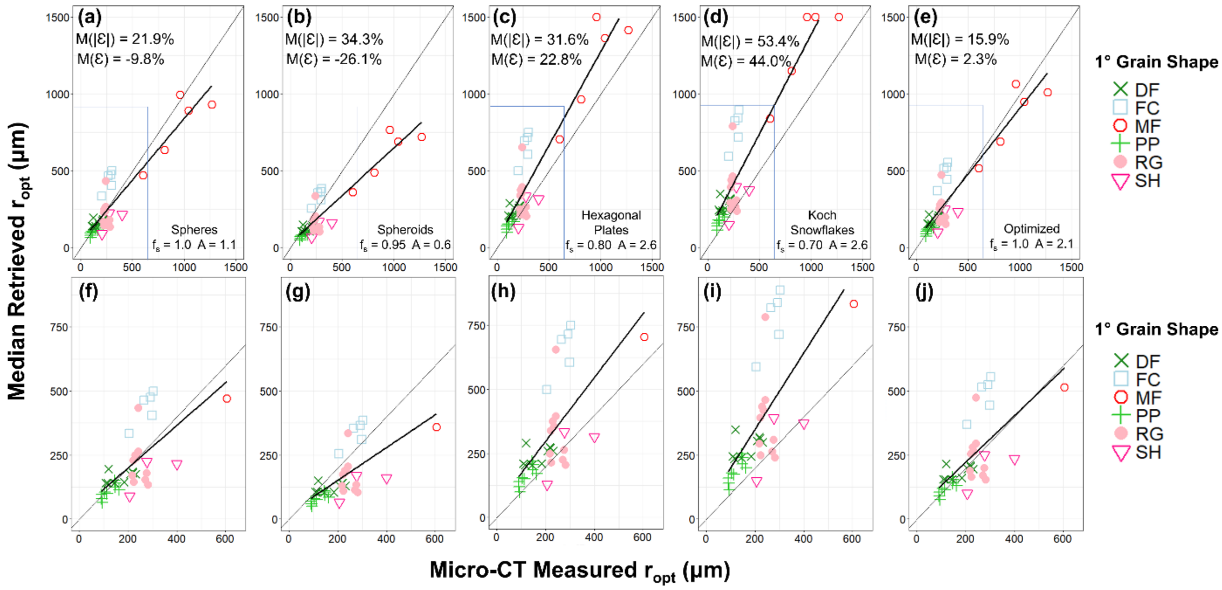

Figure 10: Samplewise median SNICAR retrieved r_{opt} values vs. Micro-CT measured r_{opt} values for different preselected shapes (a – d) and optimized shape parameters (e) using residual method retrieval technique. Grey diagonal lines are a 1:1 reference while the black lines are linear best fits. Point color and style correspond to observed grain habits, following Fierz et al. (2009). The area within the blue rectangles in (a – e) is enlarged in (f – j) with resampled trendlines, given that most samples were clustered at smaller grain sizes relative to the largest MF samples. Median error and absolute error across all samples are listed for each case. For the top row, r = 0.91 in all cases, while r = 0.66 across the bottom row.

3.2 Model and retrieval technique intercomparison

For a given radiative transfer model, we generally observed the most accurate results using retrieval techniques that leverage the more spectral data, with modest reductions associated with techniques using fewer spectral data, demonstrated by the three techniques shown in Fig. 11. This is a predictable result, although it should be noted that techniques using fewer spectral data, such as broadband and narrowband, still performed quite well using certain models (e.g., TARTES and SNICAR in Fig. 11c and 11f). While this success is almost certainly due, in part, to our idealized laboratory setup, it bodes well for applications without hyperspectral capacity as instrument SNRs and calibration techniques continue to improve. Across all retrieval techniques, a similar performance trend is apparent between models: TARTES and SNICAR produced excellent and comparable results, followed by AART, and then the model of Malinka (2014). This result likely highlights the importance of shape optimization for a particular application and/or retrieval technique (Sect. 3.1). In other words, tuning the single scattering/inherent optical properties can be quite useful for boosting accuracy. An example of using the residual method and an optimized TARTES spectral library to create pixelwise r_{opt} maps for different samples is presented in Fig. 12, demonstrating good agreement with micro-CT measurements.

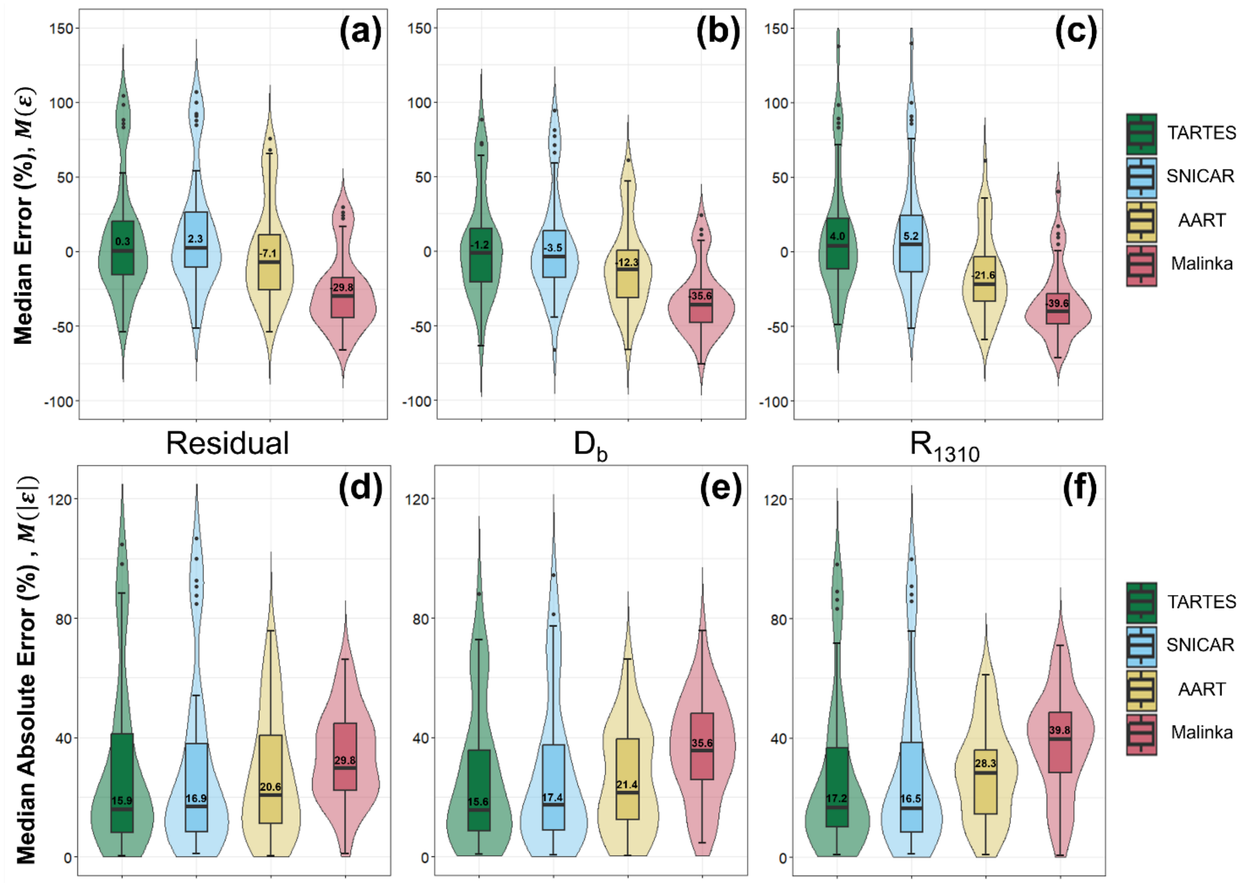

Figure 11: Violin and boxplots demonstrating distributions of samplewise median error (a – c) and absolute error (d – f) across all models for the residual method (left), scaled band depth (middle) and R1310 (right) retrieval techniques.

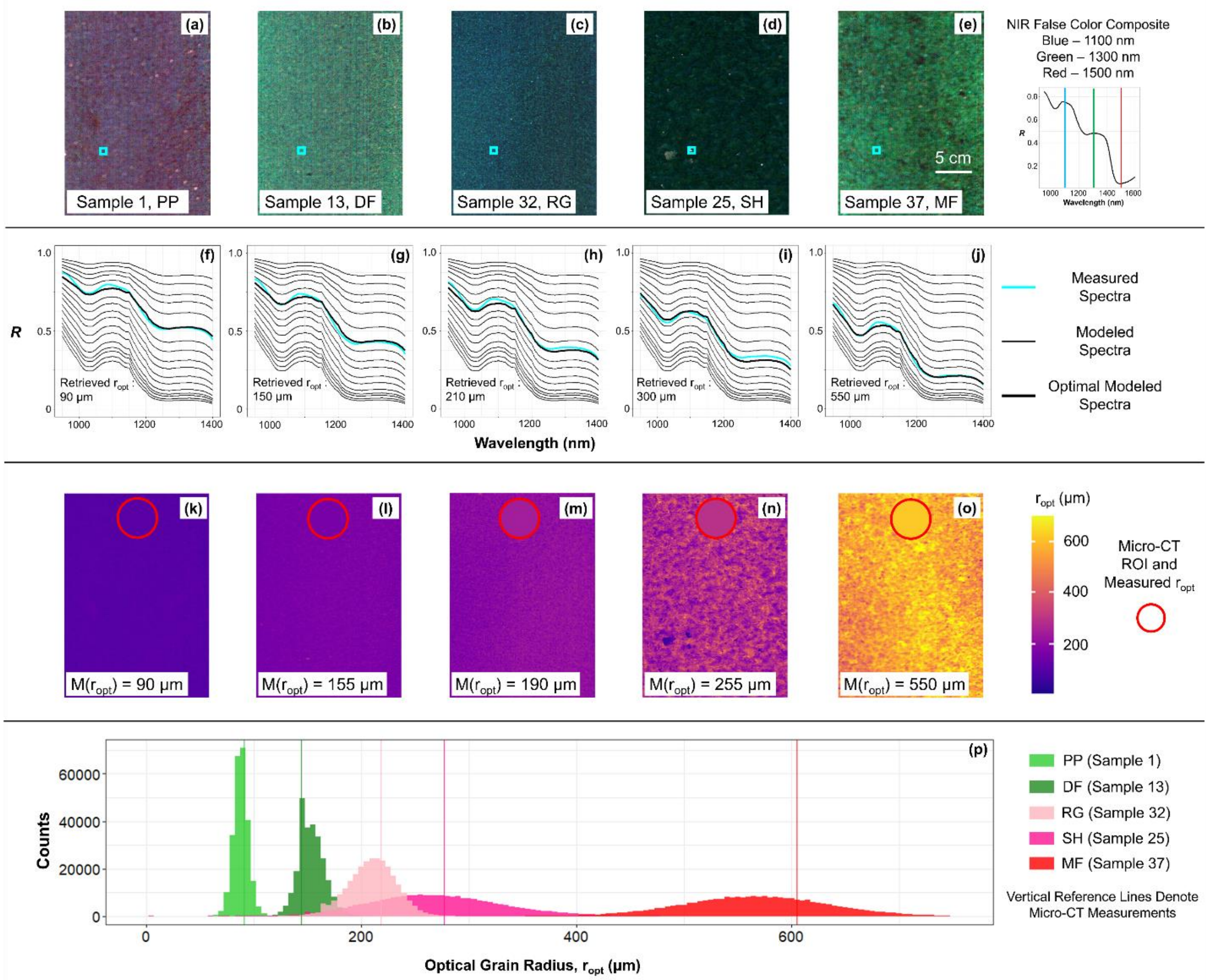

Figure 12: False color composite images for five different snow samples are shown in a – e. The residual method is demonstrated (f – j) on the measured spectra from the cyan pixel in each image. By repeating the process on all pixels, maps of r_{opt} are created and juxtaposed with micro-CT measured values (k – o). Similarly, pixelwise grain size distributions are visualized as histograms compared to vertical micro-CT reference lines (p).

Error metrics for the same three retrieval techniques are grouped by grain habit in Fig. 13. Much like the SNICAR scatterplot in Fig. 10, Fig. 13 demonstrates the difficulty in simultaneously producing accurate retrievals for a wide variety of snow grain habits. As with Fig. 11, we can see that TARTES and SNICAR, after shape optimization, generally perform the best, particularly with PP, DF, RG, and MF. Both AART and particularly the model of Malinka (2014) demonstrated a tendency to consistently underestimate grain size across most grain habits. All models struggled most with samples of a FC or SH primary grain habit, which is perhaps sensible, as chord lengths can vary dramatically in these crystals depending on the angle at which light interacts with the grain. More intriguing is the inconsistent sign of the error. For samples with a FC primary grain habit, all models overestimated grain size, although the model of Malinka (2014) was quite accurate contrary to overall results. Meanwhile, SH samples were globally overestimated. While beyond the scope of this study, it would certainly be possible to perform a similar modeled shape optimization (Sect. 3.1) towards enhancing results for a particular grain habit, if a practitioner had prior knowledge or expectation of what conditions might be encountered. However, our goal here was to optimize results across a wide range of snow microstructures.

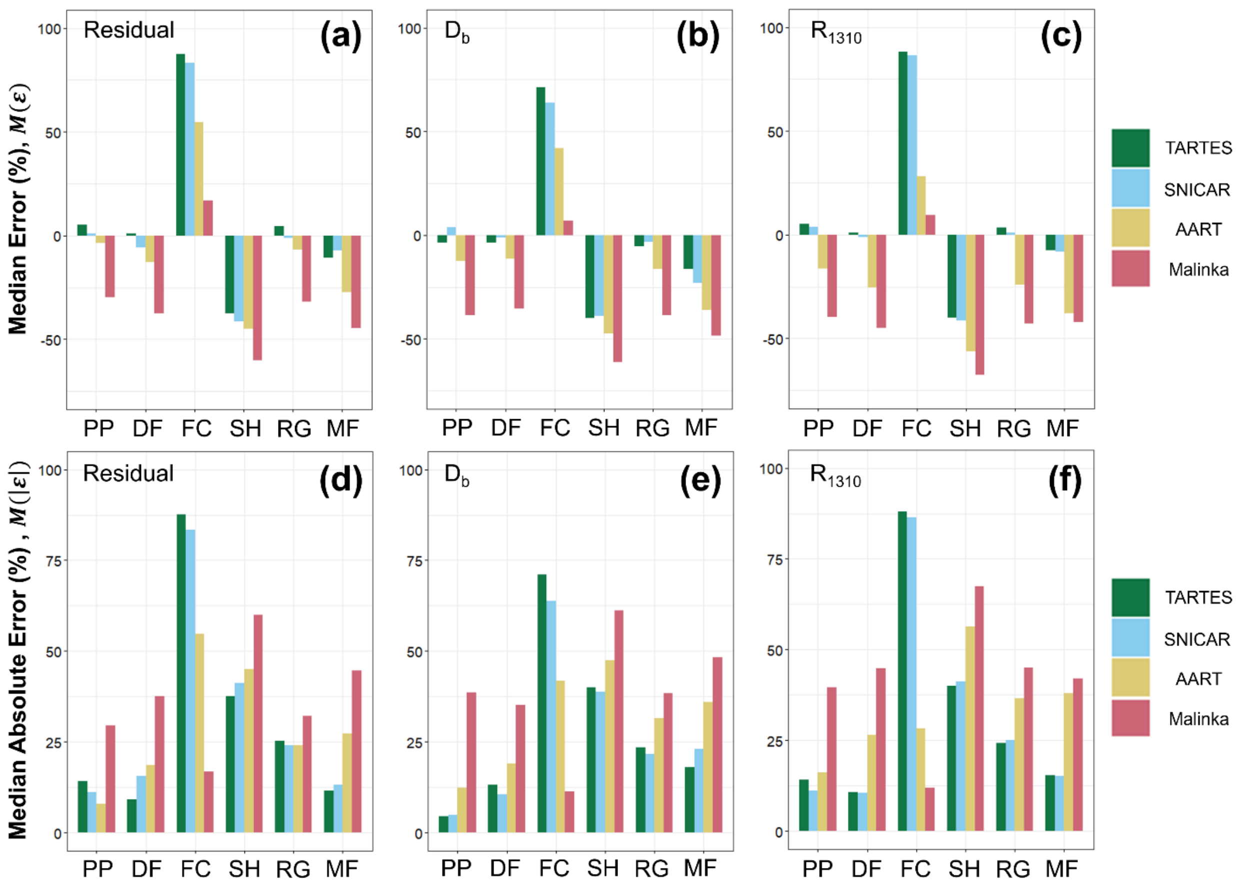

Figure 13: Column graphs grouped by grain habit depicting magnitudes of samplewise median error (a – c) and absolute error (d – f) across all models for the residual method (left), scaled band depth (middle) and R_{1310} (right) retrieval techniques.

4 Discussion

4.1 Shape Optimization

Shape optimization analysis revealed that modeled snow grain shape has a substantial influence on the quality and accuracy of grain size retrievals. For both TARTES and SNICAR, the optimal values of shape parameters were fairly constant across retrieval techniques, with similar patterns emerging (Fig. 8 and Fig. 9). Using SNICAR, optimal values of f_s ranged from 0.85 – 1.00 (M = 0.95), while aspect ratio, A, values varying from unity by a factor of ~2 – 3 proved ideal (e.g., 0.6, 2.1, 2.6, 3.1). Thus, based on our results, the ideal modeled shape for SNICAR is a somewhat flattened, elongated, asymmetrical spheroid. To the best of our knowledge, this is the first study to examine the optimal combination of SNICAR shape parameters.

In contrast to SNICAR, optimization of TARTES shape parameters, the absorption enhancement and asymmetry parameter (B and g, respectively), has received considerable attention in recent years. Libois et al. (2013) suggested $1.6 \leq B \leq 1.9$, further narrowing to the TARTES default of 1.6 in Libois et al. (2014), as they note a wide peak in their retrieved B values from 1.4 – 1.8. The most recent and thorough work on the matter, conducted by Robledano et al. (2023), suggested $B = 1.7$ and $g = 0.82$, describing the optimal modeled shape of snow as, "a collection of convex particles without symmetry…". This estimate of the asymmetry parameter, g, varies from the TARTES default of 0.86 via Meirold-Mautner and Lehning (2004) and from the suggested 0.75 value from Kokhanovsky and Zege (2004); it also sits outside the range of 0.83 – 0.87 found in Libois et al. (2013). We observe asymmetry parameter values on the lower end of these observations. Our optimal r_{opt} retrievals were achieved when running TARTES with $g = 0.750 – 0.825$ ($M = 0.775$) depending on the retrieval technique, thus spanning the values suggested by Kokhanovsky and Zege (2004) and Robledano et al. (2023). Regarding the absorption enhancement parameter, we observed optimal B values ranging from 1.6 – 1.7 depending on the retrieval technique, with a median value of 1.7, in agreement with Robledano et al. (2023) as well as Libois et al. (2014).

For future modeling efforts, we reiterate our median optimal shape parameters as a potential starting point: for SNICAR, $f_s = 0.95$ and $A = 2.1$; for TARTES, $B = 1.7$ and $g = 0.775$. However, there seems to be more to the story than single ideal values. We can observe in Fig. 8 and Fig. 9 that several combinations of shape parameters (for both TARTES and SNICAR) can produce similarly favorable r_{opt} retrievals, and that the interplay between the two variables is most important. Thus, our true recommendation is a similar optimization analysis for each individual application, considering instrument, retrieval technique, etc. Additionally, although certain pairings at extreme values still produced reasonable retrievals (e.g., $B = 2.7$, $g = 0.60$), we caution that these are outside the range of established values from most previous literature, and they may prove unreliable at differing illumination and viewing geometries. Furthermore, as mentioned in Sect. 3.2, it is evident that some modeled grain shapes/shape parameter combinations represent certain grain habits better than others (e.g., Fig. 10, Fig. 13). This finding suggests that a dynamic approach might be useful, where modeled snow grain shape is assigned based on the grain habits that researchers expect to encounter most during measurements, although a priori knowledge would be required for this approach.

4.2 Intercomparison

Though some disagreement between optical retrievals and micro-CT measurements is to be expected, it is imperative to understand how well such techniques compare to true r_{opt} measurements. Such a comparison is especially important considering that broader (airborne and spaceborne) r_{opt} mapping efforts are often validated by local optical retrievals (rather than micro-CT). Considering previous work, Matzl and Schneebeli (2006) found an uncertainty of 15% between SSA estimates from NIR photography and stereological measurements. In Gergely et al. (2014), grain size estimates from the Infrasnow integrating sphere demonstrated agreement within 25% relative to micro-CT based on seven of ten samples. Gallet et al. (2009) were able to estimate SSA with error as low as 10 - 12% using their DUFISS instrument and an empirical reflectance relationship. Donahue et al. (2021) used the scaled band area retrieval technique and a hyperspectral imager to map r_{opt} on a per-pixel basis in a cold laboratory. When comparing mean r_{opt} retrievals to five micro-CT measurements on a semi-homogeneous sample, it was found that micro-CT measurements were 23.9% larger on average. The seminal work of Nolin and Dozier (2000) reports an excellent $r^2 = 0.997$ between measured and retrieved r_{opt} from an airborne platform, though it is difficult to compare results because they do not specify which of their measured grain radii are from stereology versus hand lens. Many of these studies used a spherical modeled grain shape, and they report r_{opt} underestimations similar to those found here when using SNICAR spheres (Fig. 10a, 10f), consistent with many papers discussing the limitations of a spherical assumption (e.g., Kokhanovsky and Zege (2004), Libois et al. (2013), Malinka et al. (2014), Robledano et al. (2023)).

Once optimized shape parameter values were applied, our results depended primarily on the radiative transfer model used, and, to a lesser extent, on the retrieval technique (Fig. 11 and Fig. 13). As discussed in Sect. 3.2, the residual method was the most accurate hyperspectral retrieval technique and often the best overall performer, a sensible result considering the superior amount of spectral data leveraged. However, when using TARTES and SNICAR, excellent results were still achieved with the multispectral, narrowband, and pseudo-broadband techniques (e.g., Fig. 11c and 11f). This is probably due to the consistent illumination source and idealized laboratory condition; scaled absorption feature techniques were primarily introduced to limit uncertainty from varying strength of illumination (Nolin and Dozier, 2000). However, this result is still encouraging for broadband and multispectral applications as instrument SNRs and calibration methods continue to improve.

Regarding models, as mentioned earlier, TARTES and SNICAR performed the best, with median absolute error ranging from 15.6 – 17.4% depending on the retrieval technique, and median error of -3.5 – 5.2%. Thus, our “winning” results are on par with or improved when compared to previous literature, particularly relative to methods with mapping/scalable capacity. The AART model followed, with median error values ranging from 7.1 – 29.8%, and then the model of Malinka (2014), with median error of 29.8 – 45.2%. Though we did not go so far as to hypothesize which models would have the most success, this last result is perhaps surprising, as the novel approach put forth by Malinka seems quite robust. Examining the question of modelled grain shape in terms of chord length distribution is sensible, and the model has been validated by Malinka et al. (2016), and in the more substantial bidirectional reflectance evaluation of Dumont et al. (2021), albeit with only three snow samples spanning two grain habits in the latter. More investigation on this topic is required, as Malinka (2023) points out. For instance, the researcher demonstrates that dense packing in structures like snow, deemed only to influence light penetration depth in traditional snow radiative transfer modeling, may also result in a reduction in reflectance and albedo that has not been considered. Regardless, the efforts presented here constitute one of the most thorough comparisons between optical retrievals and micro-CT data to date. Our success highlights the importance of considering model selection, shape optimization, and retrieval technique, as well as interactions between these factors.

5 Conclusions

Our research demonstrates a novel intercomparison between radiative transfer models, modeled snow grain shapes, and retrieval techniques, towards mapping snow optical grain size. In essence, we found that:

i. Shape parameter combinations of $f_s = 0.95/A = 2.1$ and $B = 1.7/g = 0.775$ performed best for SNICAR and TARTES, respectively. However, operation-specific shape optimization would be ideal.
ii. Regarding retrieval techniques, the hyperspectral residual method performed best. Multispectral, narrowband, and "broadband" retrieval techniques produced accuracy comparable to hyperspectral techniques when using certain models, although this result should be viewed with caution given our idealized laboratory setup.
iii. Concerning radiative transfer models, SNICAR and TARTES (after shape-optimization) generally outperformed AART and the model of Malinka (2014), likely due largely to their prescribed shapes.
iv. In general, the appropriate combination of instrument, retrieval technique, and model/shape parameters is imperative.

As NIR-HSI and other NIR detectors become more economical, and as their spatial and temporal resolution become more robust, the findings presented here may provide guidance for enhanced r_{opt} (and thus snow albedo) mapping. Extending the work presented here to field operations will have immediate implications for Earth surface energy balance estimates and subsequent impacts on climate, hydrological, and even avalanche forecasting.

6 Appendix A

Results for the narrowband R_{1064} alternative retrieval technique are presented below. Shape optimization results for both TARTES and SNICAR are presented in Fig. A1, while overall and grain habit-wise error metrics are presented in Fig. A2. Optimized parameters and accuracy for the R_{1064} retrieval technique were comparable to all other retrieval techniques presented in the main text. However, interestingly, the TARTES and SNICAR models performed slightly worse at 1064 nm relative to all other retrieval techniques, while AART and the model of Malinka (2014) demonstrated their best results. It is our hope that these results can eventually be used for comparison with grain size retrievals from 1064 nm lidar.

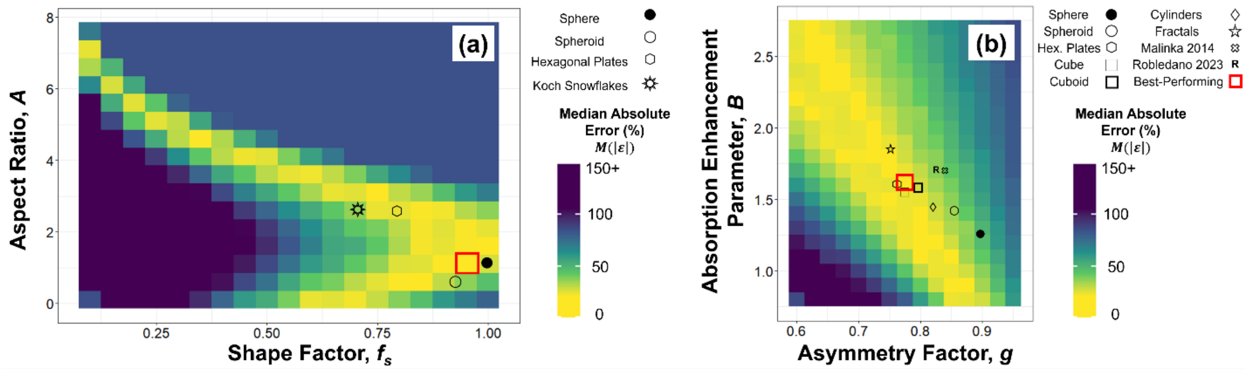

Figure A1: Heat maps depicting median r_{opt} absolute retrieval error for SNICAR (a) and TARTES (b) as a function of shape parameters for the R_{1064} retrieval technique. The best-performing combination tile for each technique is boxed in red. For TARTES, the optimal combination from Robledano et al. (2023) is marked with an "R", as well as other idealized shapes evaluated in their work. For SNICAR, the locations of idealized shapes from He et al. (2017) are denoted.

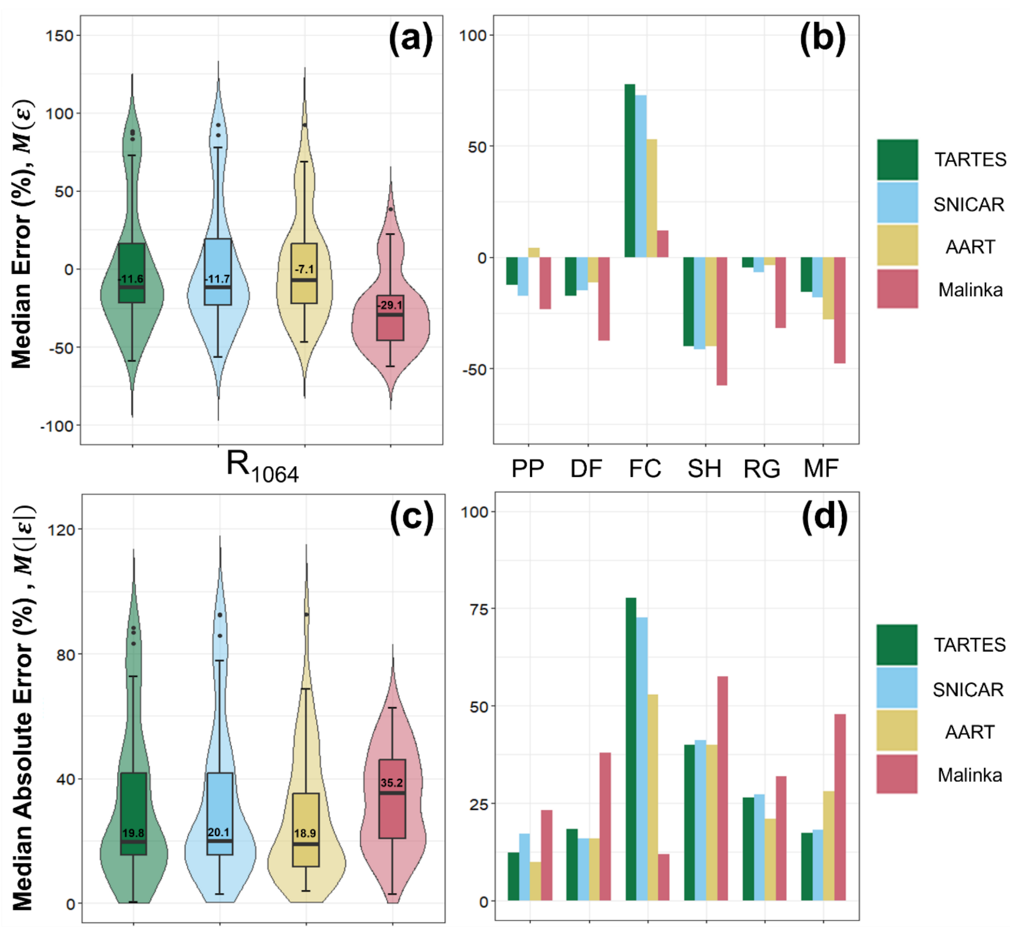

Figure A2: Violin and column graph plots depicting samplewise median error (a and b) and absolute error (c and d) for the R_{1064} retrieval technique.

7 References

Bair, E. H., Stillinger, T., & Dozier, J. (2020). Snow Property Inversion from Remote Sensing (SPIReS): A generalized multispectral unmixing approach with examples from MODIS and Landsat 8 OLI. *IEEE Transactions on Geoscience and Remote Sensing*, *59*(9), 7270-7284.

Bhandari, A., Hamre, B., Frette, Ø., Zhao, L., Stamnes, J. J., & Kildemo, M. (2011). Bidirectional reflectance distribution function of Spectralon white reflectance standard illuminated by incoherent unpolarized and plane-polarized light. Applied Optics, 50(16), 2431-2442.

Bühler, Y., Meier, L., & Ginzler, C. (2014). Potential of operational high spatial resolution near-infrared remote sensing instruments for snow surface type mapping. *IEEE Geoscience and Remote Sensing Letters*, *12*(4), 821-825.Clark and Roush, 1984

Deems, J. S., Painter, T. H., & Finnegan, D. C. (2013). Lidar measurement of snow depth: a review. *Journal of Glaciology*, *59*(215), 467-479.

Dillon, J., Donahue, C., Schehrer, E., Birkeland, K., & Hammonds, K. (2024). Mapping surface hoar from near-infrared texture in a laboratory. *EGUsphere*, *2024*, 1-36.

Donahue, C., Skiles, S. M., & Hammonds, K. (2021). In situ effective snow grain size mapping using a compact hyperspectral imager. *Journal of Glaciology*, *67*(261), 49-57.

Donahue, C., Skiles, S. M., & Hammonds, K. (2022a). Mapping liquid water content in snow at the millimeter scale: an intercomparison of mixed-phase optical property models using hyperspectral imaging and in situ measurements. *The Cryosphere*, *16*(1), 43-59.

Donahue, C., & Hammonds, K. (2022b). Laboratory observations of preferential flow paths in snow using upward-looking polarimetric radar and hyperspectral imaging. *Remote Sensing*, *14*(10), 2297.

Donahue, C. P., Menounos, B., Viner, N., Skiles, S. M., Beffort, S., Denouden, T., ... & Heathfield, D. (2023). Bridging the gap between airborne and spaceborne imaging spectroscopy for mountain glacier surface property retrievals. *Remote Sensing of Environment*, *299*, 113849.

Dumont, M., Brissaud, O., Picard, G., Schmitt, B., Gallet, J. C., & Arnaud, Y. (2010). High-accuracy measurements of snow Bidirectional Reflectance Distribution Function at visible and NIR wavelengths–comparison with modelling results. *Atmospheric Chemistry and Physics*, *10*(5), 2507-2520.

Dumont, M., Flin, F., Malinka, A., Brissaud, O., Hagenmuller, P., Lapalus, P., ... & Ando, E. (2021). Experimental and model-based investigation of the links between snow bidirectional reflectance and snow microstructure. *The Cryosphere*, *15*(8), 3921-3948.

Fierz, C. R. L. A., Armstrong, R. L., Durand, Y., Etchevers, P., Greene, E., McClung, D. M., ... & Sokratov, S. A. (2009). The international classification for seasonal snow on the ground.

Flanner, M. G., Liu, X., Zhou, C., Penner, J. E., & Jiao, C. (2012). Enhanced solar energy absorption by internally-mixed black carbon in snow grains. *Atmospheric Chemistry and Physics*, *12*(10), 4699-4721.

Flanner, M. G., Arnheim, J. B., Cook, J. M., Dang, C., He, C., Huang, X., ... & Zender, C. S. (2021). SNICAR-ADv3: a community tool for modeling spectral snow albedo. *Geoscientific Model Development*, *14*(12), 7673-7704.

Flanner, M. G., & Zender, C. S. (2005). Snowpack radiative heating: Influence on Tibetan Plateau climate. *Geophysical research letters*, *32*(6).

Gallet, J. C., Domine, F., Zender, C. S., & Picard, G. (2009). Measurement of the specific surface area of snow using infrared reflectance in an integrating sphere at 1310 and 1550 nm. *The Cryosphere*, *3*(2), 167-182.Gergely et al., 2014

Grenfell, T. C., & Warren, S. G. (1999). Representation of a nonspherical ice particle by a collection of independent spheres for scattering and absorption of radiation. *Journal of Geophysical Research: Atmospheres*, *104*(D24), 31697-31709.

He, C., Takano, Y., Liou, K. N., Yang, P., Li, Q., & Chen, F. (2017). Impact of snow grain shape and black carbon–snow internal mixing on snow optical properties: Parameterizations for climate models. *Journal of Climate*, *30*(24), 10019-10036.

Kokhanovsky, A. A., & Zege, E. P. (2004). Scattering optics of snow. *Applied optics*, *43*(7), 1589-1602.

Legagneux, L., Cabanes, A., & Dominé, F. (2002). Measurement of the specific surface area of 176 snow samples using methane adsorption at 77 K. *Journal of Geophysical Research: Atmospheres*, *107*(D17), ACH-5.

Li, D., Zhao, M. H., Garra, J., Kolpak, A. M., Rappe, A. M., Bonnell, D. A., & Vohs, J. M. (2008). Direct in situ determination of the polarization dependence of physisorption on ferroelectric surfaces. *nature materials*, *7*(6), 473-477.

Libois, Q., Picard, G., France, J. L., Arnaud, L., Dumont, M., Carmagnola, C. M., & King, M. D. (2013). Influence of grain shape on light penetration in snow. *The Cryosphere*, *7*(6), 1803-1818.

Libois, Q., Picard, G., Dumont, M., Arnaud, L., Sergent, C., Pougatch, E., ... & Vial, D. (2014). Experimental determination of the absorption enhancement parameter of snow. *Journal of Glaciology*, *60*(222), 714-724.

Lorensen, W. E., & Cline, H. E. (1987). Marching cubes: A high resolution 3D surface construction algorithm. *ACM siggraph computer graphics*, *21*(4), 163-169.

Malinka, A. V. (2014). Light scattering in porous materials: Geometrical optics and stereological approach. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *141*, 14-23.

Malinka, A., Zege, E., Heygster, G., & Istomina, L. (2016). Reflective properties of white sea ice and snow. *The Cryosphere*, *10*(6), 2541-2557.

Marks, D., & Dozier, J. (1992). Climate and energy exchange at the snow surface in the alpine region of the Sierra Nevada: 2. Snow cover energy balance. *Water Resources Research*, *28*(11), 3043-3054.

Matzl, M., & Schneebeli, M. (2006). Measuring specific surface area of snow by near-infrared photography. *Journal of Glaciology*, *52*(179), 558-564.

Meirold-Mautner, I., & Lehning, M. (2004). Measurements and model calculations of the solar shortwave fluxes in snow on Summit, Greenland. *Annals of Glaciology, 38,* 279-284.

Murphy, A. B. (2006). Modified Kubelka–Munk model for calculation of the reflectance of coatings with optically-rough surfaces. *Journal of Physics D: Applied Physics*, *39*(16), 3571.

Nolin, A. W., & Dozier, J. (2000). A hyperspectral method for remotely sensing the grain size of snow. *Remote sensing of Environment*, *74*(2), 207-216.

Painter, T. H., Molotch, N. P., Cassidy, M., Flanner, M., & Steffen, K. (2007). Contact spectroscopy for determination of stratigraphy of snow optical grain size. *Journal of Glaciology*, *53*(180), 121-127.

Painter, T. H., Rittger, K., McKenzie, C., Slaughter, P., Davis, R. E., & Dozier, J. (2009). Retrieval of subpixel snow covered area, grain size, and albedo from MODIS. *Remote Sensing of Environment*, *113*(4), 868-879.

Painter, T. H., Bryant, A. C., & Skiles, S. M. (2012). Radiative forcing by light absorbing impurities in snow from MODIS surface reflectance data. *Geophysical Research Letters*, *39*(17).

Perovich, D. K., & Govoni, J. W. (1992, December). Light reflection from a sea-ice cover during the onset of summer melt. In *Ocean Optics XI* (Vol. 1750, pp. 508-516). SPIE.

Picard, G., Arnaud, L., Domine, F., & Fily, M. (2009). Determining snow specific surface area from near-infrared reflectance measurements: Numerical study of the influence of grain shape. *Cold Regions Science and Technology*, *56*(1), 10-17.

Robledano, A., Picard, G., Dumont, M., Flin, F., Arnaud, L., & Libois, Q. (2023). Unraveling the optical shape of snow. *Nature Communications*, *14*(1), 3955.

Sassen, K. (2005). Polarization in lidar. In *Lidar* (pp. 19-42). Springer, New York, NY.

Seidel, F. C., Rittger, K., Skiles, S. M., Molotch, N. P., & Painter, T. H. (2016). Case study of spatial and temporal variability of snow cover, grain size, albedo and radiative forcing in the Sierra Nevada and Rocky Mountain snowpack derived from imaging spectroscopy. *The Cryosphere*, *10*(3), 1229-1244.

Skiles, S. M., Painter, T. H., Deems, J. S., Bryant, A. C., & Landry, C. C. (2012). Dust radiative forcing in snow of the Upper Colorado River Basin: 2. Interannual variability in radiative forcing and snowmelt rates. *Water Resources Research*, *48*(7).

Skiles, S. M., Flanner, M., Cook, J. M., Dumont, M., & Painter, T. H. (2018). Radiative forcing by light-absorbing particles in snow. *Nature Climate Change*, *8*(11), 964-971.

Skiles, S. M., Donahue, C. P., Hunsaker, A. G., & Jacobs, J. M. (2023a). UAV hyperspectral imaging for multiscale assessment of Landsat 9 snow grain size and albedo. *Frontiers in Remote Sensing*, *3*, 1038287.

Stamnes, K., Tsay, S. C., Wiscombe, W., & Jayaweera, K. (1988). Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media. *Applied optics*, *27*(12), 2502-2509.

Toon, O. B., McKay, C. P., Ackerman, T. P., & Santhanam, K. (1989). Rapid calculation of radiative heating rates and photodissociation rates in inhomogeneous multiple scattering atmospheres. *Journal of Geophysical Research: Atmospheres*, *94*(D13), 16287-16301.Warren, 1982

Warren, S. G. (1984). Impurities in snow: Effects on albedo and snowmelt. *Annals of Glaciology*, *5*, 177-179.

Warren, S. G., & Brandt, R. E. (2008). Optical constants of ice from the ultraviolet to the microwave: A revised compilation. *Journal of Geophysical Research: Atmospheres*, *113*(D14).

Wiscombe, W. J., & Warren, S. G. (1980). A model for the spectral albedo of snow. I: Pure snow. *Journal of Atmospheric Sciences*, *37*(12), 2712-2733.

Yang, Y., Marshak, A., Han, M., Palm, S. P., & Harding, D. J. (2017). Snow grain size retrieval over the polar ice sheets with the Ice, Cloud, and land Elevation Satellite (ICESat) observations. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *188*, 159-164.

CHAPTER THREE

ACCURACY ASSESSMENT OF OPTICAL GRAIN SIZE MAPPING FROM NIR LIDAR REFLECTANCE

Contribution of Authors and Co-Authors

Manuscript in Chapter 3

Author: James Dillon

Contributions: Conceptualized the study, collected laboratory data, analyzed results, and drafted the manuscript.

Co-Author: Christopher Donahue

Contributions: Provided guidance and advised during conceptualization and analysis.

Co-Author: Evan Schehrer

Contributions: Assisted with laboratory sample preparation and data collection.

Co-Author: Nathaniel Field

Contributions: Provided technical guidance regarding optical measurements.

Co-Author: Kevin Hammonds

Contributions: Acquired funding, was responsible for project administration, and supervised throughout.

Manuscript Information

James Dillon, Christopher Donahue, Evan Schehrer, Nathaniel Field, Kevin Hammonds

Status of Manuscript:
X Prepared for submission to a peer-reviewed journal
☐ Officially submitted to a peer-reviewed journal
☐ Accepted by a peer-reviewed journal
☐ Published in a peer-reviewed journal

Accuracy assessment of optical grain size mapping from NIR lidar reflectance

James Dillon[1], Christopher P. Donahue[2,3], Evan Schehrer[1], Nathaniel Field[4], Kevin Hammonds[1]
[1]Department of Civil Engineering, Montana State University, Bozeman, MT, USA
[2]Department of Geography, Earth, and Environmental Sciences, University of Northern British Columbia, Prince George, British Columbia, Canada
[3]Hakai Institute, Campbell River, Quadra Island, British Columbia V9W 0B7, Canada
[4]Department of Electrical Engineering, Montana State University, Bozeman, MT, USA

Abstract. The near-infrared (NIR) albedo of snow is dictated by the optical snow grain size (r_{opt}). Therefore, characterizing the spatial variability and temporal change in r_{opt} at the snow surface is critical. Lidar is capable of measuring NIR reflectance and subsequently can be used to retrieve r_{opt}, with significant advantages over passive sensors. However, the accuracy of using lidar in this manner has not been thoroughly evaluated. To address this knowledge gap, we conducted a laboratory reflectance study to determine if lidar is an adequate tool for r_{opt} mapping using traditional radiative transfer inversion techniques. Across forty-one snow samples of widely varied microstructure, we collected reflectance data at eight incidence angles ranging from nadir to oblique and compared subsequent r_{opt} retrievals against micro-CT benchmarks. Retrievals of the largest MF grains were dramatically underestimated, which we hypothesize is due to a gain present when lidar spot size is similar to grain diameter. Upon excluding these samples, nadir results were adequate, with retrievals from the model of Malinka (2014) producing median absolute error of 19.8%. As with passive detection, shape optimization in models that can consider shape parameters proved crucial. Regarding oblique retrievals, estimates were sufficient through incidence angles of 30°, but rapidly deteriorated thereafter with models failing to predict an observed spike in reflectance at larger angles. We demonstrate how this phenomenon can be compensated for by fitting a multivariate regression model to predict reflectance as a function of grain size and incidence. As lidar becomes more economical, it presents itself as a capable instrument for measuring NIR reflectance and corresponding retrievals of r_{opt}, with significant advantages over passive detectors. The findings presented here may provide guidance for enhanced r_{opt} (and thus snow albedo) mapping.

1 Introduction and background

Snow occupies large portions of Earth's surface and is the most reflective natural surface on the planet (with albedo up to 90%), and thus snow cover plays a key role in Earth's energy balance, climate, and hydrology. Furthermore, varied snow microstructures can produce dramatically different albedo, and this sensitivity is responsible for numerous feedback loops (Flanner et al., 2012). Regarding hydrology, snow albedo drives the timing and magnitude of snowmelt runoff, which is imperative for water forecasting (Marks and Dozier, 1992). Albedo and surface energy data can even be used to aid in avalanche forecasting. Therefore, accurate measurements and modeled estimates of snow albedo, particularly with regards to spatial and temporal variability, are key to understanding future climate, melt rates, water availability, and natural hazards.

Numerous snow radiative transfer models have been developed to predict the reflectance or albedo of snow based on optical conditions and physical snowpack parameters (e.g., Flanner and Zender, 2005; Kokhanovsky and Zege, 2004; Libois et al., 2013; Malinka, 2014; Stamnes et al., 1988). Generally, these models 1) use the optical properties of ice (e.g., Picard et al., 2016; Warren and Brandt, 2008; Warren, 1982,1984) and the theory of light scattering to obtain scattering characteristics of one particle, 2) average them over a particle ensemble, and 3) put them into the radiative transfer equation or an approximation (Malinka 2023). Across the visible wavelengths snow is highly reflective, and albedo is primarily driven by impurities near the snow surface (Skiles et al., 2012, 2018). In contrast, ice is more absorptive in the near-infrared (NIR) wavelength range, and thus the principal driver of snow albedo is the average chord length of ice in the scattering medium. This average chord length is often characterized by the optical grain size, expressed here as a radius r_{opt}. Therefore, characterizing the spatiotemporal variability in r_{opt} at the snow surface is critical for accurately estimating albedo. Using a spherical assumption (Grenfell and Warren, 1999), the optical grain size can be related to the physical snow microstructure through the ice surface per unit mass (Legagneux et al., 2002), or specific surface area, SSA (Eq. 1).

$$r_{opt} = \frac{3}{\rho_{ice} * SSA} \tag{1}$$

In summary, NIR albedo and reflectance are controlled by the optical snow grain size, and specifically there exists an inverse relationship between the two: as grain size increases, the albedo decreases due to intensified ice absorption. This relationship is the basis from which snow reflectance measurements can be used to retrieve estimates of r_{opt}. A common practice is to simulate snow reflectance for a wide range of r_{opt} values using a radiative transfer model, creating a library termed a "lookup table", to which measured reflectance values can be compared. By matching the measured reflectance to the closest simulated value or spectra, a corresponding grain size value can be retrieved, which can then be used to model albedo, forecast runoff, etc. Over the last several decades, numerous methods have been developed to retrieve r_{opt} from measured reflectance using different radiative transfer models, retrieval techniques, and instruments. These efforts range from in situ (e.g., Donahue et al., 2021, 2022a; Gallet et al., 2009; Gergely et al., 2013; Matzl and Schneebeli, 2006; Painter et al., 2007) to airborne platforms (e.g., Donahue et al., 2023; Nolin and Dozier, 2000; Seidel et al., 2016; Skiles et al., 2023) to spaceborne sensors (e.g., Bair et al., 2020; Bohn et al., 2021; Painter et al., 2009, 2012). However, these undertakings have overwhelmingly utilized passive detection as opposed to active remote sensing.

The use of Light Detection and Ranging (lidar) in the cryosphere has almost exclusively been used as a spatial data tool for high-resolution snow depth mapping (Deems et al., 2013). However, lidar has the potential to be an ideal instrument for measuring NIR reflectance and subsequent retrievals of r_{opt}. Lidars operate by emitting rapid pulses of light and record both the relative strength of backscattered signal after reflecting off a target, known as the return intensity, as well as the target's range. Each data point, or "return", has spatial coordinates (X, Y and Z) and a return intensity, which can be converted to a reflectance factor with calibration. Thus, lidar generates both an elevation and a radiometric data product. Many lidar units operate at a wavelength of 1064 nm, which sits near the center (1030 nm) of a prominent ice absorption feature in the NIR spectra that is commonly leveraged for r_{opt} retrievals (Nolin and Dozier, 2000).

Furthermore, as an active remote sensing system, lidar holds several advantages relative to passive detectors, where often solar irradiance is utilized as an illumination source (Yang et al., 2017). Perhaps most importantly, because lidar generates both reflectance measurements and an elevation data product, the coincident elevation data, or Digital Terrain Model (DTM), can be used for topographic correction. Hence, the DTM is used to determine the true illumination and viewing zenith angles when encountering mountainous terrain, in a manner that minimizes uncertainty. In the absence of a coincident lidar DTM, often only coarse and (temporally) non-coincident DTMs are available, which can substantially increase uncertainty (Donahue et al., 2023). Additionally, when using lidar there are far fewer potentially confounding factors to consider, such as varying strength of illumination due to clouds or atmosphere, solar zenith angle, relative azimuth between source and sensor, diffuse radiation, and variable polarization state of incident irradiance. When deployed from airborne or spaceborne platforms, cloud screening and atmospheric correction are much more circumventable problems when using lidar, and data collection is not constrained by the availability of sunlight. If accurate, lidar-based grain size retrievals have the potential to be a simpler workflow with reduced uncertainty relative to passive imagery (Fig. 1). However, lidar also has some disadvantages, including only measuring at a single wavelength, proper calibration, power requirements that often exceed those of passive detectors, and potential visual safety hazards.

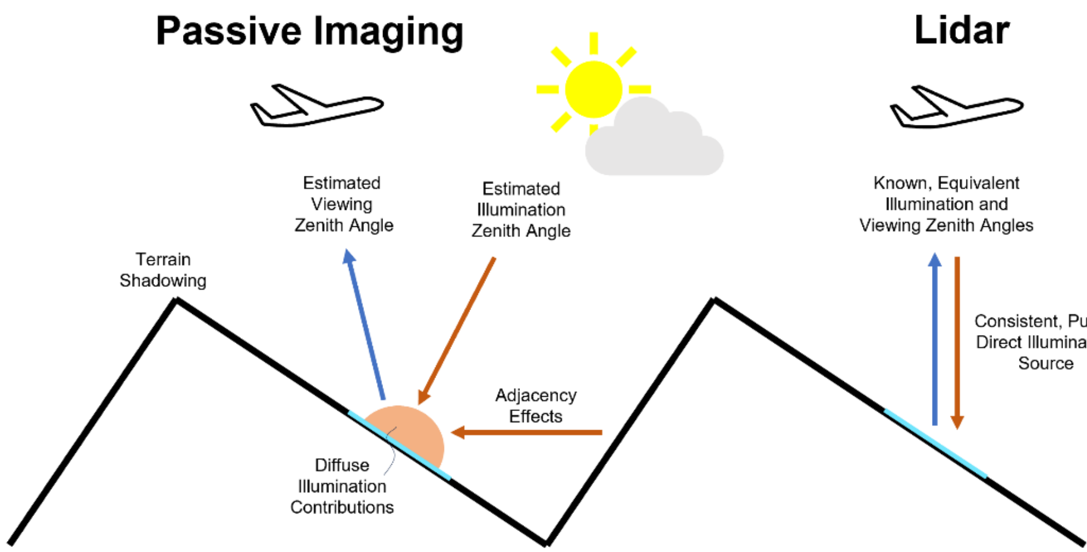

Figure 1: Schematic portrayal of challenges when evaluating bidirectional reflectance from passive imagery, compared to the simpler case of lidar measurements.

To the best of our knowledge, the only study to date that has used lidar to quantify r_{opt} was conducted by Yang et al. (2017). Reflectance observed by the Geoscience Laser Altimeter System aboard the Ice, Cloud, and Land Elevation Satellite was leveraged in conjunction with the Approximate Asymptotic Radiative Transfer model (AART; Kokhanovsky and Zege, 2004) to map r_{opt} over the polar ice sheets. However, limited reference data availability (i.e., ground truth) restricted the researchers to a qualitative assessment of performance. Therefore, the accuracy of using lidar in this manner has not been thoroughly evaluated. This is thought to be a salient point, as lidar scanning represents a unique bidirectional reflectance scenario. For instance, the beam emitted from lidar units is collimated, and thus it experiences less loss due to scattering and absorption compared to direct solar irradiance, and results in higher irradiance at the wavelength of interest. Collimation has been shown to alter scattering and reflectance in optically rough materials like snow relative to an un-collimated illumination source (Murphy, 2006). Additionally, lidar beams are predominately linearly polarized, which contrasts with the often randomly polarized nature of solar illumination (Sassen, 2005). The scattering process is partly polarizing, and therefore multiple scattering can have significant polarization dependence (Li et al., 2008; Bhandari et al., 2011). Last, lidar presents a monostatic geometric condition on bidirectional reflectance, such that every measurement strictly observes direct backscatter. This is beneficial in that it limits the number of characterizations required compared to a full bidirectional reflectance investigation. However, observations of snow reflectance at this geometry are lacking, and radiative transfer models are not well-validated for the case of direct backscatter. Thus, lidar is an optically unique case of bidirectional reflectance that requires careful examination.

To address this knowledge gap, we conducted a laboratory reflectance study to determine if lidar is an adequate tool for r_{opt} mapping using traditional radiative transfer inversion techniques. Towards informing future r_{opt} mapping efforts with additional guidance, we sought to address the following questions:

v. How accurate are lidar-based retrievals?
vi. Does the accuracy deteriorate with lidar illumination/viewing zenith angle?
vii. Which radiative transfer model pairs best with lidar reflectance?
viii. What combination of modeled snow grain shape parameters is the most effective with lidar, and does this differ from previous literature using passive detection?

Thus, we created snow samples across a wide range of grain habits and physical properties, each characterized using microscopy and X-ray computed microtomography (micro-CT). We obtained lidar reflectance data at various angles, and determined subsequent r_{opt} retrievals using numerous radiative transfer models and shape parameter combinations via a method akin to the intercomparison of Dillon et al. (2024b). We analyzed resulting values statistically against reference r_{opt} measurements from micro-CT.

2 Methodology

We prepared laboratory snow samples with a wide range of microstructures, acquired optical measurements, and leveraged radiative transfer modeling to compare r_{opt} retrievals. Section 2.1 describes snow sample preparation and physical characterization, Sect. 2.2 outlines the acquisition of lidar data, and Sect. 2.3 covers radiative transfer modeling and statistical analyses.

2.1 Sample preparation and physical characterization

The samples used here, and thus the methods for sample preparation and physical characterization, are identical to those from Dillon et al. (2024a). Sample creation and characterization are briefly summarized here, but we refer the reader to the aforementioned publication for a full description.

2.1.1 Sample preparation

We prepared samples in Montana State University's Subzero Research Laboratory (SRL). The snow used in these experiments was a combination of crystals produced in the SRL's snowmaking apparatus and natural undisturbed snow from the surrounding area. From twelve batches of differing snow grains, we created forty-one samples by sieving snow through various mesh sizes to further promote disparate microstructures (Table 1). One exception was surface hoar, which we grew following the methods used by Stanton et al. (2016). Sample grain habits included precipitation particles (PP), decomposing and fragmented precipitation particles (DF), rounded grains (RG), melt forms (MF), faceted crystals (FC), depth hoar (DH), and surface hoar (SH) (Fierz et al., 2009). To the best of our ability, we produced samples to be microstructurally homogeneous.

Table 1: Physical snow sample characteristics organized by primary grain habit and listed in order of decreasing surface area-to-volume ratio therein. Adapted from Dillon et al. (2024a) in *The Cryosphere*.

Sample #	Batch ID	Primary Grain Habit	Secondary Grain Habit(s)	Micro-CT SSA (kg m^{-2})	Micro-CT r_{opt} (μm)	Micro-CT ρ (kg m^{-3})	Sieve Size (mm) Passed	Sieve Size (mm) Caught	Notes
1	A	PP	PPrm, DF	35.85	91.3	176	2.38	1.18	
2	A	PP	PPrm, DF	31.60	103.5	217	2.38	-	
3	A	PP	PPrm, DF	28.69	114.0	211	1.18	0.42	
4	B	PP	PPgp	34.67	94.4	160	2.38	1.18	
5	C	PP	DF	36.10	90.6	94	-	-	In situ fresh PP
6	C	PP	DF	22.40	146.1	286	2.38	1.18	
7	C	PP	DF	22.30	146.7	280	0.85	0.42	
8	C	PP	DF	21.94	149.1	275	2.38	-	
9	C	PP	DF	20.05	163.1	303	1.18	0.85	
10	D	DF	RG	29.92	109.3	293	2.38	1.18	
11	D	DF	RG	28.19	116.1	323	0.85	0.42	
12	D	DF	RG	27.38	119.5	351	1.18	0.85	
13	D	DF	RG	22.65	144.4	365	2.38	-	
14	E	DF	DFbk, RGwp	17.74	184.4	374	0.85	-	
15	F	DF	PP	15.69	208.5	322	2.38	-	
16	F	DF	PP	15.04	217.5	312	2.38	1.18	
17	F	DF	PP	14.89	219.8	309	1.18	0.85	
18	F	DF	PP	14.17	230.9	382	0.85	-	
19	G	FC	DH	16.00	204.5	407	1.18	0.42	
20	G	FC	DH	12.34	265.0	448	2.38	1.18	
21	G	FC	DH	11.19	292.4	417	6.3	3.35	
22	G	FC	DH	10.96	298.5	472	6.3	-	
23	G	FC	DH	10.76	304.0	404	3.35	2.38	
24	H	SH	RG	15.83	206.6	213	6.3	-	Re-sieved SH grains
25	H	SH	RG	11.80	277.3	65	-	-	In situ SH atop RGs
26	H	SH	RG	8.18	400.0	94	-	-	Smaller than S25
27	I	RG	DF	14.75	221.7	381	2.38	1.18	
28	I	RG	DF	14.26	229.4	419	1.18	0.85	
29	I	RG	DF	13.93	234.9	431	2.38	-	
30	I	RG	DF	13.56	241.4	489	0.85	0.42	
31	I	RG	DF	13.52	241.9	452	-	-	S29 melt-refreeze
32	J	RG	DF	15.01	218.0	394	0.85	-	
33	J	RG	DF	14.67	223.0	355	0.42	-	
34	J	RG	DF	11.62	281.4	460	1.18	-	
35	K	RG	DF	12.14	269.5	404	1.18	0.85	
36	K	RG	DF	11.82	276.8	428	0.85	0.42	
37	L	MF	RG	5.41	604.8	582	2.38	0.42	
38	L	MF	RG	4.02	813.0	545	6.3	-	
39	L	MF	RG	3.41	958.5	512	3.15	2.38	
40	L	MF	RG	3.14	1041.7	467	-	-	Refrozen in situ
41	L	MF	RG	2.58	1265.8	433	6.3	3.15	

2.1.2 Physical characterization

To begin sample characterization, we first performed microscopy on representative grains from each batch prior to sieving (Fig. 2), and visually classified grain habits following Fierz et al. (2009). After sieving and preparation, we collected micro-CT data from each sample generally following the protocol outlined by Donahue et al. (2021). We used a cylindrical micro-CT sample holder with 3 cm diameter x 4 cm length, which allowed for a voxel size of 14.5 μm, the finest spatial resolution achievable with the relatively large holder. The sample holder was chosen to allow for larger-grained samples (e.g., surface hoar) to be encapsulated and transported to micro-CT for measurement. We note that this will result in neglection or mischaracterization of features smaller than this size. Therefore, it is possible that the SSA of samples with a PP primary grain habit (particularly Samples 1 – 5) was underestimated by micro-CT, and thus r_{opt} overestimated. After scanning, we binarized grey-scale images into ice and air phases (e.g., Fig. 2), and created 3D reconstructions via the marching cubes method (Lorensen and Cline, 1998), which allowed us to determine the SSA (and thus r_{opt}) and density of each sample. Micro-CT r_{opt} values served as benchmarks for comparison to retrievals.

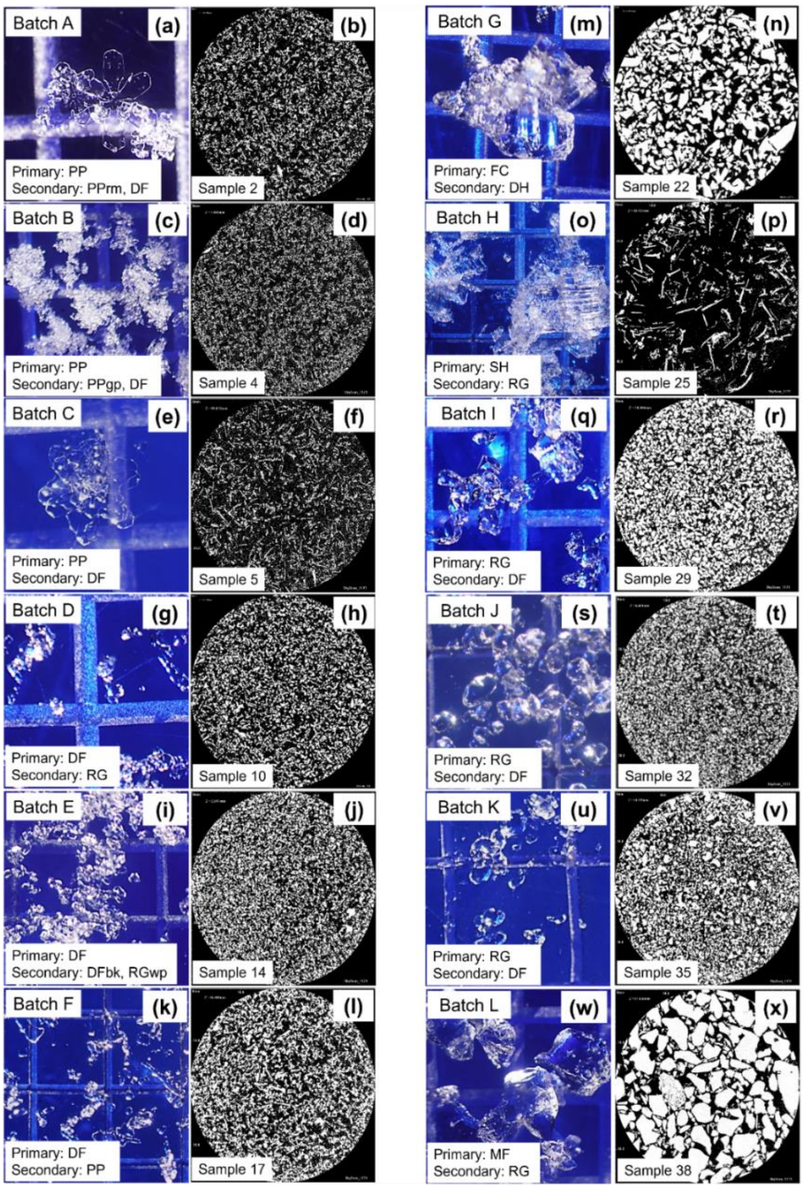

Figure 2: Microscopy images of grains from each initial batch (left columns) and binary micro-CT cross-sections from representative samples (right columns). In the microscopy images, the grid size on the underlying blue grain card is 2 mm. Originally published by *The Cryosphere* in Dillon et al. (2024a).

2.2 Lidar data collection and processing

We used a Riegl VZ-6000 terrestrial laser scanner mounted to a tripod for lidar data acquisition. The scanner achieves vertical (line) scanning via an oscillating mirror while moving horizontally on a rotating head (Fig. 3a). The maximum vertical scan field-of-view (sFOV) is 60° - 120° from zenith and the horizontal sFOV can range from 0° to a full 360° panorama. The laser operates in the NIR range narrowly around a central wavelength of 1064 nm. We set vertical and horizontal angular increments to 0.01° to maximize resolution (point density). The initial laser beam diameter upon exiting the scanner is 15 mm with a divergence of 0.12 mrad, and thus at nadir the incident diameter was essentially still 15 mm. We positioned the scanner about 2.5 m from the snow samples.

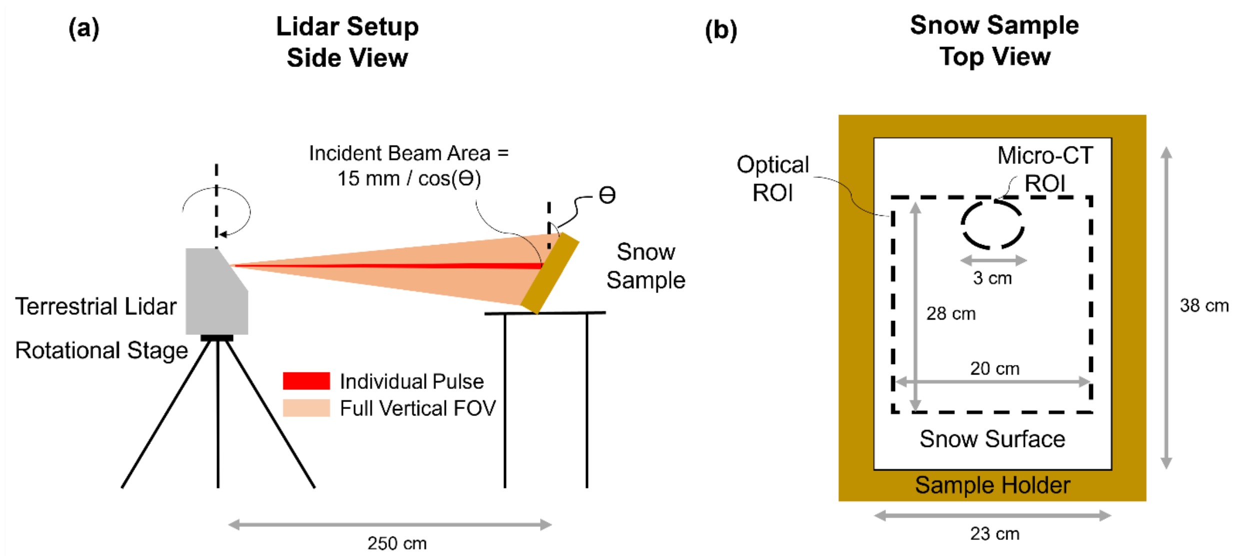

Figure 3: Laboratory data collection schematic for lidar scanning (a). Data regions-of-interest (ROIs) within the snow sample are illustrated in (b). Adapted from Dillon et al. (2024a).

The resulting data product for lidar is a "cloud" of discrete vector data points, called returns. We used a Lambertian reflectance standard to convert return power to bidirectional reflectance factor, R_{1064}, for each individual point in the clouds. As discussed in Sect. 1, a strength of lidar is that it produces a coincident spatial (X, Y, and Z) and radiometric data product; the former can be used to determine the appropriate illumination and viewing zenith angle (equivalent due to the monostatic nature of lidar) for radiative transfer modeling. Hereafter, this illumination and viewing zenith angle for the case of lidar backscatter will simply be referred to as Θ (Fig. 3a). We collected lidar data from Θ = 0 - 70° at 10° increments, beyond which reliable data could not be acquired. Initial processing of point clouds took place in Riegl's RiScan application, and analyses thereafter in Rstudio. We truncated point clouds to a central region-of-interest (ROI) encapsulating the micro-CT ROI (Fig. 3b) to exclude edge effects from mixed points, considering lidar beam dilation when samples were tilted off-nadir. After truncation, the average point cloud contained 80,000 returns at nadir (spacing of ~ 1.4 pts/mm^2), decreasing to about 30,000 returns when $\Theta = 70°$. We acquired all optical data immediately prior to micro-CT analysis at a constant temperature of -10° C.

2.3 Radiative transfer modeling

To model snow reflectance, we utilized four commonly used snow radiative transfer models: the Two-stream Analytical Radiative TransfEr in Snow (TARTES) model (Libois et al., 2013), the SNow, ICe, and Aerosol Radiative – Adding-Doubling (SNICAR-AD) model (Flanner et al., 2021), the Approximate Asymptotic Radiative Transfer (AART) model (Kokhanovsky and Zege, 2004), and the model of Malinka (2014). For each model, we created R_{1064} lookup tables to retrieve r_{opt} from lidar reflectance measurements. Both TARTES and SNICAR are albedo models, and thus they can only be used to approximate bidirectional reflectance at nadir viewing (Dumont et al., 2010). Therefore, these models were only used for the case of $\Theta = 0°$.

Furthermore, TARTES and SNICAR each have two tunable shape parameters which can dramatically vary the inherent single scattering properties and thus the modeled reflectance and subsequent grain size retrievals. In TARTES, the asymmetry parameter, g, which determines how forward-scattering grain are, and the absorption enhancement parameter, B (akin to single scattering albedo), can be tuned. In SNICAR, a spherical assumption is built into absorption as it pertains to grain shape. However, altering the combination of the shape factor, f_s, and the aspect ratio, A, essentially controls g. To further investigate the influence of modeled snow grain shape, we produced numerous lookup tables for both TARTES and SNICAR using modulated combinations of shape parameters, following Dillon et al. (2024b). For TARTES, we evaluated absorption enhancement parameter, B, values from 0.8 – 2.7 at increments of 0.1, and asymmetry factor, g, values from 0.60 – 0.95 at increments of 0.025. These ranges spanned all reasonable values based on previous literature (e.g., Libois et al., 2013, 2014; Robledano et al., 2023). Similarly, for SNICAR we varied the shape parameter (f_s) from 0.1 – 1.0 at 0.05 increments and Aspect Factor (A) from 0.1 – 7.6 with steps of 0.5, again spanning all reasonable values (e.g., He et al., 2017) and nearly the full range selectable values in the model.

For AART and the model of Malinka (2014), the modeled snow grain shape is fixed. However, because these are true bidirectional reflectance models that can consider any illumination and viewing geometry, we were able to model R_{1064} for each angle of Θ to examine off-nadir retrievals. All models were run for r_{opt} values of 30 – 1500 µm at 5 µm increments. Once the lookup tables were created, we matched the lidar R_{1064} measurements with the closest value for a given table, and “retrieved” the corresponding grain size. We repeated this process for all points in each sample across all lookup tables. For the radiative transfer models that can consider shape (TARTES and SNICAR), we identified the optimal shape parameter pairing. We calculated samplewise medians of retrieved r_{opt} and compared them to each other and across models, as well as to reference micro-CT values.

3 Results

3.1 Shape parameter optimization

Beginning with the TARTES modeled reflectance tables, we calculated samplewise median values of retrieved r_{opt} for each absorption enhancement/asymmetry (*B/g*) parameter combination. To visualize the influence of shape parameters, we extracted the median absolute error (relative to micro-CT r_{opt}) across all samples and colored the heat map in Fig. 4a by these error values. The optimal shape parameter combination (B = 1.6, g = 0.85) yielded a median absolute error of 29.5%. These optimal values are in good agreement with Robledano et al. (2023) and most literature. The substantial dependence of median absolute error values on shape parameters highlights the importance of selecting an optimal shape parameter combination. As with the intercomparison of passive detection techniques in Dillon et al. (2024b), we observe that interplay between the two parameters is a salient consideration, and thus the best selection for one shape parameter depends on the value of the other. The inverse trend between B and g is understandable; as individual grains become more absorptive (via an increase in B) accurate results are still achieved by reducing the extent to which grains preferentially scatter forward, hence a decrease in g, resulting in a larger portion of the (unabsorbed) light escaping the snowpack.

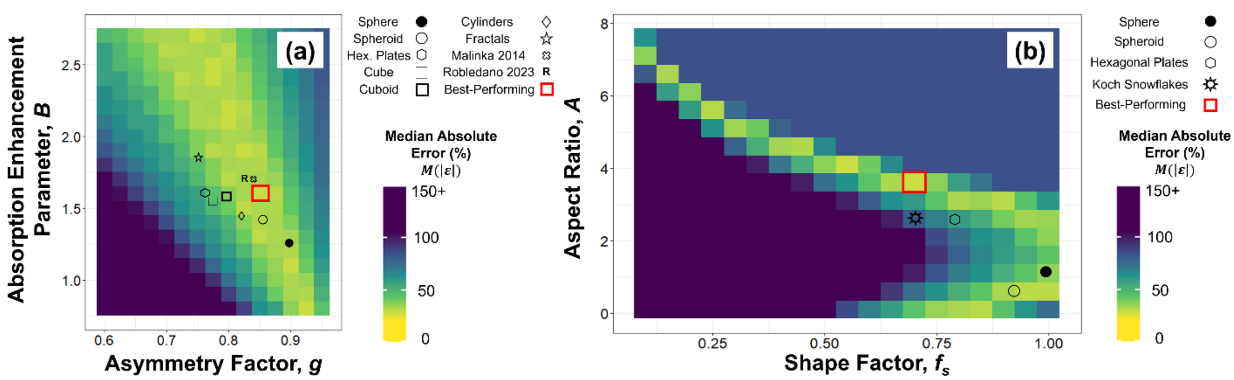

Figure 4: Heat maps depicting median r_{opt} absolute retrieval error for TARTES (a) and SNICAR (b) as a function of shape parameters when using lidar R_{1064} for retrieval. The best-performing combination tile for each technique is boxed in red. For TARTES, the optimal combination from Robledano et al. (2023) is marked with an "R", as well as other idealized shapes evaluated in their work. For SNICAR, the locations of idealized shapes from He et al. (2017) are denoted. The color scales are identical to those from Dillon et al. (2024b) for comparison to passive detection methods.

We performed the same heat map optimization analysis on shape parameter combinations in SNICAR (Fig. 4b). The optimal shape parameter combination yielded median absolute error values of 29.3%, nearly identical to TARTES. Again, significant shape dependence and patterns of optimal accuracy are apparent in the heat maps. The observed "stripe" of maximum accuracy reveals combinations of f_s/A that result in high values of g, offsetting the spherical assumption built into absorption/single scattering albedo in SNICAR (see the location of the idealized sphere in Fig. 4a, where B = 1.25). For more discussion on this, we refer the reader to Dillon et al. (2024b). The optimal shape in SNICAR was essentially an elongated Koch snowflake, with f_s = 0.70 and A = 3.6. Interestingly, this optimal shape factor is considerably lower than the median value of 0.95 determined using passive techniques in Dillon et al. (2024b). Optimized parameters for each model are used hereafter in Sect. 3.2.

3.2 Model comparison

3.2.1 Nadir results

Across all four models, the samplewise median retrieved r_{opt} values were within 20% of micro-CT benchmarks (Fig. 5a), with TARTES, SNICAR, and the model of Malinka (2014) demonstrating a tendency to underestimate grain size, while AART overestimated. Median absolute error values were higher (Fig. 5c), with the model of Malinka (2014) producing the best accuracy, slightly outperforming even the shape-optimized TARTES and SNICAR models. A closer look at median error and absolute error aggregated by primary grain habit (Fig. 5b and 5d, respectively) reveals that all models struggled with SH samples and, especially, MF samples, with global underestimation. The lidar-measured values of R_{1064} for MF samples were substantially larger than any model predicted, and thus retrieved r_{opt} values were much lower than those measured by micro-CT.

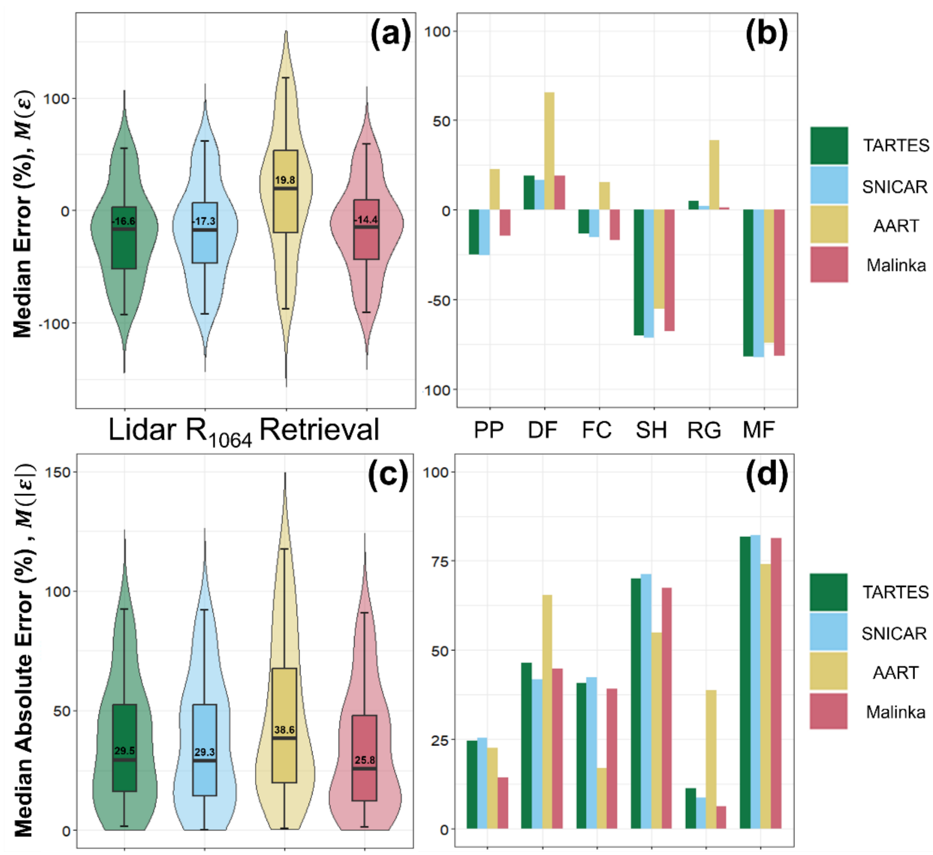

Figure 5: Violin and column graph plots depicting samplewise median error (a and b) and absolute error (c and d) for lidar r_{opt} retrievals using four radiative transfer models.

The difficulty with MF samples is exceptionally pronounced when examining the scatterplot in Fig. 6, and even more perplexing is the negative trend in retrieved r_{opt} as measured r_{opt} increases amongst MF samples. However, in the absence of MFs, the results are adequate (Fig. 6e – 6h). After removing MF samples, the model of Malinka (2014) produced the only case where median absolute error dropped beneath 20%. Hereafter, samples with a MF primary grain habit (S37 – S41) are excluded from lidar results. Uncertainty related to this grain habit is further discussed in Sect. 4. Despite substantial error associated with certain grain habits, we note that these efforts still constitute the novel observation that NIR lidar reflectance is sensitive to snow microstructure, generally in the manner one would predict considering radiative transfer theory. As with passive techniques, this sensitivity can be leveraged with radiative transfer model inversion to map and quantify spatial variability in snow grain size. We demonstrate this mapping capacity on a sample with mixed grain habits in Fig. 7.

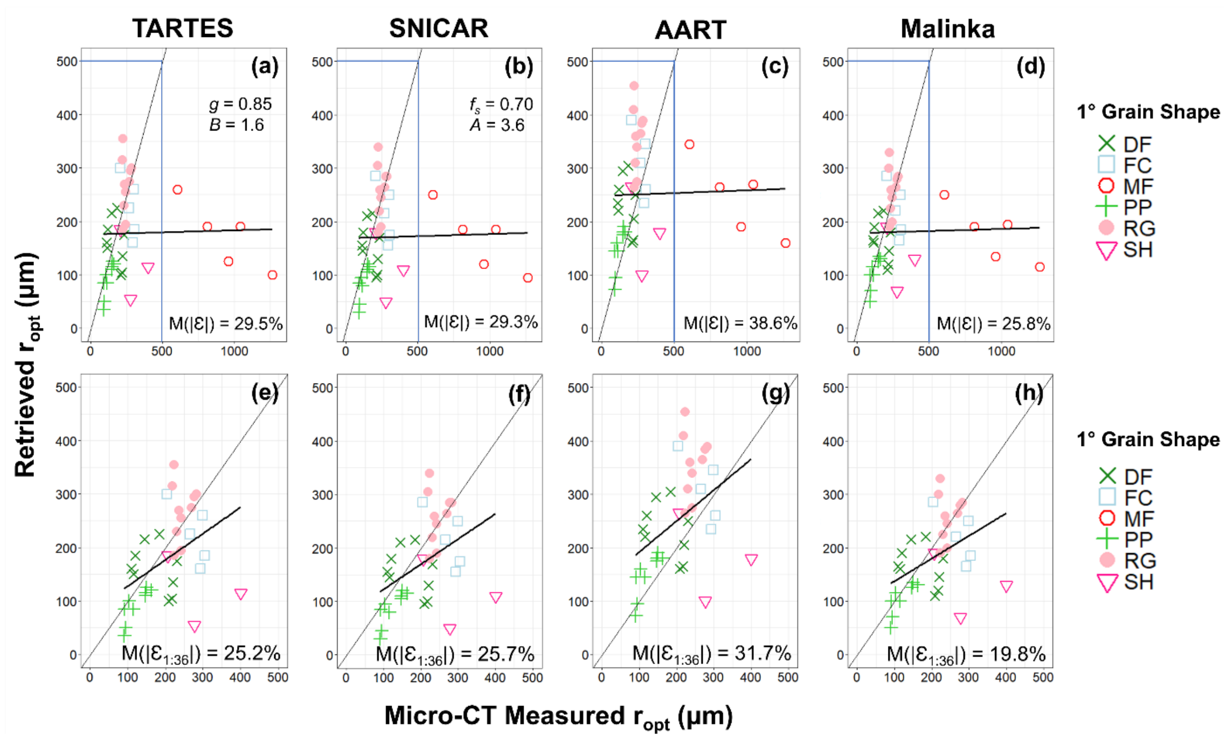

Figure 6: Samplewise median retrieved r_{opt} values at nadir vs. Micro-CT measured r_{opt} values for the four radiative transfer models (a – d) using R_{1064}. Grey diagonal lines are a 1:1 reference while the black lines are linear best fits. Point colors and styles correspond to observed grain habits, following Fierz et al. (2009). The area within the blue rectangles in (a – d) is enlarged in (e – h) given that most samples were clustered at smaller grain sizes relative to the largest MF samples. Median absolute error for a given case is listed on each plot. For the top row, r = 0.04 in all cases, while r = 0.44 across the bottom row.

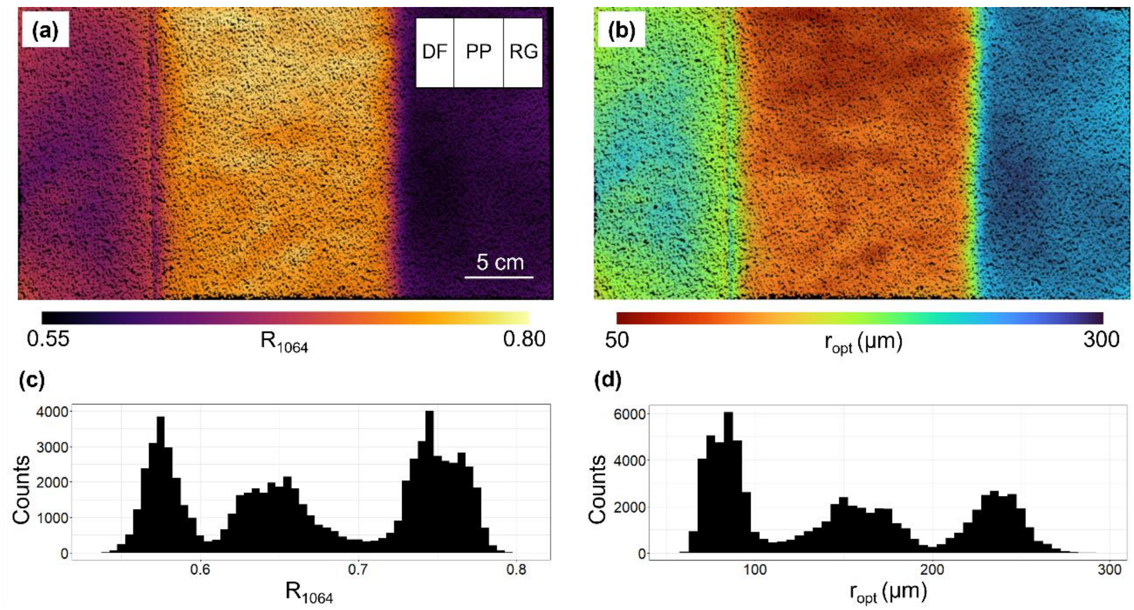

Figure 7: A mixed sample featuring three distinct grain habits to demonstrate grain size mapping from lidar reflectance. In (a) the point cloud is colored by reflectance (R1064), and then converted to pointwise estimates of ropt (b) via the model of Malinka (2014). Associated trimodal distributions are depicted in the histograms in (c) and (d).

3.2.2 Oblique results

Using the two bidirectional, asymptotic models (AART and Malinka, 2014) we examined how increasing the lidar illumination and viewing zenith angle, ϴ, influences accuracy (Fig. 8). Progressing from nadir, $\Theta = 0°$ (Fig. 8a), to $\Theta = 70°$ (Fig. 8h), we can see that accuracy remains relatively constant through $\Theta = 20°$, then plummets at angles upwards of $\Theta = 40°$, with retrievals eventually converging to zero. Although the model of Malinka (2014) performs better at nadir, accuracy deteriorates with increasing ϴ more rapidly than AART. We can further examine this perplexing result by juxtaposing measured reflectance from an individual sample and modeled spectra (Fig. 9). The AART-modeled spectra in Fig. 9a demonstrate that an increase in reflectance (including R_{1064}) is anticipated as values of ϴ increase. For a sample with an identical r_{opt} value, an increase in R_{1064} with ϴ is also observed, but the rate of increase is far steeper (Fig. 9b). Because of this incongruity, modeled values of R_{1064} diverge from measurements at larger angles, resulting in substantial error. The model of Malinka (2014) either predicts a slight decrease (Fig. 9b) or slight increase in R_{1064} with ϴ depending on the grain size, although even in the case of the latter, the same divergence occurs.

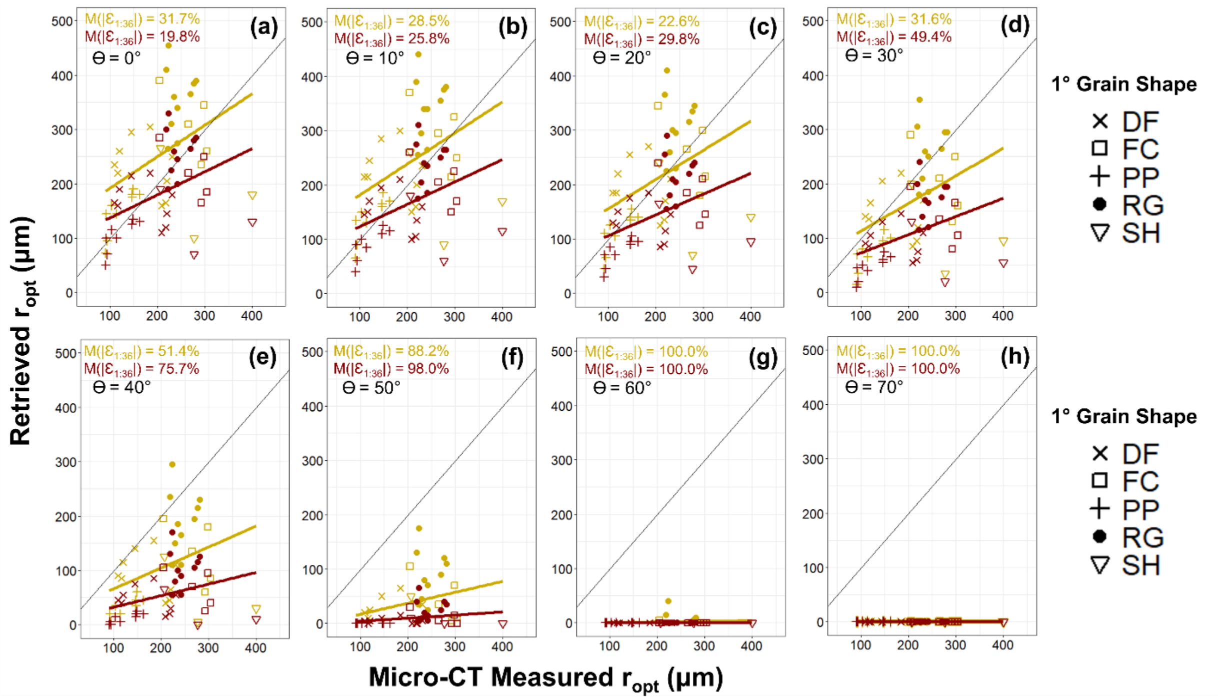

Figure 8: Scatterplots comparing micro-CT measured r_{opt} to retrieved r_{opt} from the AART (gold) and Malinka, 2014 (red) models for different values of Θ. Grey diagonal lines are a 1:1 reference while the gold and red lines are linear best fits. Point styles correspond to observed grain habits, following Fierz et al. (2009). Samples with a MF primary grain habit are excluded.

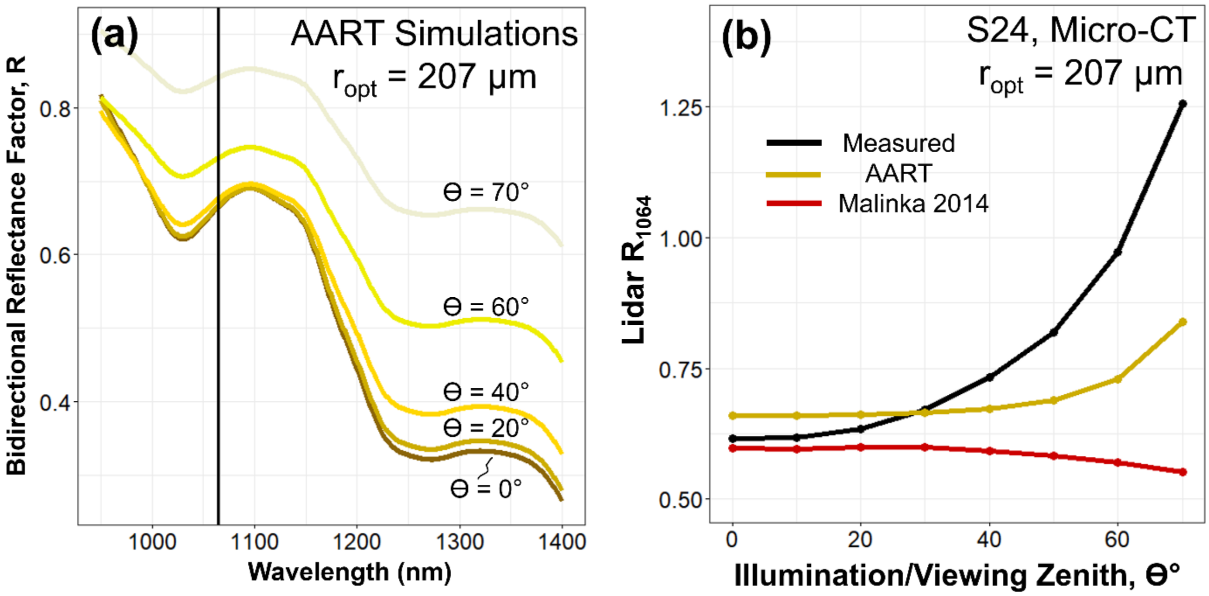

Figure 9: AART-simulated spectra for r_{opt} = 207 µm, with lighter shades of gold corresponding to increasing values of Θ (a). The black vertical reference line sits at 1064 nm. These modeled R_{1064} values (as well as those from the Malinka, 2014 model) are compared to measured R_{1064} values from Sample 24, with a micro-CT measured r_{opt} of 207 µm (b). The observed divergence is the cause of large retrieval error at oblique angles.

The discrepancy between measured and modeled R_{1064} values can further be examined in Fig. 10, along with a possible solution. Both AART (Fig. 10a) and the model of Malinka, 2014 (Fig. 10b) boast reasonable median error values across measurements at all angles (14.1% and 19.3%, respectively) due to excellent accuracy at and near-nadir. However, the "drift" away from the 1:1 line at larger angles of Θ (lighter shades) is evident, even resulting in a negative correlation coefficient for the Malinka model. Because of this limitation, it may be advisable to instead leverage a statistical model when large values of Θ are anticipated; we demonstrate one such example in Fig. 10c. We created a multivariate, polynomial statistical model to estimate R_{1064} as a function of both r_{opt} and Θ. The model replicates in observed increase in R_{1064} with increasing Θ better than either radiative transfer model and yields reduced error. Similarly, one could create multiple univariate models for increments of Θ. While we do not anticipate that this parameterization will remain effective across different instruments, scales, and operations, the fit and accuracy achieved are encouraging. Similar parameterizations could, perhaps, be created for different instruments and applications when large angles are encountered.

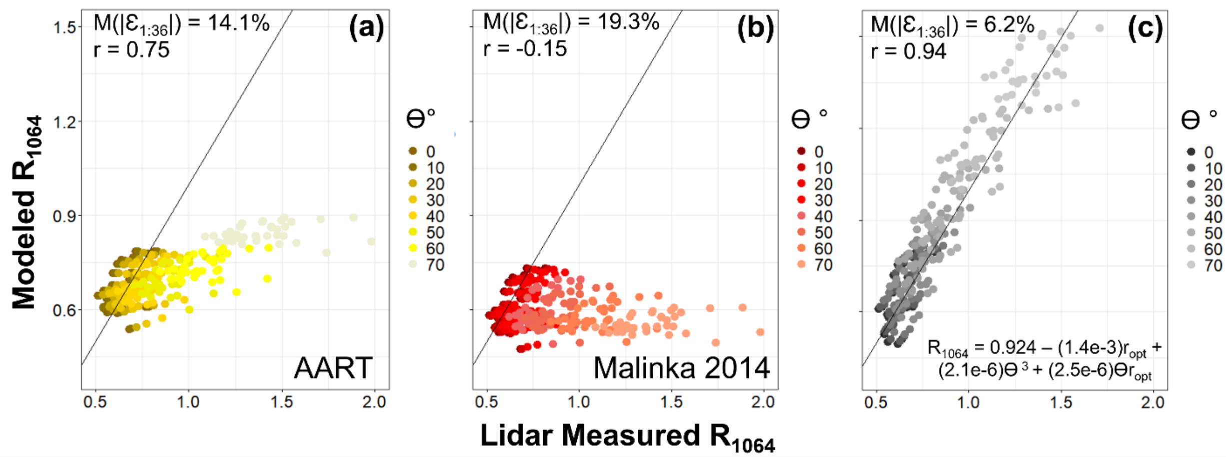

Figure 10: Scatterplots of modeled vs. measured R_{1064} values for AART (gold, a), Malinka 2014 (red, b) and a statistical model (black/grey, c). Lighter shades correspond to increasing values of Θ.

4 Discussion

4.1 Shape optimization

Our results echo the findings presented in the passive detection intercomparison of Dillon et al. (2024b): modeled snow grain shape has a substantial influence on the quality and accuracy of grain size retrievals. The researchers found that numerous combinations of shape parameters in both SNICAR and TARTES could produce adequate results, and we observe very similar patterns in Fig. 4. Therefore, as they advised, performing a shape optimization analysis for a particular instrument, application, etc. seems prudent. In Dillon et al. (2024b), which was the first to examine the optimal combination of shape parameters in SNICAR, it was found that the ideal modeled shape is a somewhat flattened, elongated, asymmetrical spheroid (f_s = 0.95, A = 2.1). Our results here also emphasize elongation via aspect ratio (A = 3.6), but with lower shape factor values (f_s = 0.70), thus closer to the idealized Koch snowflake.

Optimization of parameters in TARTES has received considerable attention, and our lidar results are in good agreement with previous research. Regarding the absorption enhancement parameter, B, Libois et al. (2014) observed values ranging from 1.4 to 1.8, resulting in the TARTES default of 1.6. Dillon et al. (2024b) suggested B = 1.7, which concurred with Robledano et al. (2023), who described the optimal modeled shape of snow as, "a collection of convex particles without symmetry…" Regarding the asymmetry parameter, Robledano et al. (2023) observed g = 0.82 via ray tracing methods, while Libois et al. (2013) suggested a range of 0.83 – 0.87, which includes the TARTES default of 0.86 determined by Meirold-Mautner and Lehning (2004). Thus, our lidar values of B = 1.6 and g = 0.85 are well-supported. The corresponding values from the random mixture method in Malinka (2014) are also quite close (B = 1.7, g = 0.84), consistent with the excellent performance of that model here.

4.2 Lidar retrievals

This study constitutes the first attempt to compare lidar-retrieved r_{opt} to other optically retrieved values or micro-CT measurements. Retrieval of r_{opt} from lidar reflectance is essentially a novel technique, and thus we have little to compare our results to. Yang et al. (2017) performed, to our knowledge, the only other grain size extraction from lidar reflectance in the literature, albeit on the very different scale of spaceborne lidar over the polar regions. Unfortunately, limited reference data (i.e., ground truth) availability restricted the researchers to a qualitative assessment of performance, though our measured R_{1064} values appear comparable. We can, however, compare our results to the abundance of retrievals from passive techniques in the literature, where error values of 10 – 25% (when compared to micro-CT) appear standard.

Of particular interest are the results of Dillon et al. (2024b), who compared various passive retrieval techniques to micro-CT using the same samples presented here. Depending on the radiative transfer model, we observed median absolute error values of 25.8 – 38.6% (Fig. 5 and Fig. 6). This is slightly degraded relative to the optimal results found in Dillon et al. (2024b), which are closer to 15% when using TARTES and SNICAR. Our lidar results are more reasonable after excluding samples of a MF grain habit, for which r_{opt} retrievals were woefully underestimated (e.g., Fig. 6). In fact, after removing MF samples, the accuracy obtained by the Malinka model using lidar R_{1064} was equivalent to the best results (median absolute error of 19.8%) achieved with R_{1064} from a NIR imager in Dillon et al. (2024b). This finding provides considerable optimism that lidar, despite its optical ambiguities, can achieve retrieval accuracy on par with its passive detection counterparts. Another interesting comparison between studies is that of relative model performance. In Dillon et al. (2024b), TARTES and SNICAR were the most accurate, followed by AART, and then the model of Malinka (2014). Here, the Malinka model proved to be the most accurate. It is possible that the chord length distribution approach is particularly well-suited for lidar illumination via a collimated beam, and this perhaps explains the excellent agreement between Malinka (2014) and the ray tracing method of Robledano et al. (2023) as well, although this is purely speculative.

Regarding the substantial uncertainty surrounding samples with a MF primary grain habit, it is unclear why these samples produced such a large magnitude of R_{1064} reflectance. We note that, because these grains were so large (several millimeters in some cases), and because our laboratory range was so small, the lidar spot size incident on the snow samples (15 mm) was comparable to the grain size. It is possible that this resulted in a greater number of specular backscatter contributions, resulting in a gain, but the topic requires further investigation. This result runs contrary to the not-yet published work of Ackroyd (2023), who found suitable agreement between lidar vs. hyperspectrally retrieved r_{opt} values on larger snow grains.

In contrast to lidar results at nadir, retrieval accuracy decreased substantially as we proceeded to increasingly oblique angles of Θ (Fig. 8). Median accuracy for both AART and the Malinka model surpassed 30% by $\Theta = 30°$, and eventually reached 100% by $\Theta = 60°$. This reduction with illumination/viewing incident angle was due to a discrepancy between radiative transfer models and measurements. Conceptually, it is difficult to predict the behavior of R_{1064} as a function of Θ. It is well-understood that albedo increases with illumination zenith angle. However, given that snow is increasingly forward-scattering at larger incidence angles (Dumont et al., 2010), one might expect that a lesser proportion of light would be reflected in the direct backscatter orientation. That said, direct backscatter is the next most probable snow scattering orientation after directly forward scattering. In its totality, it is a non-trivial matter of dueling phenomenon. The AART model predicts an increase in R_{1064} with Θ; the Malinka model predicts either a slight increase or decrease, depending on r_{opt}. In either case, neither model anticipates the steep increase in reflectance that we observe in our measurements, at least not at $\Theta \leq 70°$. Simply put, the models underestimated the rate of increase in backscattered reflectance with Θ relative to our lidar observations (Fig. 9), and thus a parameterization may prove useful when larger incidence angles are encountered (Fig. 10c). However, we caution that this should be investigated and calibrated to a specific instrument and application; the empirical relationship presented here is purely a means of demonstration, but represents an encouraging result for operations in complex topography.

Future work on lidar retrievals should focus on scaling the laboratory findings presented here to field operations. The accuracy of r_{opt} retrievals from lidar reflectance should be evaluated at longer range and/or for larger incident beam diameters. To replicate the radiometric calibration performed here, a Lambertian standard large enough to overfill the scanner's iFOV would be required, ideally oriented perpendicular to the incident beam. The measured return intensity from this standard would then need to be range-corrected to allow for appropriate calibration of points at varying range. Alternatively, if parameters such as outgoing power and beam divergence are well-known, atmospheric correction in conjunction with the lidar equation could be used to calculate bidirectional reflectance factors in lieu of a calibration standard.

5 Conclusions

Our research demonstrates the novel capacity of lidar to accurately map snow optical grain size, while also comparing accuracy across radiative transfer models, in a controlled laboratory environment. Listed as a set of general conclusions, we found that:

v. At nadir, lidar-based r_{opt} retrievals were adequate and comparable to passive detection, but only after excluding samples of a MF grain habit, where underestimation was substantial.

vi. Concerning radiative transfer models, the model of Malinka (2014) slightly outperformed its counterparts (TARTES, SNICAR, and AART).

vii. Shape parameter combinations of $f_s = 0.70/A = 3.6$ and $B = 1.6/g = 0.85$ performed best for SNICAR and TARTES, respectively. However, operation-specific shape optimization would be ideal.

viii. Lidar-based retrievals off-nadir were satisfactory from illumination/viewing zenith angles of 0 - 30°, but fell off considerably at higher values of Θ due to a discrepancy between models and observations. We suggest a statistical parameterization to resolve this dilemma.

As lidar becomes more economical, it presents itself as a capable instrument for measuring NIR reflectance and corresponding retrievals of r_{opt}, with significant advantages over passive detectors. The findings presented here may provide guidance for enhanced r_{opt} (and thus snow albedo) mapping. Extending this work to field operations will have immediate implications for Earth surface energy balance estimates and subsequent impacts on climate, hydrological, and even avalanche forecasting.

6 References

Ackroyd, C. S. (2023). *Airborne Lidar Intensity Correction for Mapping Snow Cover Extent and Grain Size in Mountainous Terrain* (Doctoral dissertation, The University of Utah).

Bair, E. H., Stillinger, T., & Dozier, J. (2020). Snow Property Inversion from Remote Sensing (SPIReS): A generalized multispectral unmixing approach with examples from MODIS and Landsat 8 OLI. *IEEE Transactions on Geoscience and Remote Sensing*, *59*(9), 7270-7284.

Bhandari, A., Hamre, B., Frette, Ø., Zhao, L., Stamnes, J. J., & Kildemo, M. (2011). Bidirectional reflectance distribution function of Spectralon white reflectance standard illuminated by incoherent unpolarized and plane-polarized light. *Applied Optics*, *50*(16), 2431-2442.

Bohn, N., Painter, T. H., Thompson, D. R., Carmon, N., Susiluoto, J., Turmon, M. J., ... & Guanter, L. (2021). Optimal estimation of snow and ice surface parameters from imaging spectroscopy measurements. *Remote Sensing of Environment*, *264*, 112613.

Deems, J. S., Painter, T. H., & Finnegan, D. C. (2013). Lidar measurement of snow depth: a review. *Journal of Glaciology*, *59*(215), 467-479.

Dillon, J., Donahue, C., Schehrer, E., Birkeland, K., & Hammonds, K. (2024). Mapping surface hoar from near-infrared texture in a laboratory. *The Cryosphere*, *18*(5), 2557-2582.

Donahue, C., Skiles, S. M., & Hammonds, K. (2021). In situ effective snow grain size mapping using a compact hyperspectral imager. *Journal of Glaciology*, *67*(261), 49-57.

Donahue, C., Skiles, S. M., & Hammonds, K. (2022). Mapping liquid water content in snow at the millimeter scale: an intercomparison of mixed-phase optical property models using hyperspectral imaging and in situ measurements. *The Cryosphere*, *16*(1), 43-59.

Donahue, C. P., Menounos, B., Viner, N., Skiles, S. M., Beffort, S., Denouden, T., ... & Heathfield, D. (2023). Bridging the gap between airborne and spaceborne imaging spectroscopy for mountain glacier surface property retrievals. *Remote Sensing of Environment*, *299*, 113849.

Dumont, M., Brissaud, O., Picard, G., Schmitt, B., Gallet, J. C., & Arnaud, Y. (2010). High-accuracy measurements of snow Bidirectional Reflectance Distribution Function at visible and NIR wavelengths–comparison with modelling results. *Atmospheric Chemistry and Physics*, *10*(5), 2507-2520.

Fierz, C., Armstrong, R., Durand, Y., Etchevers, P., Greene, E., McClung, D., ... & der Künste ZHdK, Z. (2009). IACS international classification for seasonal snow on the ground. *International Association of Cryospheric Sciences, UNESCO, Paris*.

Flanner, M. G., & Zender, C. S. (2005). Snowpack radiative heating: Influence on Tibetan Plateau climate. *Geophysical research letters*, *32*(6).

Flanner, M. G., Liu, X., Zhou, C., Penner, J. E., & Jiao, C. (2012). Enhanced solar energy absorption by internally-mixed black carbon in snow grains. *Atmospheric Chemistry and Physics*, *12*(10), 4699-4721.

Flanner, M. G., Arnheim, J. B., Cook, J. M., Dang, C., He, C., Huang, X., ... & Zender, C. S. (2021). SNICAR-ADv3: a community tool for modeling spectral snow albedo. *Geoscientific Model Development*, *14*(12), 7673-7704.

Gallet, J. C., Domine, F., Zender, C. S., & Picard, G. (2009). Measurement of the specific surface area of snow using infrared reflectance in an integrating sphere at 1310 and 1550 nm. *The Cryosphere*, *3*(2), 167-182.

Gergely, M., Wolfsperger, F., & Schneebeli, M. (2013). Simulation and validation of the InfraSnow: An instrument to measure snow optically equivalent grain size. *IEEE transactions on geoscience and remote sensing*, *52*(7), 4236-4247.

Grenfell, T. C., & Warren, S. G. (1999). Representation of a nonspherical ice particle by a collection of independent spheres for scattering and absorption of radiation. *Journal of Geophysical Research: Atmospheres*, *104*(D24), 31697-31709.

He, C., Takano, Y., Liou, K. N., Yang, P., Li, Q., & Chen, F. (2017). Impact of snow grain shape and black carbon–snow internal mixing on snow optical properties: Parameterizations for climate models. *Journal of Climate*, *30*(24), 10019-10036.

Kokhanovsky, A. A., & Zege, E. P. (2004). Scattering optics of snow. *Applied optics*, *43*(7), 1589-1602.

Legagneux, L., Cabanes, A., & Dominé, F. (2002). Measurement of the specific surface area of 176 snow samples using methane adsorption at 77 K. *Journal of Geophysical Research: Atmospheres*, *107*(D17), ACH-5.

Li, D., Zhao, M. H., Garra, J., Kolpak, A. M., Rappe, A. M., Bonnell, D. A., & Vohs, J. M. (2008). Direct in situ determination of the polarization dependence of physisorption on ferroelectric surfaces. *nature materials*, *7*(6), 473-477.

Libois, Q., Picard, G., France, J. L., Arnaud, L., Dumont, M., Carmagnola, C. M., & King, M. D. (2013). Influence of grain shape on light penetration in snow. *The Cryosphere*, *7*(6), 1803-1818.

Libois, Q., Picard, G., Dumont, M., Arnaud, L., Sergent, C., Pougatch, E., ... & Vial, D. (2014). Experimental determination of the absorption enhancement parameter of snow. *Journal of Glaciology*, *60*(222), 714-724.

Lorensen, W. E., & Cline, H. E. (1998). Marching cubes: A high resolution 3D surface construction algorithm. In *Seminal graphics: pioneering efforts that shaped the field* (pp. 347-353).

Malinka, A. V. (2014). Light scattering in porous materials: Geometrical optics and stereological approach. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *141*, 14-23.

Malinka, A. (2023). Stereological approach to radiative transfer in porous materials. Application to the optics of snow. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *295*, 108410.

Marks, D., & Dozier, J. (1992). Climate and energy exchange at the snow surface in the alpine region of the Sierra Nevada: 2. Snow cover energy balance. *Water Resources Research*, *28*(11), 3043-3054.

Matzl, M., & Schneebeli, M. (2006). Measuring specific surface area of snow by near-infrared photography. *Journal of Glaciology*, *52*(179), 558-564.

Meirold-Mautner, I., & Lehning, M. (2004). Measurements and model calculations of the solar shortwave fluxes in snow on Summit, Greenland. *Annals of Glaciology*, *38*, 279-284.

Nolin, A. W., & Dozier, J. (2000). A hyperspectral method for remotely sensing the grain size of snow. *Remote sensing of Environment*, *74*(2), 207-216.

Painter, T. H., Barrett, A. P., Landry, C. C., Neff, J. C., Cassidy, M. P., Lawrence, C. R., ... & Farmer, G. L. (2007). Impact of disturbed desert soils on duration of mountain snow cover. *Geophysical Research Letters*, *34*(12).

Painter, T. H., Rittger, K., McKenzie, C., Slaughter, P., Davis, R. E., & Dozier, J. (2009). Retrieval of subpixel snow covered area, grain size, and albedo from MODIS. *Remote Sensing of Environment*, *113*(4), 868-879.

Painter, T. H., Skiles, S. M., Deems, J. S., Bryant, A. C., & Landry, C. C. (2012). Dust radiative forcing in snow of the Upper Colorado River Basin: 1. A 6 year record of energy balance, radiation, and dust concentrations. *Water Resources Research*, *48*(7).

Picard, G., Libois, Q., & Arnaud, L. (2016). Refinement of the ice absorption spectrum in the visible using radiance profile measurements in Antarctic snow. *The Cryosphere*, *10*(6), 2655-2672.

Robledano, A., Picard, G., Dumont, M., Flin, F., Arnaud, L., & Libois, Q. (2023). Unraveling the optical shape of snow. *Nature Communications*, *14*(1), 3955.

Sassen, K. (2005). Polarization in lidar. In *LIDAR: Range-resolved optical remote sensing of the atmosphere* (pp. 19-42). New York, NY: Springer New York.

Seidel, F. C., Rittger, K., Skiles, S. M., Molotch, N. P., & Painter, T. H. (2016). Case study of spatial and temporal variability of snow cover, grain size, albedo and radiative forcing in the Sierra Nevada and Rocky Mountain snowpack derived from imaging spectroscopy. *The Cryosphere*, *10*(3), 1229-1244.

Skiles, S. M., Painter, T. H., Deems, J. S., Bryant, A. C., & Landry, C. C. (2012). Dust radiative forcing in snow of the Upper Colorado River Basin: 2. Interannual variability in radiative forcing and snowmelt rates. *Water Resources Research*, *48*(7).

Skiles, S. M., Flanner, M., Cook, J. M., Dumont, M., & Painter, T. H. (2018). Radiative forcing by light-absorbing particles in snow. *Nature Climate Change*, *8*(11), 964-971.

Skiles, S. M., Donahue, C. P., Hunsaker, A. G., & Jacobs, J. M. (2023). UAV hyperspectral imaging for multiscale assessment of Landsat 9 snow grain size and albedo. *Frontiers in Remote Sensing*, *3*, 1038287.

Stamnes, K., Tsay, S. C., Wiscombe, W., & Jayaweera, K. (1988). Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media. *Applied optics*, *27*(12), 2502-2509.

Stanton, B., Miller, D., Adams, E., & Shaw, J. A. (2016). Bidirectional-reflectance measurements for various snow crystal morphologies. *Cold Regions Science and Technology*, *124*, 110-117.

Warren, S. G. (1982). Optical properties of snow. *Reviews of Geophysics*, *20*(1), 67-89.

Warren, S. G. (1984). Optical constants of ice from the ultraviolet to the microwave. *Applied optics*, *23*(8), 1206-1225.

Warren, S. G., & Brandt, R. E. (2008). Optical constants of ice from the ultraviolet to the microwave: A revised compilation. *Journal of Geophysical Research: Atmospheres*, *113*(D14).

Yang, Y., Marshak, A., Han, M., Palm, S. P., & Harding, D. J. (2017). Snow grain size retrieval over the polar ice sheets with the Ice, Cloud, and land Elevation Satellite (ICESat) observations. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *188*, 159-164.

CHAPTER FOUR

MAPPING SURFACE HOAR FROM NEAR-INFRARED TEXTURE IN A LABORATORY

Contribution of Authors and Co-Authors

Manuscript in Chapter 4

Author: James Dillon

Contributions: Conceptualized the study, collected laboratory data, analyzed results, and drafted the manuscript.

Co-Author: Christopher Donahue

Contributions: Provided guidance and advised during conceptualization and analysis.

Co-Author: Evan Schehrer

Contributions: Assisted with laboratory sample preparation and data collection.

Co-Author: Karl Birkeland

Contributions: Provided guidance and advised during conceptualization and analysis.

Co-Author: Kevin Hammonds

Contributions: Acquired funding, was responsible for project administration, and supervised throughout.

Manuscript Information

James Dillon, Christopher Donahue, Evan Schehrer, Karl Birkeland, Kevin Hammonds

The Cryosphere

Status of Manuscript:
☐ Prepared for submission to a peer-reviewed journal
☐ Officially submitted to a peer-reviewed journal
☐ Accepted by a peer-reviewed journal
X Published in a peer-reviewed journal

Publisher: Copernicus
Submitted: 24 December 2023
Published: 24 May 2024
Volume: 18
DOI: https://doi.org/10.5194/tc-18-2557-2024

Mapping surface hoar from near-infrared texture in a laboratory

James Dillon[1], Christopher Donahue[2], Evan Schehrer[1], Karl Birkeland[3], Kevin Hammonds[1]
[1]Department of Civil Engineering, Montana State University, Bozeman, MT, USA
[2]Department of Geography, Earth, and Environmental Sciences, University of Northern British Columbia, Prince George, British Columbia, Canada
[3]USDA Forest Service National Avalanche Center, Bozeman, MT, USA

Correspondence to: James Dillon (james.dillon013@gmail.com)

Abstract. Surface hoar crystals are snow grains that form when water vapor deposits on the snow surface. Once buried, surface hoar creates a weak layer in the snowpack that can later cause large avalanches to occur. The formation and persistence of surface hoar are highly spatiotemporally variable making its detection difficult. Remote sensing technology capable of detecting the presence and spatial distribution of surface hoar would be beneficial for avalanche forecasting, however this capability has yet to be developed. Here, we hypothesize that near-infrared (NIR) texture, defined as the spatial variability of reflectance magnitude, may produce an optical signature unique to surface hoar due to the grains distinct shape and orientation. We tested this hypothesis by performing reflectance experiments in a controlled cold laboratory environment to evaluate the potential and accuracy of surface hoar mapping from NIR texture using a near-infrared hyperspectral imager (NIR-HSI) and a lidar operating at 1064 nm. We analyzed forty-one snow samples; three of which were surface hoar and 38 that consisted of other grain morphologies. When using NIR-HSI under direct and diffuse illumination, we found that surface hoar displayed higher NIR texture relative to all other grain shapes across numerous spectral bands and a wide range of spatial resolutions (0.5 - 50 mm). Due to the large number of spectral and spatial resolution combinations, we conducted a detailed samplewise case study at 1324 nm spectral and 10 mm spatial resolutions. The case study resulted in the median texture of surface hoar being 1.3 to 8.6 times greater than the 38 other samples under direct and diffuse illumination ($p < 0.05$ in all cases). Using lidar, surface hoar also exhibited significantly increased NIR texture in 30 out of 38 samples, but only at select (5 – 25 mm) spatial resolutions. Leveraging these results, we propose a simple binary classification algorithm to map the extent of surface hoar on a pixelwise basis using both the NIR-HSI and lidar instruments. The NIR-HSI under direct and diffuse illumination performed best, with a median accuracy of 96.91% and 97.37%, respectively. Conversely, median classification accuracy with lidar was only 66.99%. Further, to assess the repeatability of our method and demonstrate mapping capacity, we ran the algorithm on a new sample with mixed microstructures, with accuracy of 99.61% and 96.15% for direct and diffuse, respectively. As NIR-HSI detectors become increasingly available, our findings demonstrate the potential of a new tool for avalanche forecasters to remotely assess the spatiotemporal variability of surface hoar, which would improve avalanche forecasts and potentially save lives.

1 Introduction and background

Mountainous and polar snowpacks are commonly blanketed by surface hoar, unique ice crystals that grow when water vapor deposits on the snow surface (Horton and Jamieson, 2017). Hoar crystals can grow to several centimeters in length and stand in a predominately vertical (but often quite variable) orientation atop the snow surface. Surface hoar is a prominent concern for avalanche forecasters due to its propensity for creating weak layers in the snowpack that can cause large avalanches. Generally, a weak layer is formed due to a snow layer being weakly bonded to the slab above, as in the case of an ice lens, or being of low shear strength, such as with surface hoar. Once buried, surface hoar layers are prone to fracture propagation and avalanche release (Horton and Jamieson, 2017; Jamieson and Schweizer, 2000). For instance, Birkeland (1998) found that nearly one-third of large natural avalanches in southwestern Montana failed on buried layers of surface hoar, while Jamieson and Schweizer (2000) similarly observed in a Swiss dataset that 40% of avalanches released on a surface hoar weak layer.

Surface hoar formation and persistence is highly spatiotemporally variable. Its presence is difficult to predict because it depends on the complex spatial distributions of atmospheric water vapor, precipitation, wind, radiation, and vegetation that are present in polar and mountainous environments. Temporally, surface hoar can be promptly destroyed by environmental influences, such as wind, or it can persist for weeks (Champollion et al., 2013; Lutz and Birkeland, 2011). The formation of surface hoar can occur across the landscape or in isolated sub-slope regions. Therefore, before becoming a buried weak layer in the snowpack, surface hoar detection is an ideal application for remote sensing.

Optical remote sensing retrievals of snow and ice commonly use near-infrared (NIR) wavelengths where ice is absorptive, and reflectance is sensitive to microstructure. In the NIR, snow grain size is the primary driver of snow reflectance and albedo; the increased path length of light through ice yields a reduction in reflectance. Despite the complex shapes of snow grains, representing snow as a collection of ice spheres with radius (r_e), known as the optical or effective snow grain size, is commonly used for simulating snow reflectance and albedo (Grenfell and Warren, 1999). The effective grain size is related to the physical snow microstructure through the ice surface area-to-volume ratio (or specific surface area; SSA) by the following relationship:

$$r_e = \frac{3}{SSA} \quad (1)$$

In practice, the non-linear inverse relationship between NIR reflectance and r_e is leveraged to map r_e from reflectance measurements (e.g., Nolin and Dozier, 2000). Because grain size controls broadband NIR albedo, estimates of r_e have historically dominated physical snow surface characterization, constituting a major goal of snow optics.

In recent decades, the use of NIR hyperspectral imaging (NIR-HSI), and (to a lesser extent) light detection and ranging (lidar), have enhanced r_e mapping efforts. Hyperspectral instruments have superior spectral resolution relative to broadband or multispectral reflectance measurements, producing a nearly continuous measured spectra for each pixel in an image. Lidar, on the other hand, holds key advantages as an active remote sensor. Lidar units scan a surface by emitting rapid pulses of light (most commonly at a NIR wavelength) and record both the relative strength of backscattered light after reflecting off a target, and the two-way travel time. These two measurement types can be used independently; the travel time is used to measure surface topography, whereas the backscattered magnitude has been demonstrated as a useful measure of optical properties. Although far less validated than traditional passive reflectance measurements, lidar backscatter may provide adequate r_e estimates (Yang et al., 2017).

Despite advances in snow optics, Horton and Jamieson (2017) note the fundamental disconnect between snow surface characterizations conducted by the remote sensing community (i.e., r_e mapping), and the physical properties relevant to avalanche release. For avalanche forecasters, characterizing snow surface microstructure with the morphological grain shapes defined in the International Classification for Seasonal Snow on the Ground (ICSSG, Fierz et al., 2009) and their associated mechanical properties is critically important because these mechanical properties help determine how well new snow will bond to the old snow surface. Hence, a forecaster would much prefer a map identifying a potential future weak layer, like surface hoar, than a map of r_e. Relating morphological grain shape to r_e is difficult because effective grain size only considers the path length of ice which is a complex function of grain shape, traditional grain size, bulk density, and other physical characteristics. Therefore, r_e is not particularly useful for avalanche forecasting operations, although room for this adaptation does exist. While certain magnitudes of r_e are generally related to grain shape (Domine et al., 2007; Matzl and Schneebeli, 2010), few studies have formally attempted to use NIR reflectance for mapping morphological grain shape instead of r_e. Further, the studies that have attempted this (e.g., Bühler et al., 2014; Horton and Jamieson, 2017) have found that surface hoar crystals produce moderate reflectance signatures relative to other grain shapes, making them difficult to delineate from less-concerning snow microstructures based on NIR reflectance magnitude.

To the best of our knowledge, the only study to date that successfully identified surface hoar formation from NIR remote sensing was conducted by Champollion et al. (2013) in Antarctica. As opposed to evaluating magnitudes of NIR reflectance, the researchers leveraged a NIR texture signature, defined as the localized spatial variability in reflectance, to classify the presence of surface hoar using an infrared camera and an 850 nm artificial light source. Simply put, the researchers found that a large, localized variance in NIR reflectance, as quantified by a contrast index, was strongly correlated with surface hoar crystals. Similarly, in the preliminary findings presented by Walter et al. (2023), the researchers quantified a 600% increase in NIR reflectance spatial variability during surface hoar formation, measured with a 905 nm lidar unit.

Here, we postulate that the physical phenomena for these findings include increasing specular contributions with surface hoar growth (Walter et al., 2023), as well as variable ice absorption and path length (Fig. 1b). Depending on how an incoming photon interacts with the relatively large, often vertically oriented surface hoar crystals, the photon could experience a wide range of ice path lengths and thus absorption. Further, it is thought that surface hoar plates promote specular reflectance, which can cause large portions of radiation to either return to the sensor or forward scatter toward the snowpack and into the "radiation trap" of hoar crystals. As a result, *we hypothesize that the presence of surface hoar will coincide with a quantifiable increase in localized reflectance variance*, a texture signature, which could be used to map its distribution. If this is the case, it would enable fine spatial and temporal resolution mapping of surface hoar extent (prior to burial) in challenging to observe environments, particularly as NIR remote sensors and uncrewed aerial vehicles (UAVs) become more cost-effective as forecasting tools. However, such a texture analysis has never been fully evaluated, rigorously quantified for accuracy, or compared to a wide variety of well-defined microstructures to ensure that the NIR texture is indeed a unique defining feature of surface hoar.

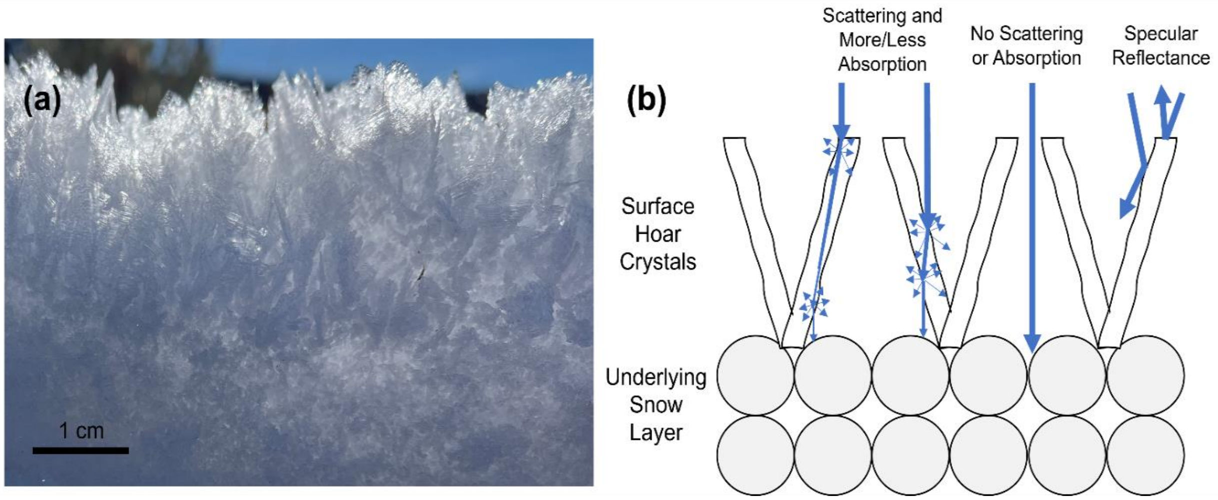

Figure 1: Profile view of surface hoar atop an underlying snow layer (a) juxtaposed with an idealized schematic of light scattering when encountering a surface hoar layer (b). Because surface hoar crystals are typically large and vertically oriented, incoming photons may experience a very long path length and thus substantial absorption before reaching the underlying snow layer (leftmost case). Depending on the angle of interaction, that path may be considerably shortened (middle left case), or, as these crystals tend to be modestly spaced, photons may evade the surface hoar crystals entirely and pass straight into the underlying snow layer (middle case). Further, specular contributions are thought to increase in surface hoar layers (rightmost cases), which can produce particularly high or low reflectance depending on the angle of interaction. Last, because the surface hoar crystals will almost certainly have a different SSA than the underlying layer, light scattering back to the sensor directly from the surface hoar layer should further enhance the spatial variability of reflectance.

To address this knowledge gap, we performed controlled cold laboratory experiments to determine whether NIR texture can delineate the extent of surface hoar. We first created snow samples with varying grain shapes and physical properties and quantified their microstructures using X-ray computed microtomography (micro-CT). Using both a compact NIR-HSI and a terrestrial lidar unit, we scanned each sample under a variety of illumination conditions. We subsequently analyzed the resulting maps of reflectance to produce measurements of NIR texture. Finally, we assessed the statistical significance of increased texture in surface hoar samples. We used our results to inform optimal thresholds for classifying surface hoar on a per-pixel basis, before analyzing the accuracy of our resulting classified data products.

2 Methodology

We aimed to prepare laboratory snow samples with a wide variety of well-defined grain shapes and microstructures, acquire optical measurements, and perform a texture analysis towards delineation of surface hoar from other snow surface grain shapes. Section 2.1 describes snow sample preparation and physical characterization, Sect. 2.2 describes the acquisition of NIR-HSI and lidar data, Sect. 2.3 covers image texture analysis, Sect. 2.4 describes classification, and Sect. 2.5 evaluates the repeatability of our work.

2.1 Sample preparation and physical characterization

2.1.1 Sample preparation

We utilized Montana State University's Subzero Research Laboratory (SRL), a controlled cold laboratory environment, for sample preparation, testing and assessment. The snow used in these experiments was a combination of laboratory-grown crystals produced in the SRL's snowmaking apparatus, which is similar to the systems presented in Schleef et al. (2014) and Abe and Kosugi (2019), and natural undisturbed snow that we collected from the surrounding area. To ensure the snow was completely dry, we kept all samples in a cold room at -30° C and allowed them to equilibrate for at least 24 hours prior to evaluation. We prepared forty-one snow samples from twelve batches of differing snow grains (Fig. 2). From the bulk batches, we sieved snow grains through differing mesh sizes to further promote disparate microstructures (Table 1). The exception to this was surface hoar, which was grown in the laboratory atop rounded grains following the methods used by Stanton et al. (2016). Samples 25 and 26 consisted of in situ surface hoar growth at differing stages. Meanwhile, Sample 24 featured these surface hoar grains redistributed through a large (6.30 mm) sieve, in an effort to further examine the optical behavior of these grains after disrupting their typical vertical structure. Sample grain shapes included precipitation particles (PP), decomposing and fragmented precipitation particles (DF), rounded grains (RG), melt forms (MF), faceted crystals (FC), depth hoar (DH), and surface hoar (SH) (Fierz et al., 2009). We prepared snow samples to be microstructurally homogeneous, both laterally across the sample and vertically over sample depth, in a rectangular 38 cm x 23 cm sample holder. Sample thickness was 3.8 cm, over twice the largest optically active depth estimates for our sample microstructures in the NIR spectrum (Nolin and Dozier, 2000).

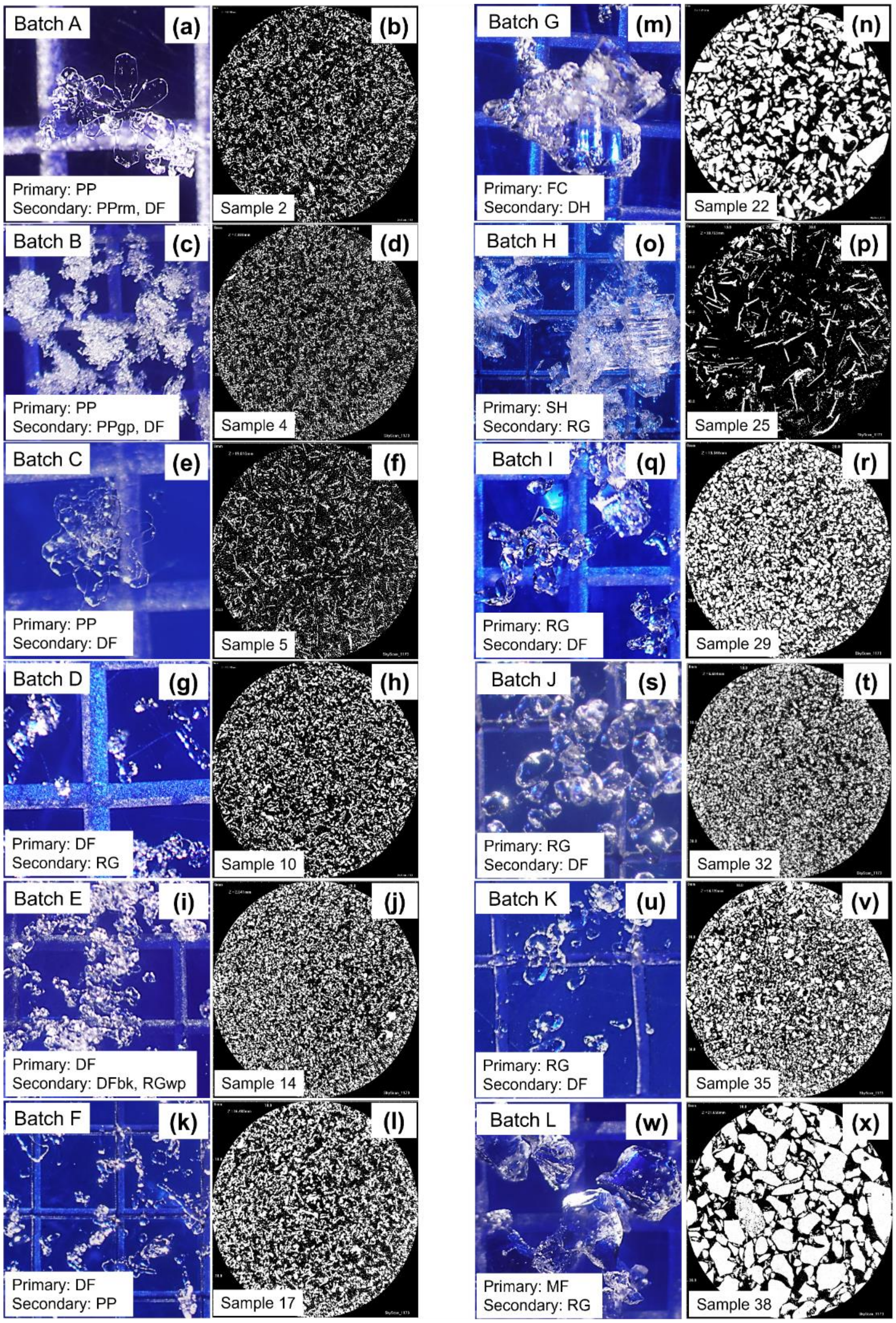

Figure 2: Microscopy images of grains from each initial batch (left columns) and binary micro-CT cross-sections from representative samples (right columns). In the microscopy images, the grid size on the underlying blue grain card is 2 mm.

Table 1: Physical snow sample characteristics organized by primary grain shape and listed in order of decreasing surface area-to-volume ratio therein.

Sample #	Batch ID	Primary Grain Shape	Secondary Grain Shape(s)	Micro-CT SSA (mm^{-1})	Micro-CT ρ ($kg\ m^{-3}$)	Sieve Size (mm)		Notes
						Passed	Caught	
1	A	PP	PPrm, DF	32.87	176	2.38	1.18	
2	A	PP	PPrm, DF	28.98	217	2.38	-	
3	A	PP	PPrm, DF	26.31	211	1.18	0.42	
4	B	PP	PPgp, DF	31.79	160	2.38	1.18	
5	C	PP	DF	33.10	94	-	-	In situ fresh PP
6	C	PP	DF	20.54	286	2.38	1.18	
7	C	PP	DF	20.45	280	0.85	0.42	
8	C	PP	DF	20.12	275	2.38	-	
9	C	PP	DF	18.39	303	1.18	0.85	
10	D	DF	RG	27.44	293	2.38	1.18	
11	D	DF	RG	25.85	323	0.85	0.42	
12	D	DF	RG	25.11	351	1.18	0.85	
13	D	DF	RG	20.77	365	2.38	-	
14	E	DF	DFbk, RGwp	16.27	374	0.85	-	
15	F	DF	PP	14.39	322	2.38	-	
16	F	DF	PP	13.79	312	2.38	1.18	
17	F	DF	PP	13.65	309	1.18	0.85	
18	F	DF	PP	12.99	382	0.85	-	
19	G	FC	DH	14.67	407	1.18	0.42	
20	G	FC	DH	11.32	448	2.38	1.18	
21	G	FC	DH	10.26	417	6.30	3.35	
22	G	FC	DH	10.05	472	6.30	-	
23	G	FC	DH	9.87	404	3.35	2.38	
24	H	SH	RG	14.52	213	6.30	-	Re-sieved SH grains
25	H	SH	RG	10.82	65	-	-	In situ SH atop RGs
26	H	SH	RG	7.50	94	-	-	Smaller than S25
27	I	RG	DF	13.53	381	2.38	1.18	
28	I	RG	DF	13.08	419	1.18	0.85	
29	I	RG	DF	12.77	431	2.38	-	
30	I	RG	DF	12.43	489	0.85	0.42	
31	I	RG	DF	12.40	452	-	-	S29 melt/refreeze
32	J	RG	DF	13.76	394	0.85	-	
33	J	RG	DF	13.45	355	0.42	-	
34	J	RG	DF	10.66	460	1.18	-	
35	K	RG	DF	11.13	404	1.18	0.85	
36	K	RG	DF	10.84	428	0.85	0.42	
37	L	MF	RG	4.96	582	2.38	0.42	
38	L	MF	RG	3.69	545	6.30	-	
39	L	MF	RG	3.13	512	3.15	2.38	
40	L	MF	RG	2.88	467	-	-	Refrozen in-situ
41	L	MF	RG	2.37	433	6.30	3.15	

2.1.2 Physical characterization

We thoroughly characterized the physical properties of each sample, as summarized in Table 1. First, we performed microscopy on representative grains from each batch prior to sieving (Fig. 2), and classified grain shapes using a crystal card and lens following Fierz et al. (2009). After sieving and sample preparation, we collected micro-CT data from each sample using a Bruker SkyScan 1173 housed in a -10° C chamber within the SRL, generally following the protocol outlined by Donahue et al. (2021). To prepare samples, we used a cylindrical holder with 3 cm diameter x 4 cm length, which allowed for a voxel size of 14.5 μm. We obtained measurements using a 42 kV, 190 mA X-ray beam, 100 ms exposure time, with each sample rotated 180° at 0.7° increments. After scanning, we performed thresholding of grey-scale images into ice and air phases by visual inspection (e.g., Fig. 2), and used a despeckling filter to remove white and black speckles. Reconstructions via the marching cubes method (Lorensen and Cline, 1987) allowed us to determine the volume and surface area in 3D, which we used to calculate the SSA and density of each sample.

2.2 Optical measurements

We examined NIR texture under a variety of illumination and viewing conditions, using both lidar and NIR-HSI independently. For NIR-HSI, we constructed laboratory setups for both direct (Sect. 2.2.1.1) and diffuse (Sect. 2.2.1.2) illumination conditions. As lidar (Sect. 2.2.2) produces its own direct illumination via laser irradiance, this was the only possible illumination condition for lidar analysis. Furthermore, it is well-understood that incidence angle impacts snow reflectance in the NIR spectral region; snow is nearly Lambertian under nadir incidence and predominantly forward scattering at off-nadir incidence. To consider this, for each of the aforementioned illumination configurations/instruments, we collected data with: 1) the sample container perpendicular to the detector (nadir), and 2) tilted 10° away from the detector. Hereafter, these incidence angles of 0° and 10° are termed Θ. The result was six datasets representing each illumination/instrument/incidence combination for every sample. We acquired all optical data immediately prior to micro-CT analysis at a constant temperature of -10° C.

2.2.1 Hyperspectral imaging

We used a Resonon Inc. Pika NIR-640 near-infrared hyperspectral imager to map snow reflectance in the NIR (www.resonon.com). Donahue et al. (2021) provide a detailed description of the instrument. Briefly, the imager's spectral resolution ranges from 2.39 to 2.50 nm, and measures 336 bands across the NIR region from 891–1711 nm. It constructs a 2D image containing the full spectrum in each pixel by collecting the image line by line, known commonly as a "push broom" or "line" scanner. Thus, to collect an image, the camera must move (translating or rotating) relative to the scene, or the scene must move relative to the imager. We used a Resonon benchtop linear scanning stage to move the sample beneath the sensor.

2.2.1.1 Direct illumination data collection

We positioned the hyperspectral imager above the linear translating stage that held the samples. For more details on the benchtop apparatus, see Donahue et al. (2022). The lens of the imager is surrounded by a set of four halogen lamps that produce direct illumination (Fig. 3a). The halogen lamps and lens of the imager are at a height of 38 and 47 cm above the snow surface, respectively. We used a large Spectralon white diffuse reflectance panel to perform a pixel-by-pixel calibration, resulting in a reflectance factor (R) measured for each band in every individual pixel of the image.

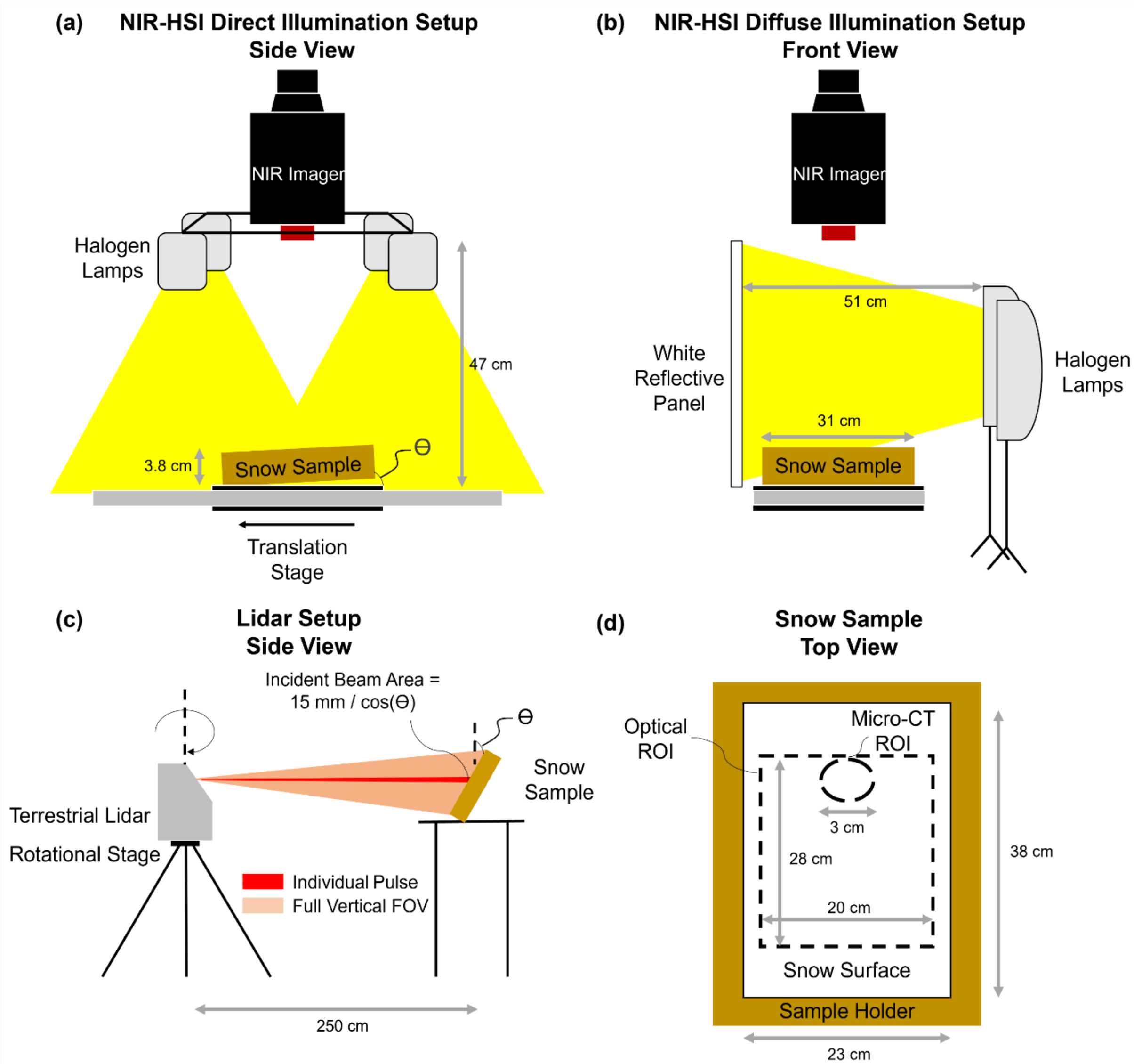

Figure 3: Laboratory data collection schematic for hyperspectral imaging under direct (a) and diffuse (b) illumination conditions, as well as for lidar (c). Data regions-of-interest (ROIs) within the snow sample are illustrated in (d). Angle Θ refers to the viewing incidence angle (as well as the illumination incidence angle, in the case of both lidar and NIR-HSI under direct illumination).

2.2.1.2 Diffuse illumination data collection

For the diffuse illumination setup, we kept the imager mounted in the same orientation, but removed the set of four halogen lights used for direct illumination. Instead, we positioned two larger softbox diffuse halogen lamps (Westcott uLite) perpendicular to the snow sample surface (Fig. 3b). Aiming the lights at a large white panel on the opposite side of the snow sample, 51 cm away mimicked diffuse hemispheric illumination conditions with no direct component. Once again, we used a Spectralon panel to convert raw data to reflectance images.

2.2.1.3 Hyperspectral data processing

Initial processing took place in Resonon's proprietary Spectronon software, and analyses thereafter performed in Rstudio. We began by truncating the reflectance data from each dataset to a central region-of-interest (ROI) encapsulating the micro-CT ROI (Fig. 3d). Our goal was to exclude edge effects from mixed pixels/points, particularly considering lidar beam dilation when samples were tilted off-nadir. Resulting NIR-HSI ROIs contained 224,000 pixels with a spatial resolution of 0.5 mm. Reflectance images were produced from 188 of the 336 available bands, ranging from 951 – 1403 nm, trimmed to reduce noise at the lower end of the imager spectral range and to exclude regions where snow and ice are scarcely reflective. Example snow sample reflectance maps from a single band centered at 1030 nm, the location of a prominent ice absorption feature, (hereafter R_{1030}) are illustrated in Fig. 4c and 4d. The last factor we sought to examine was spatial resolution, considering that if a NIR texture signature specific to surface hoar does exist, then it is likely resolution-dependent. Thus, we coarsened all datasets to spatial resolutions of 1, 2.5, 5, 10, 25, and 50 mm (hence two orders of magnitude), as an attempt to mimic the finer spatial resolutions achievable by UAV-mounted systems.

2.2.2 Lidar

2.2.2.1 Data collection

We used a Riegl VZ-6000 terrestrial laser scanner mounted to a tripod. The scanner achieves vertical (line) scanning via an oscillating mirror while moving horizontally on a rotating head (Fig. 3c). The maximum vertical scan field-of-view (sFOV) is 60° - 120° from zenith and selectable therein, while the horizontal sFOV can range from 0° to a full 360° panorama. The laser operates in the NIR range narrowly around a central wavelength of 1064 nm. We set vertical and horizontal angular increments to 0.01° to maximize resolution (point density) and selected a pulse repetition frequency of 300 kHz. The initial laser beam diameter upon exiting the scanner is 15 mm with a divergence of 0.12 mrad. We positioned the scanner about 2.5 m from the snow samples. The duration of each scan was approximately 1 minute. The resulting data product for lidar is a "cloud" of discrete vector data points, called returns. Similar to the hyperspectral imaging setup, we used a Lambertian reflectance standard to convert return power to bidirectional reflectance factor for each individual point in the clouds.

2.2.2.2 Lidar data processing

Initial processing of point clouds took place in Riegl's RiScan application, and analyses thereafter in Rstudio. As with NIR-HSI data, we trimmed point clouds to a central ROI (Fig. 3d). After truncation, the average point cloud contained 80,000 returns, with spacing of ~ 1.4 pts/mm^2. In order to perform pixelwise operations and to better compare with hyperspectral imagery, we converted the point cloud into an image of continuous pixels. Using a Delaunay triangulation, we interpolated point clouds into images with 1 mm spatial resolution, resulting in 56,000 pixels per dataset. The Riegl VZ-6000 lidar unit operates narrowly around a central wavelength of 1064 nm. Therefore, unlike NIR-HSI, the lidar is only capable of measuring bidirectional reflectance specifically at 1064 nm (hereafter R_{1064}; e.g., Fig. 4e and 4f). This wavelength occurs on the shoulder of the ice absorption feature centered at 1030 nm (Inset of Fig. 4e). Last, as with NIR-HSI, we coarsened all lidar datasets to the resolutions listed in Sect. 2.2.1.3 to examine the influence of spatial resolution on NIR texture.

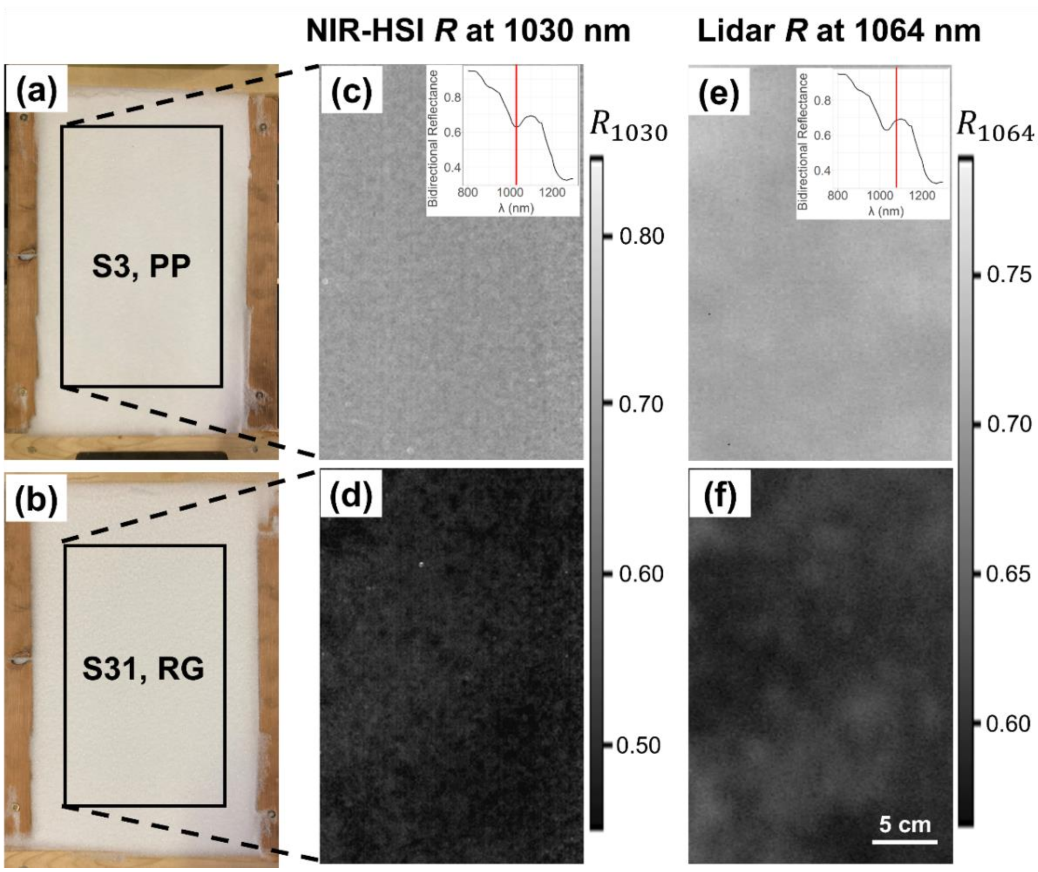

Figure 4: Visible photographs of two different snow samples (a and b) are juxtaposed with greyscale images of hyperspectral reflectance at 1030 nm (c and d), as well as lidar-derived reflectance at 1064 nm (e and f). Samples are shown under direct illumination at ϴ = 0° and at their original resolutions, prior to any coarsening. The data in the top row depicts Sample 3, a snow microstructure with relatively high specific surface area (26.31 mm^{-1}) compared to Sample 31 in the bottom row, which had a specific surface area of 12.40 mm^{-1}. The insensitivity of reflectance to snow microstructure in the visible range, as opposed to the dramatically different reflected magnitudes in the NIR, is apparent. The inset figures illustrate the location of the bidirectional reflectance measurements relative to the ice absorption spectra. These example spectra were produced using the Asymptotic Radiative Transfer model (Kokhanovsky and Zege, 2004).

2.3 Texture analysis

2.3.1 Moving window focal analysis

To restate our hypothesis, we anticipated that surface hoar would display heightened NIR variability, or texture, relative to other snow surface grain shapes due to the physical phenomenon illustrated in Fig. 1. Therefore, we sought to evaluate localized variability of reflectance. To achieve this, we performed a moving window focal analysis to create maps of local standard deviation. Using R_{1030} from NIR-HSI, a demonstration of both coarsening and subsequent moving window analysis is presented in Fig. 5 for Sample 20. Beginning with a map of R at either the native 0.5 mm resolution (Fig. 5a), or coarsened resolutions of 5 mm and 50 mm (Fig. 5b – 5c, respectively), a 9-pixel neighborhood is placed around a central pixel. The standard deviation of R_{1030} (hereafter σ_{1030}) is calculated within the window via Equation 2:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{3}\sum_{j=1}^{3}(R_{ij} - \bar{R})^2}{N}} \tag{2}$$

where R_{ij} is the reflectance value at row i and column j in the 3x3 grid, $\bar{R}$, is the mean of the reflectance values in the 3x3 grid, and N is the total number of pixel values (nine in this case). The resulting value of σ is assigned to the central pixel.

We handled edges by truncating the window size when necessary, such that corner pixels only considered three neighboring cells, for example. Moving the window across each R_{1030} map and evaluating every pixel independently yields a map of σ_{1030} (Fig. 5d – 5f). Hence, these maps depict NIR reflectance texture, rather than magnitude. We underwent this process for all datasets, thus each of the 188 NIR-HSI bands and lidar-derived R_{1064}, for each illumination condition and incidence angle.

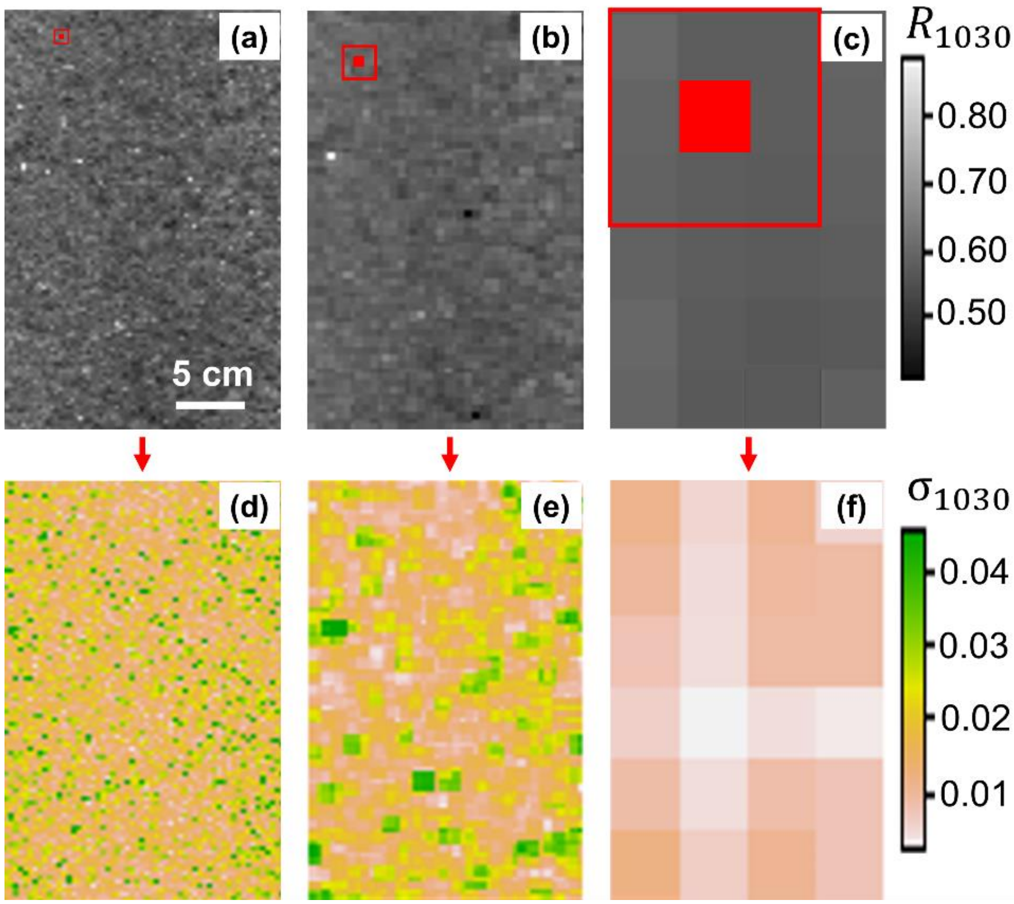

Figure 5: A map of R_{1030} from NIR-HSI at its raw resolution of 0.5 mm (a) is coarsened by two orders of magnitude to 5 mm (b) and 50 mm (c). For each resolution, the localized standard deviation, σ_{1030}, is calculated using a nine-pixel moving window analysis (d – f). The data shown is from Sample 26 under direct illumination at an incidence angle of Θ = 0°.

2.3.2 Spatial and spectral texture analysis

Once texture maps were derived for each dataset, we determined which spatial resolutions were best for surface hoar delineation, assuming some degree of resolution-dependence. Furthermore, leveraging the spectral data provided by NIR-HSI, we examined the relationship between NIR texture and wavelength. We began by grouping the pixelwise σ distributions from surface hoar samples 24 – 26 for each dataset. Similarly, we grouped σ distributions of all other samples (hereafter termed "Other", meaning a microstructure other than SH). Next, we calculated median values of both grouped distributions for each instrument, band, and spatial resolution, and determined the percent difference in medians for each case. Using heat maps and line graphs, we evaluated σ distributions and resulting differences in medians (i.e., Δ M(σ)). This allowed us to determine the spatial resolutions and, in the case of NIR-HSI, the wavelengths, that maximized the difference between SH and Other microstructure medians. We note that a larger difference in texture medians between SH and Other samples should allow for greater ease of SH classification.

2.3.3 Samplewise analysis and significance testing

We next examined how texture varied between individual samples, rather than considering SH samples against an aggregation of other microstructures (Sect. 2.3.2). As a case study, we selected a single band and spatial resolution for each instrument. For NIR-HSI, we chose the band centered at 1324 nm. The results of our spatial and spectral analysis (described in Sect. 2.3.2; results to be discussed in Sect. 3.1) indicated that this was an optimal wavelength under both direct and diffuse illumination. The lidar unit has only one band, centered at 1064 nm. We elected to optimize spatial resolutions based on the results in Sect. 3.1 as well, and thus we selected 10 mm for NIR-HSI and 5 mm for lidar. To determine if surface hoar textures are significantly larger than those of differing microstructures, we compared distributions and median values of σ across samples. As in Sect. 2.3.2, we grouped the pixelwise σ distributions of surface hoar (Samples 24 – 26) for each dataset. We then compared the median value of this grouped SH distribution against samplewise medians. Specifically, we performed one-sided, one-sample t-tests to assess the grouped median against the median σ of each sample. The flow chart in Fig. 6 describes the processing workflow from raw reflectance images through statistical analyses for the case study. Our hypotheses can be summarized as follows, where X describes an individual sample number:

$$H_0:\ M(\sigma^{S24:S26}) = M\left(\sigma^{S(X)}\right)$$

$$H_A:\ M(\sigma^{S24:S26}) > M\left(\sigma^{S(X)}\right);\quad \alpha = 0.05$$

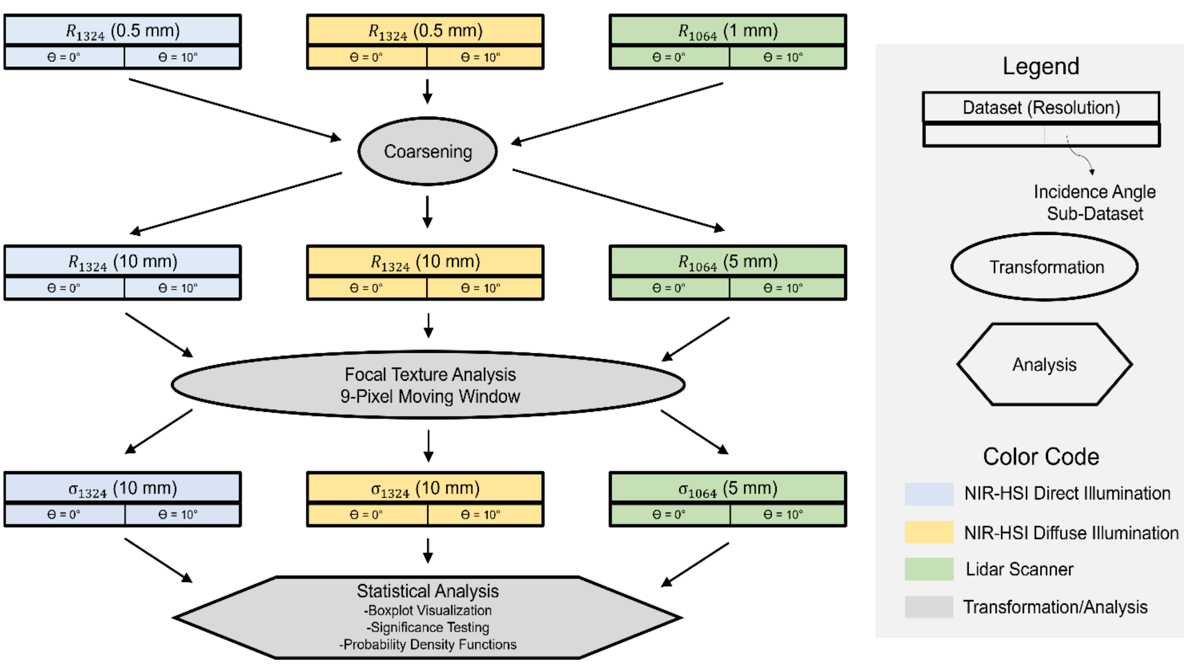

Figure 6: Flowchart of samplewise data processing and statistical analysis workflow.

2.4 Classification algorithm

Continuing our samplewise case study (Sect. 2.3.3), we next produced classified maps of surface hoar. We determined optimal threshold values of σ to delineate surface hoar from other snow surface shapes on a per-pixel basis. In addition to visualizing distributions (across all pixels) of σ with boxplots for each sample, we also constructed probability density functions (PDFs). As in Sect. 2.3.2, we grouped the distributions of surface hoar samples (Samples 24 – 26) and compared them to the aggregated distributions of all other samples. We selected values of σ corresponding to the intersection of the grouped PDFs as the optimal thresholds of delineation for each dataset (i.e., each combination of instrument, illumination condition, and incidence angle), termed σ_{crit}. Using these threshold values, we performed a binary pixelwise classification; pixels with σ values above σ_{crit} were classified as surface hoar, while values beneath were designated as “Other” microstructures. We ran the classification algorithm on all samples using the appropriate σ_{crit} value for each dataset. To evaluate the success of the classification algorithm, we calculated the true positive rate (TPR), true negative rate (TNR), and overall accuracy (A) for each sample using the following equations:

$$TPR = \frac{TP}{TP + FN} \tag{3}$$

$$TNR = \frac{TN}{TN + FP} \tag{4}$$

$$A = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

where TP is a true positive, TN is a true negative, etc. In this case, a TP refers to the correct identification of surface hoar when it is present, a TN corresponds to the correct identification of an “Other” microstructure, a FP is when the algorithm makes an incorrect surface hoar classification, and a FN is when surface hoar is misclassified as “Other”.

2.5 Method repeatability assessment

Our final investigation was to test the spatial mapping capacity and repeatability of our texture-based classification on a new snow sample, one that was not involved in the initial analysis. Using the techniques outlined in Sect. 2.1.1, we prepared a snow sample of rounded grains and grew surface hoar atop roughly *half* of the surface area, while the other half remained covered. Thus, the resulting snow surface grain shape was a 1:1 ratio of SH:RG. We proceeded to run the classification algorithm (Sect. 2.4) on the mixed sample to produce a map of surface hoar extent, using the appropriate values of σ_{crit} for each dataset. Unfortunately, the lidar unit was no longer available at the time of this assessment, so only NIR-HSI was evaluated. We again quantified accuracy using Equations 3 – 5.

3 Results

Here, we discuss results at nadir orientations, hence ϴ = 0° (as defined in Fig. 3). The results of each case at ϴ = 10° were very similar to their nadir counterparts, and so for clarity we address only the latter here. Results for ϴ = 10° are presented in Appendix A.

3.1 Spatial and spectral texture analysis

We evaluated a grouped distribution of σ for surface hoar (Samples 24 – 26) against a grouped distribution containing all other samples for each instrument/illumination condition. We performed this evaluation across a variety of spatial resolutions spanning two orders of magnitude. Further, in the case of NIR-HSI where spectral data were available, we examined results over 188 bands from 951 – 1403 nm. Using NIR-HSI under both direct and diffuse illumination, we found that the median value of σ for surface hoar was nearly always greater than that of "Other" microstructure group, but with considerable spatial and spectral dependence (Fig. 7a and 7b). When evaluating the difference in medians (i.e., SH minus Other), a normal distribution is evident along the spatial (vertical) axis, with the largest differences observed at the 10 mm spatial resolution for both illumination cases. Spectrally, the difference in medians remains fairly constant, peaking in the ~1250 – 1350 nm range. The maximum increase was 383% at 1246 nm for diffuse illumination and 294% at 1369 nm for direct illumination. The differences in median values tend to be larger in the case of diffuse illumination relative to direct, and therefore we anticipate greater ease of surface hoar classification with diffuse lighting.

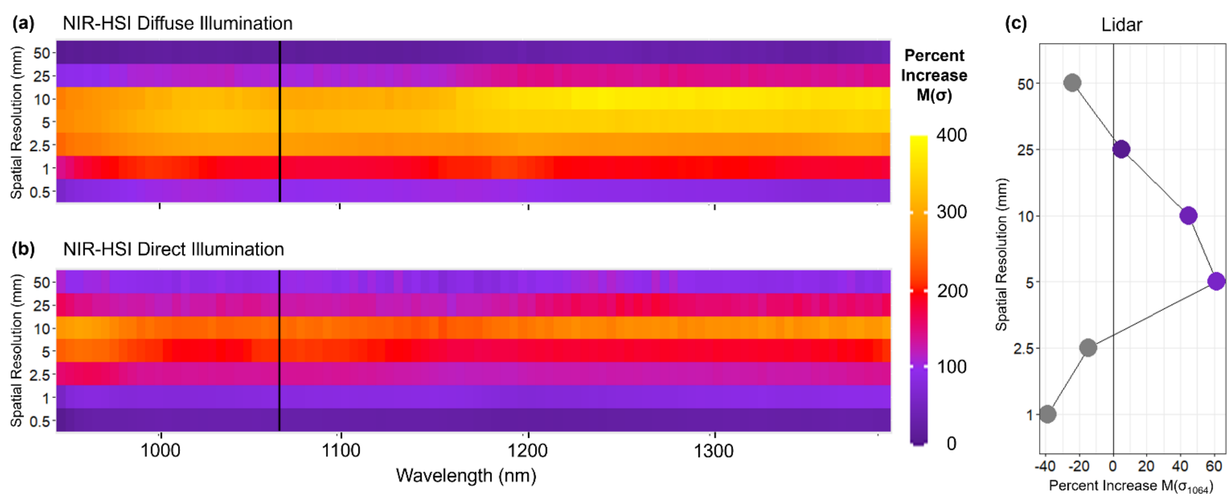

Figure 7: Heat maps depicting the percent difference in median values of σ for surface hoar samples versus all other microstructures (SH minus other) across a variety of spatial resolutions and spectral bands (a and b). Data were acquired via NIR-HSI under diffuse (a) and direct (b) illumination conditions. In (c), a line graph illustrates the same percent difference for lidar-derived σ_{1064} with points colored by the same scale when an increase is observed. Black vertical lines in (a) and (b) are located at the lidar wavelength of 1064 nm, for reference, while the grey vertical line in (c) is centered at zero.

To reiterate, when using NIR-HSI, median values of σ were larger in grouped SH distributions than other microstructure distributions across nearly all cases (Fig. 7a and 7b); at worst the medians were essentially equal. This provides evidence for our hypothesis that surface hoar will produce increased NIR reflectance texture under a variety of conditions. Our lidar observations, however, were less consistent. The mere presence of heightened texture in SH was dependent on spatial resolution (Fig. 7c). At the lowest (1 mm and 2.5 mm) and highest (50 mm) spatial resolutions, we observed larger medians in the "Other" microstructure group than in SH, a result that contrasts our hypothesis. A normal distribution of delta median values corresponding to spatial resolution is evident, as in NIR-HSI, with 5 mm and 10 mm datasets producing the largest (positive) differences. As discussed, we were limited to evaluation of a single band with our lidar unit (R_{1064}), although our NIR-HSI spectral results (Fig. 7a and 7b) indicate that 1064 nm is a suitable wavelength.

3.2 Samplewise statistical analysis and significance testing

We conducted a samplewise case study using NIR-HSI derived σ_{1324} (both direct and diffuse) at 10 mm and lidar σ_{1064} at 5 mm spatial resolution. These selections were based on our findings in Sect. 3.1; the spatial resolutions were the optimal choice for each instrument, and 1324 nm maximized NIR-HSI texture increases under both illumination conditions. When using NIR-HSI, we found that surface hoar exhibited larger values of σ_{1324} relative to other sample microstructures in both direct and diffuse illumination conditions. This is evident in Fig. 8d and 8e, while also outlining the problem with using NIR reflectance magnitude to delineate surface hoar. In both cases, the median reflectance magnitude (R_{1324}) gradually increased proceeding from lower to higher SSA (Fig. 8a and 8b), as expected. This leaves the reflectance magnitude of surface hoar "hidden" in the middle of all other grain shapes, with moderate median R_{1324} values (fuchsia horizontal reference lines). When examining values of σ_{1324} (Fig. 8d and 8e), a different pattern is evident. In Fig. 8d and 8e, median values of σ_{1324} are fairly constant across all samples with the exception of surface hoar, where a spike is present. The separation between the median value of σ_{1324} for surface hoar samples compared to the best-fit line of all other samples (black lines) further illustrates this and may allow for delineation of surface hoar based on a texture parameter. Our results for lidar R_{1064} indicate a similar trend, although the distinction is less pronounced (Fig. 8c and 8f).

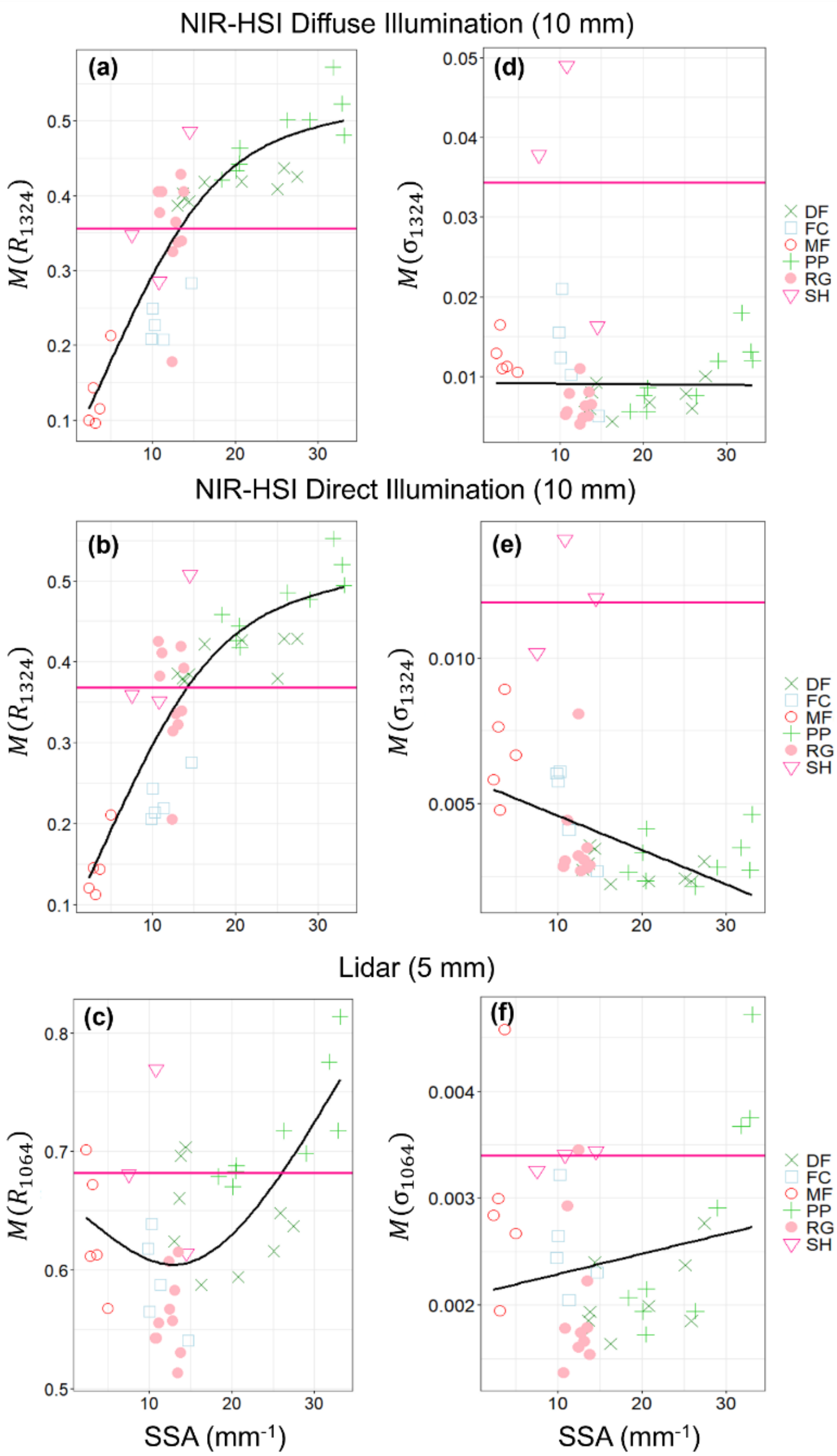

Figure 8: Samplewise median values of reflectance (a – c) and σ (d – f) for each instrument/illumination case study. The black lines are spline or linear best fits to all samples other than surface hoar, while fuchsia horizontal reference lines depict SH median values. Colors and point shapes correspond to those suggested by the ICSSG (Fierz et al., 2009).

To restate our central finding, when using NIR-HSI under both direct and diffuse illumination, surface hoar exhibited larger median values of σ_{1324} relative to other sample microstructures in all cases (Table 2, Fig. 9a and 9b), thus confirming our hypothesis. Furthermore, this increase was statistically significant in all cases. The in situ Samples 25 and 26 were particularly distinct, with interquartile ranges rarely overlapping those of any other sample, making these easily discernible. Interestingly, Sample 24, which consisted of sieved SH grains (Table 1), displayed heightened texture under direct illumination but not diffuse. As with Fig. 8, our lidar results were similar to those of NIR-HSI, but less pronounced (Fig. 9c), with particular confusion occurring with PP and MF samples. Still, the grouped SH median value of σ_{1064} was significantly larger than sample medians in 30/38 cases. Table 2 provides complete significance findings for all scenarios, with greyed cells denoting a case where the sample median σ value is significantly less than the surface hoar median.

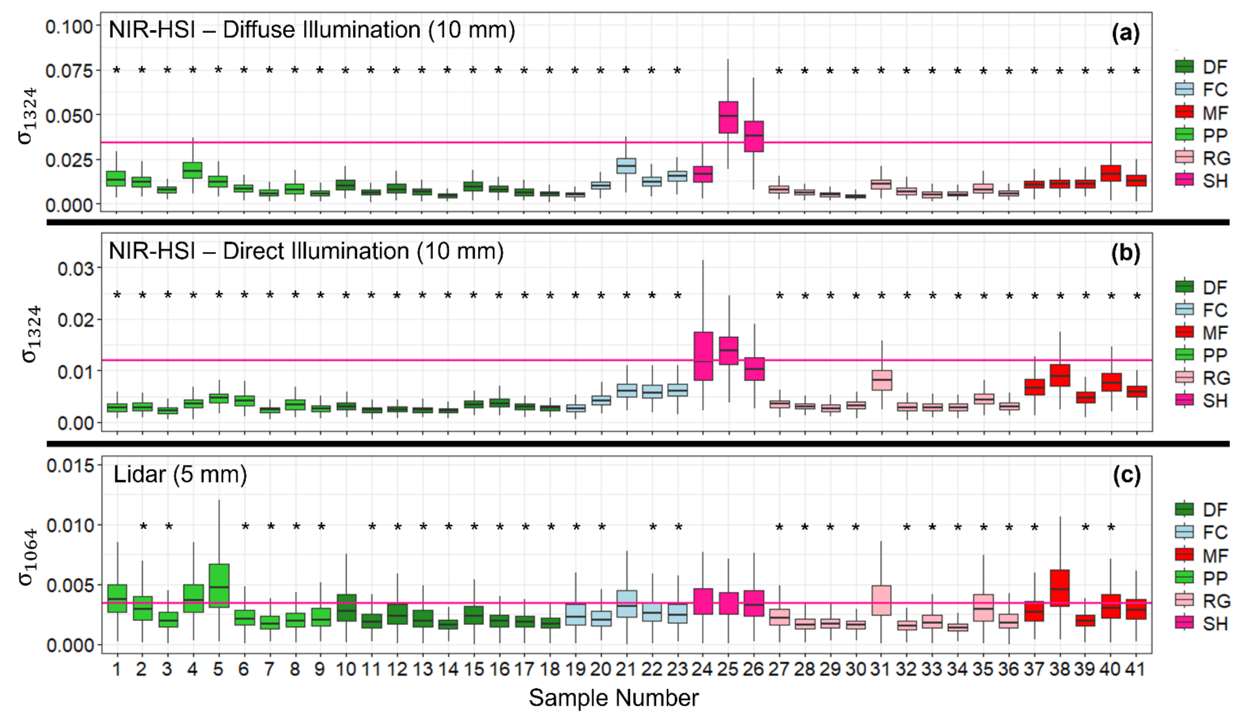

Figure 9: Samplewise boxplot analysis for all three instrument/illumination scenarios. Boxes are colored by primary grain shape following the ICSSG. Fuchsia horizontal reference lines depict the median value of σ for surface hoar Samples 24 – 26. Samples with an asterisk had median values significantly lower than surface hoar medians based on one-sample, one-sided t-tests at α = 0.05.

Table 2: Sample median values of σ for each instrument and illumination condition. Median σ values of surface hoar Samples 24 – 26 are shown at the top. Sample median σ values shaded in grey denote values significantly lower than the corresponding surface hoar medians at a significance level of α = 0.05.

		NIR-HSI Diffuse	NIR-HSI Direct	Lidar
		M (σ) for SH Samples 24 - 26		
		0.0343	0.0119	0.0034
Sample #	1° Grain Shape	M (σ_{1324})	M (σ_{1324})	M (σ_{1064})
1	PP	0.0132	0.0027	0.0037
2	PP	0.0120	0.0028	0.0029
3	PP	0.0076	0.0022	0.0019
4	PP	0.0180	0.0035	0.0037
5	PP	0.0121	0.0046	0.0047
6	PP	0.0086	0.0041	0.0021
7	PP	0.0057	0.0024	0.0017
8	PP	0.0076	0.0033	0.0019
9	PP	0.0057	0.0026	0.0021
10	DF	0.0100	0.0030	0.0028
11	DF	0.0061	0.0023	0.0019
12	DF	0.0079	0.0024	0.0024
13	DF	0.0068	0.0023	0.0020
14	DF	0.0044	0.0022	0.0016
15	DF	0.0093	0.0034	0.0024
16	DF	0.0080	0.0036	0.0019
17	DF	0.0059	0.0029	0.0019
18	DF	0.0056	0.0027	0.0017
19	FC	0.0051	0.0027	0.0023
20	FC	0.0102	0.0041	0.0020
21	FC	0.0211	0.0061	0.0032
22	FC	0.0124	0.0058	0.0026
23	FC	0.0155	0.0060	0.0024
24	SH	-	-	-
25	SH	-	-	-
26	SH	-	-	-
27	RG	0.0081	0.0035	0.0022
28	RG	0.0064	0.0031	0.0017
29	RG	0.0049	0.0027	0.0017
30	RG	0.0040	0.0032	0.0016
31	RG	0.0111	0.0081	0.0035
32	RG	0.0065	0.0029	0.0015
33	RG	0.0050	0.0028	0.0018
34	RG	0.0053	0.0028	0.0014
35	RG	0.0079	0.0044	0.0029
36	RG	0.0056	0.0030	0.0018
37	MF	0.0105	0.0067	0.0027
38	MF	0.0113	0.0089	0.0046
39	MF	0.0110	0.0048	0.0019
40	MF	0.0165	0.0076	0.0030
41	MF	0.0129	0.0058	0.0028

3.3 Classification algorithm

We used PDFs to compare the probability density distributions of surface hoar samples (Samples 24 – 26) against all other samples, which allowed us to determine critical texture thresholds (σ_{crit}) for surface hoar delineation under each instrument/illumination scenario. This process is illustrated in Fig. 10. Selecting the intersection point of the two distributions, following Champollion et al. (2013), allows for optimal binary classification of a given pixel. The resulting critical values for each condition are listed on the PDF plots. Using these σ_{crit} values, we conducted pixelwise classification of all samples). Examples of this binarization for several samples of varying primary grain shape are shown in Fig. 11. All samplewise accuracy values (A) are listed in Table 3, while median values of A, TPR, and TNR for each scenario are included on the PDF plots in Fig. 10. Classification results generally mirrored those of the statistical analysis. That is, results were excellent for NIR-HSI under both direct and diffuse illumination, but dwindled substantially when using lidar-derived σ_{1064}. The classification algorithm generally struggled most with samples of a FC or MF primary grain shape.

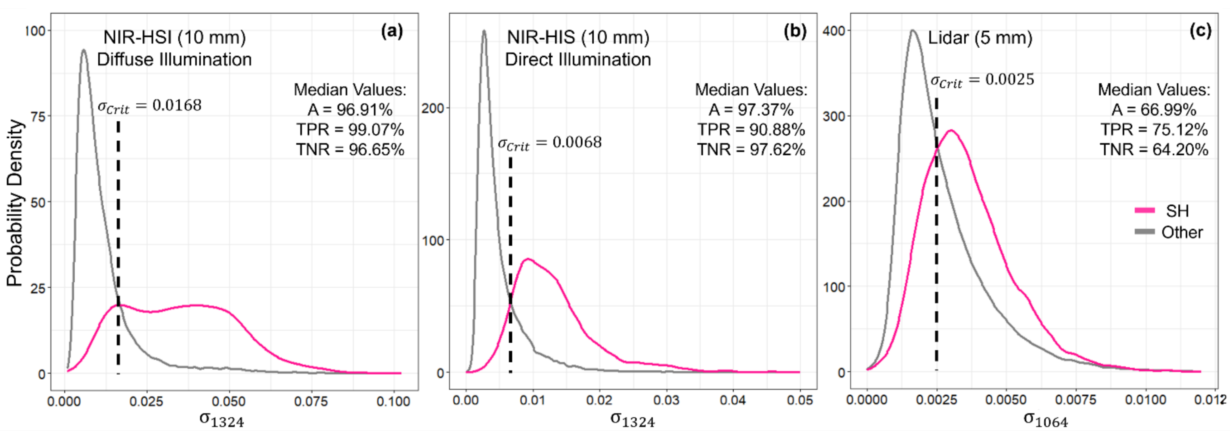

Figure 10: Probability density functions juxtaposing the σ distributions of surface hoar, Samples 24 – 26 (fuchsia curves), with those of all other samples (grey curves). Dotted vertical reference lines represent the distribution intersection, where the critical threshold values (σ_{crit}) were extracted. Median accuracy metrics from the resulting samplewise binary classifications are also listed for each scenario.

Table 3: Samplewise and median accuracy (A) values for each instrument/illumination combination. Results for surface hoar, Samples 24 – 26, are shaded in fuchsia.

	NIR-HSI Diffuse	NIR-HSI Direct	Lidar
		σ_{crit}	
	0.0168	0.0068	0.0025
Sample #		A (%)	
1	71.67	99.85	20.26
2	80.81	98.02	38.31
3	99.48	100.00	69.69
4	42.19	95.45	20.89
5	80.67	91.09	15.60
6	97.27	93.51	63.71
7	98.62	100.00	80.09
8	88.01	99.31	72.02
9	100.00	100.00	64.68
10	89.50	97.34	43.93
11	100.00	100.00	74.31
12	91.18	100.00	53.45
13	96.91	100.00	66.99
14	100.00	100.00	90.67
15	93.35	97.37	54.83
16	99.15	94.01	75.43
17	88.97	98.81	80.64
18	100.00	100.00	84.12
19	100.00	99.71	56.27
20	94.16	95.44	67.89
21	22.71	66.05	32.27
22	83.87	68.61	44.99
23	60.34	65.59	51.89
24	47.76	86.10	75.12
25	100.00	97.20	75.97
26	99.07	90.88	71.58
27	98.04	95.24	60.70
28	100.00	100.00	86.56
29	100.00	99.87	87.04
30	100.00	97.63	93.47
31	93.83	31.25	25.54
32	98.66	97.62	91.21
33	100.00	100.00	75.55
34	100.00	98.99	96.40
35	97.30	89.08	39.34
36	100.00	99.57	74.68
37	96.39	52.66	44.08
38	93.42	22.10	12.32
39	91.74	85.85	76.60
40	52.24	37.43	35.75
41	79.27	72.25	37.70
Median	**96.91**	**97.37**	**66.99**

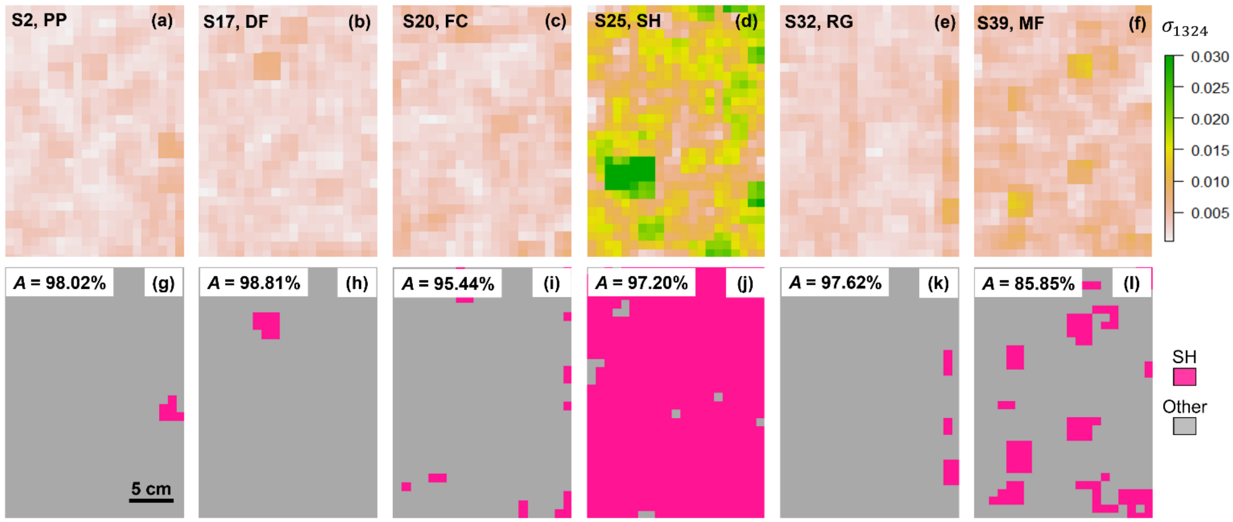

Figure 11: Example transformations from maps of σ_{1324} (a – f) to pixelwise classifications (g – l) via binarization with a critical threshold. The data shown here are from direct illumination at 10 mm spatial resolution.

3.4 Classification assessment

Classification accuracy assessed on a new sample, comprised of a 1:1 ratio of RG and SH surface grain shapes, proved consistent, demonstrating repeatability of the texture phenomenon when using NIR-HSI (Fig. 12). Further, we observe the capacity to map surface hoar extent amid mixed surface conditions. In the R_{1324} maps (Fig. 12b and 12c) the heterogeneity of surface hoar reflectance is already apparent compared to the RGs, and then quantified via a moving window analysis (Fig. 12d and 12e). Using the appropriate values of σ_{crit} (Sect. 3.3), we binarized each texture image to create a classified data product (Fig. 12f and 12g). In the classified maps, correct rejection of the RG surface (hence TNR) was perfect (100.00%) for diffuse illumination and suitable (92.23%) under direct illumination. Accurate identification of SH (i.e., TPR) was also excellent, at 99.35% and 99.32% for diffuse and direct, respectively, while overall accuracy was 99.61% and 96.15%.

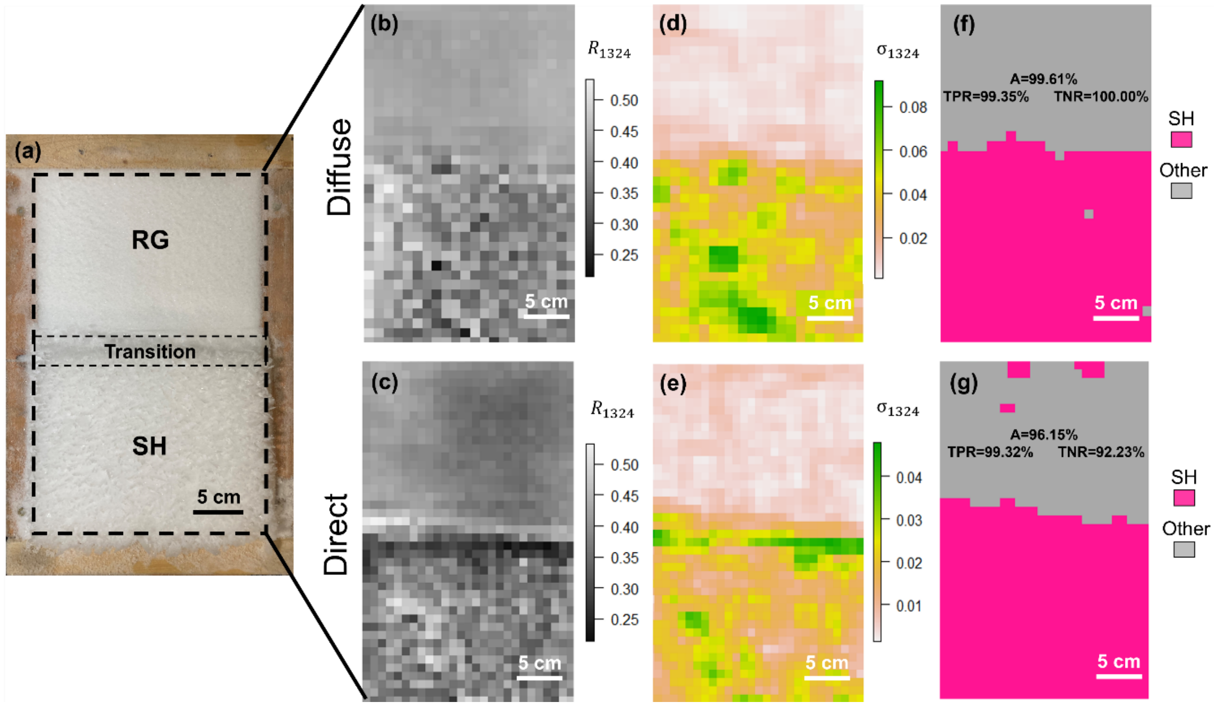

Figure 12: A sample with a 50:50 ratio of RG:SH surface grain shapes is displayed in the visible (a). Under both diffuse (upper row) and direct (lower row) illumination conditions, the bidirectional reflectance factor at 1324 nm is extracted (b and c) from NIR-HSI. Via a moving window analysis, localized standard deviation (σ_{1324}), or NIR texture, is quantified (d and e), and then binarized using critical thresholds to produce classified maps (f and g). Overall accuracy, true positive rates, and true negative rates are listed on the classified maps. The transitional zone was excluded in accuracy analyses.

4 Discussion

4.1 NIR-HSI

We found that, when measured with NIR-HSI, surface hoar exhibited larger median values of σ, a NIR texture metric, than other snow microstructures across a variety of spatial resolutions and spectral bands. Furthermore, in a samplewise case study evaluating σ_{1342} at 10 mm spatial resolution, we determined that median surface hoar values were significantly larger than other snow surface structures (primary grain habits of PP, DF, FC, RG and MF) in all cases under both direct and diffuse illumination (one-sided, one-sample t-test, $\alpha = 0.05$). Our findings are consistent with Champollion et al. (2013), who found that under artificial lighting conditions the NIR texture (in this case a contrast index) of the Antarctic snow surface was higher when surface hoar was present as compared to when it was absent. These researchers used a passive NIR camera and downward-looking lights, all mounted ~ 2 m above a flat snow surface, to create a field setup akin to the direct illumination ($\Theta = 0°$) case presented here. Although they do not explicitly mention the spatial resolution of their images, they imply that individual surface hoar crystals spanned 5 – 10 pixels, and thus their resolution was likely on the order of millimeters. While Champollion et al. (2013) simply juxtaposed the presence versus absence of surface hoar, we build on these foundational findings by quantifying the texture phenomenon against a variety of differing, thoroughly characterized snow microstructures in a controlled laboratory environment. Further, our work features the explicit inclusion of diffuse illumination, varied illumination/viewing incidence angle, and a range of spatial resolutions and spectral bands. Last, whereas Champollion et al. (2013) used texture metrics across an entire image to confirm the presence of hoar crystals on a given day (with accuracy of 94%), we extend this classification by demonstrating a per-pixel mapping methodology, allowing us to map the spatial extent of surface hoar within an image (Fig. 12). Median sample classification accuracy was 96.91% and 97.37% for diffuse and direct illumination, respectively. As discussed in Sect. 3.1 and illustrated in Fig. 8, surface hoar exhibits moderate NIR reflectance when compared to other grain shapes, and thus leveraging reflectance magnitude alone is insufficient for surface hoar delineation. These results highlight the importance of our texture-based approach.

Spatial resolution was perhaps the factor with the most influence on our results, with the maximal difference between surface hoar and other median values occurring at 10 mm resolution for both illumination cases (Fig. 7). However, consistent differences in median values can be observed at nearly all resolutions beneath 50 mm, indicating that classification at coarser resolutions may still be suitable under the right conditions. This trend makes sense physically; we attribute the rise in texture metrics associated with surface hoar to be related to the lateral spacing between large hoar crystals (Fig. 1). Given that this spacing is often on the order of millimeters, it is possible that the raw resolution is too fine to optimally observe this variability, while 50 mm resolution is too coarse. Further, the fact that results were fairly consistent spectrally from 951 – 1403 nm implies that hyperspectral capacity is likely unnecessary; a broadband passive NIR sensor could likely observe the same texture increases. Samples that proved the most challenging to discern from SH were FC (Samples 19 – 23), and the large MF grains from Batch L (Samples 37 – 41). The latter was likely due to enhanced surface roughness, as these grains were several mm in length (Fig. 2), a rather extreme case, and indicates that caution should be used when conditions favor wet snow metamorphism. Sporadic misclassification of FC (e.g., Fig. 11i) is perhaps understandable, given that these grains form from kinetic snow metamorphism, which is somewhat similar to surface hoar growth. In practice, it may be useful to try to identify these surfaces as well, as near-surface FC also tend to act as weak layers once buried. The relatively lower median σ_{1342} value of the sieved SH Sample 24 under diffuse illumination (Fig. 9a) is an interesting result. This anomaly perhaps implies that the texture signature dissipates when the predominately vertical orientation of SH grains is interrupted, and thus surface hoar that has been blown over is likely more difficult to detect. The low accuracy value of this sample under diffuse illumination (47.76%, Table 3) is evidence of this. The uncertainty involved in SH classification for any case can be quantified by examining the area under the intersecting PDF curves (Fig. 10). Furthermore, uncertainty with regards to a specific microstructure (such as FC) could be observed by comparing PDFs between individual samples, rather than grouping SH samples versus all other samples. However, our goal was to determine robust thresholds for delineation of SH from any other microstructure.

4.2 Lidar

For as impressive as the performance of NIR-HSI was, the results of our lidar analysis were more perplexing. In Fig. 7c, we can observe that the median value of σ_{1064} for SH Samples 24 – 26 was only greater than that of other microstructures at spatial resolutions of 5, 10, and 25 mm. At the largest and smallest resolutions, much like with NIR-HSI, performance diminished. However, unlike NIR-HSI, in these cases the median SH value was actually *lower* than the other microstructure median, a result that runs counter to our hypothesis. Even at 5 mm resolution, our lidar classification results were considerably less robust than those of NIR-HSI, with median sample accuracy of 66.99%. Unexpected results in texture analyses are not unheard of. While Champollion et al. (2013) noted increases in NIR texture when surface hoar was present under direct, "artificial" illumination at nadir incidence (as we did with NIR-HSI), they also documented a reversal of this trend under natural (solar) illumination. However, this was at very large (80° - 85°) solar zenith angles, and thus represents an extreme case of grazing incidence, and the cause of this observation was not thoroughly explained. For the laboratory setup used here, we expected lidar to reproduce similar results to those of our NIR-HSI analysis and to the findings of Champollion et al. (2013) using artificial light. Like NIR-HSI, the lidar classification algorithm struggled with the large MFs of Batch L, as well as PP (Fig. 9c).

In general, the use of lidar reflectance to ascertain snow surface properties is far less validated relative to passive imagery, like NIR-HSI. While passive NIR imagery has been used to estimate SSA or r_e for decades, leveraging the well-established sensitivity of NIR reflectance to snow microstructure, very few studies (e.g., Yang et al., 2017) have attempted to do so with lidar. Therefore, the dependence of NIR lidar reflectance on snow microstructure is substantially less understood. This is important because lidar scanning represents a unique bidirectional reflectance scenario. For instance, the beam emitted from lidar units is collimated, and thus it experiences less loss due to scattering and absorption compared to direct solar irradiance, and results in higher irradiance at the wavelength of interest. Collimation alters scattering and reflectance in optically rough materials like snow relative to an un-collimated illumination source (Murphy, 2006). Additionally, lidar beams are predominately linearly polarized, which contrasts with the often randomly polarized nature of solar illumination (Sassen, 2005). The scattering process is partly polarizing, and therefore multiple scattering can have significant polarization dependence (Li et al., 2008; Bhandari et al., 2011). Last, lidar presents a monostatic geometric condition on bidirectional reflectance, such that every measurement strictly observes direct backscatter. This is beneficial in that it limits the number of characterizations required compared to a full bidirectional reflectance investigation. However, observations of snow reflectance at this geometry are lacking, and radiative transfer models are not well-validated for the case of direct backscatter. Thus, lidar is an optically unique case of bidirectional reflectance that requires careful examination, so it is not entirely surprising that a texture analysis of lidar reflectance produced peculiar results.

One possible explanation for the relatively poor performance of our algorithm when using lidar reflectance can be found in the work of Walter et al. (2023), who leveraged a laboratory setup much like the apparatus presented here. These researchers focused on a temporal analysis, using a 905 nm lidar to observe changes in the mean reflectance magnitude and standard deviation (across an entire image) at prescribed time intervals as they grew surface hoar atop a layer of compressed PP. Consistent with our spatial analysis, they noted that reflectance magnitude was insufficient to characterize surface hoar growth, as the mean reflectance changed only 4%, while standard deviation increased as much as 600%. Although this juxtaposition features only a single microstructure other than SH (compressed PP), these lidar-based results seem more encouraging than our own, and are more consistent with our NIR-HSI findings. A key distinction is, perhaps, the lidar spot size. Their spot size of 3 mm is much smaller than that of our lidar/setup (15 mm), and likely closer to the size and spacing of individual surface hoar crystals in many cases. Indeed, the authors predict that surface hoar crystals smaller than the lidar beam diameter will not be detectable, and thus it is possible that our beam diameter is simply too coarse for this application.

4.3 Future work and speculations on scaling to field applications

Though scaling to field applications presents considerable challenges and uncertainty, it is likely that our NIR-HSI findings can be extended to operational avalanche forecasting in the near future. A simple setup like that of Champollion et al. (2013), where a downward-looking NIR camera acquired daily and nightly images, could be installed at remote weather stations to monitor the formation and persistence of surface hoar layers prior to burial. Such remote measurements are not currently available but are critically important for avalanche forecasters. Further, UAV snow mapping using NIR-HSI has recently been demonstrated as a useful tool to measure snow grain size and albedo at the slope scale (Skiles et al. 2023) and future studies should consider texture analyses. To accomplish this, a greater variety of incidence angles must be evaluated. Although it is encouraging that our results remained promising at $\Theta = 10°$ (Appendix A), more oblique angles will inevitably be encountered in the field. Even on overcast days, where diffuse solar illumination is prevalent, variance in terrain slope beneath a downward looking imager would still alter the viewing incidence angle and thus the magnitude, and likely texture, of NIR snow reflectance. Using direct solar illumination would add another factor, the illumination incidence angle, although this could be kept consistent by using artificial lighting. Ideal thresholds of σ_{crit} will likely continue to vary between incidence angles, as well as between instruments/applications. Further, the flight plan and/or optical logistics would need to be controlled to ensure adequate spatial resolution. Another factor that may be worth investigating is the window size during focal analysis, although a preliminary study determined that the neighborhood size was of little consequence for our data.

Although we realized limited success with lidar, the use of lidar for surface hoar mapping via texture requires further evaluation, particularly with regards to beam diameter. A better physical understanding regarding the resolution-dependence of our lidar results is needed. At a minimum, lidar could be useful in conjunction with NIR-HSI or other passive NIR detection. Using the spatial capacity of lidar, values of slope angle (and thus incidence angle if using a downward-oriented imager), could be determined on a per-pixel basis, allowing for the proper threshold value to be used when analyzing NIR-HSI texture data. Though challenging, slope, or even basin-scale mapping of surface hoar extent is likely possible with current technology, though a field campaign would be required to tune our method to a wider variety of illumination conditions and incidence angles. Maps of surface hoar extent over avalanche terrain prior to burial would be vital in making slope-specific avalanche forecasts. Such maps could also identify likely avalanche trigger points, improving avalanche mitigation with explosives, for example. While our work focused on surface hoar classification, texture analysis will likely provide a useful tool for evaluating other physical phenomena as well, such as surface roughness.

5 Conclusion

Our research demonstrates a novel method for mapping the spatial extent of surface hoar using NIR texture in a cold laboratory. In essence, we found that:

I. Hyperspectral imaging can robustly measure the texture of snow and ice by computing pixelwise variability in reflectance at any NIR wavelength.

II. When evaluated with NIR-HSI, surface hoar has significantly heightened NIR texture relative to other snow microstructures across a range of spectral bands and spatial resolutions, likely as a result of variable ice absorption and specular contributions.

III. Near-infrared texture thresholds can be used to binarize NIR-HSI texture measurements, resulting in accurate maps of surface hoar spatial extent on a per-pixel basis.

IV. Similar results were achieved with 1064 nm lidar, although the phenomenon was resolution-dependent and performance was substantially less robust. The use of lidar for this purpose requires further investigation, and is likely dependent on beam diameter.

As NIR-HSI and lidar become more economical, these may provide capable methods for measuring NIR texture and subsequently mapping surface hoar. Extending the work presented here to field operations will have immediate implications for broad-scale snow surface mapping and avalanche forecasting.

6 Appendix A

Results for the Θ = 10° incidence viewing angle case are presented below.

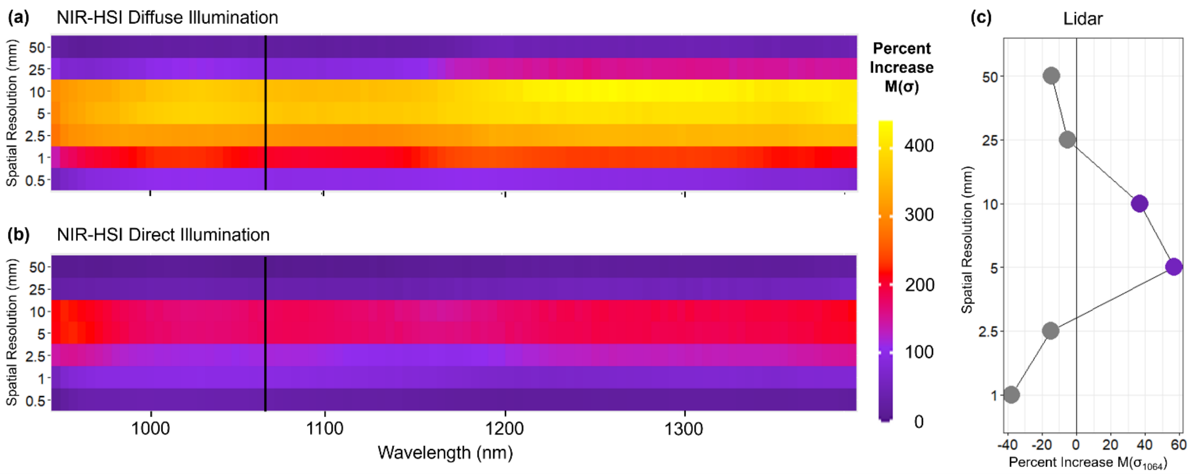

Figure A1: Heat maps depicting the percent difference in median values of σ for surface hoar samples versus all other microstructures (SH minus other) across a variety of spatial resolutions and spectral bands (a and b). Data were acquired via NIR-HSI under diffuse (a) and direct (b) illumination conditions. In (c), a line graph illustrates the same percent difference for lidar-derived σ_{1064} with points colored by the same scale when an increase is observed. Black vertical lines in (a) and (b) are located at the lidar wavelength of 1064 nm, for reference, while the grey vertical line in (c) is centered at zero. Relative to the nadir case, performance generally improved for NIR-HSI under diffuse illumination, but decreased under direct illumination for both NIR-HSI and lidar.

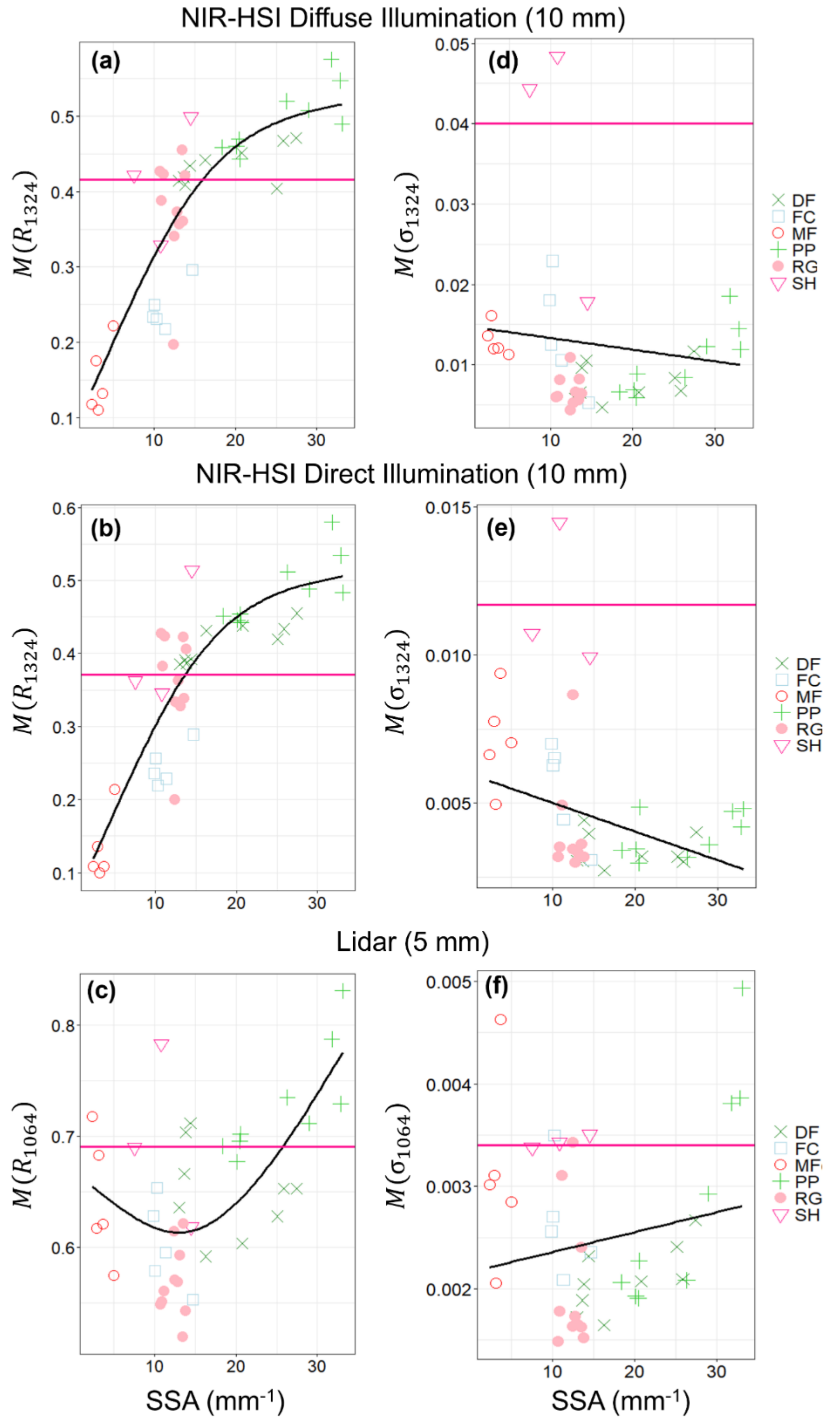

Figure A2: Samplewise median values of reflectance (a – c) and σ (d – f) for each instrument/illumination case study. The black lines are spline or linear best fits to all samples other than surface hoar, while fuchsia horizontal reference lines depict SH median values. Colors and point shapes correspond to those suggested by the ICSSG (Fierz et al., 2009).

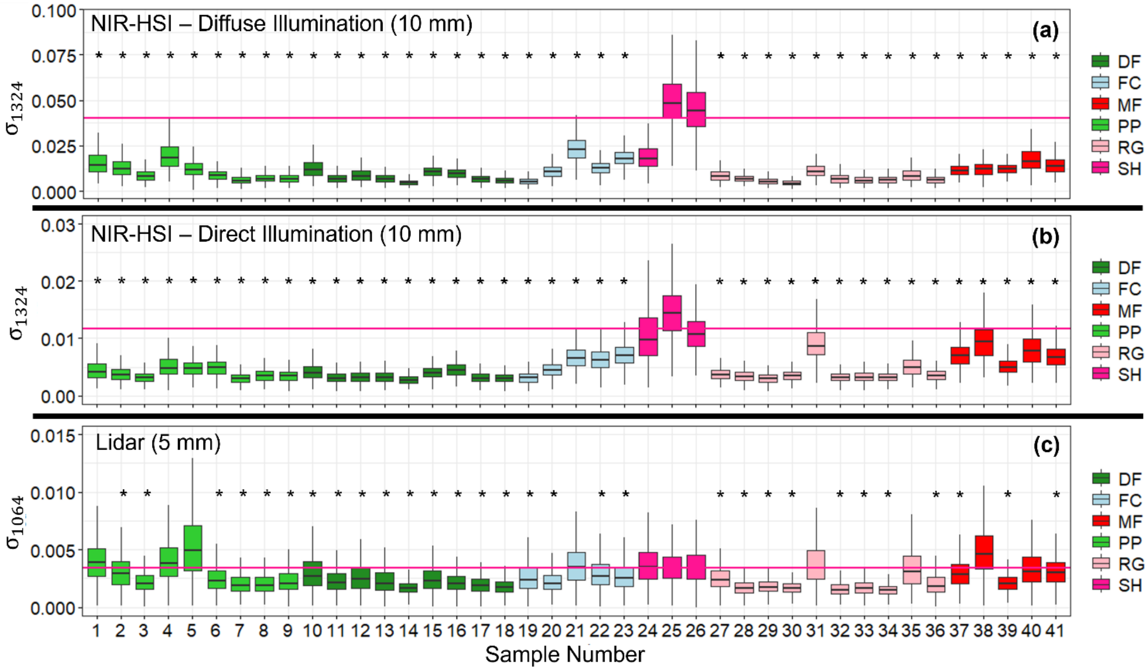

Figure A3: Samplewise boxplot analysis for all three instrument/illumination scenarios. Boxes are colored by primary grain shape following the ICSSG. Fuchsia horizontal reference lines depict the median value of σ for surface hoar Samples 24 – 26. Samples with an asterisk had median values significantly lower than surface hoar medians based on one-sample, one-sided t-tests at α = 0.05.

Table A1: Sample median values of σ for each instrument and illumination condition. Median σ values of surface hoar Samples 24 – 26 are shown at the top. Sample median σ values shaded in grey denote values significantly lower than the corresponding surface hoar medians at a significance level of α = 0.05.

		NIR-HSI Diffuse	NIR-HSI Direct	Lidar
		M (σ) for SH Samples 24 - 26		
		0.0400	0.0117	0.0034
Sample #	1° Grain Shape	M (σ_{1030})	M (σ_{1030})	M (σ_{1064})
1	PP	0.0145	0.0042	0.0039
2	PP	0.0123	0.0036	0.0029
3	PP	0.0084	0.0032	0.0021
4	PP	0.0186	0.0047	0.0038
5	PP	0.0119	0.0048	0.0049
6	PP	0.0089	0.0049	0.0023
7	PP	0.0059	0.0030	0.0019
8	PP	0.0069	0.0034	0.0019
9	PP	0.0066	0.0034	0.0021
10	DF	0.0117	0.0040	0.0027
11	DF	0.0067	0.0030	0.0021
12	DF	0.0084	0.0032	0.0024
13	DF	0.0066	0.0032	0.0021
14	DF	0.0047	0.0027	0.0016
15	DF	0.0105	0.0040	0.0023
16	DF	0.0097	0.0044	0.0020
17	DF	0.0065	0.0030	0.0019
18	DF	0.0058	0.0031	0.0017
19	FC	0.0053	0.0031	0.0024
20	FC	0.0106	0.0045	0.0021
21	FC	0.0229	0.0065	0.0035
22	FC	0.0125	0.0063	0.0027
23	FC	0.0180	0.0070	0.0026
24	SH	-	-	-
25	SH	-	-	-
26	SH	-	-	-
27	RG	0.0082	0.0036	0.0024
28	RG	0.0066	0.0033	0.0016
29	RG	0.0053	0.0030	0.0017
30	RG	0.0044	0.0035	0.0016
31	RG	0.0109	0.0087	0.0034
32	RG	0.0065	0.0032	0.0015
33	RG	0.0057	0.0032	0.0016
34	RG	0.0060	0.0032	0.0015
35	RG	0.0082	0.0049	0.0031
36	RG	0.0061	0.0035	0.0018
37	MF	0.0113	0.0070	0.0028
38	MF	0.0121	0.0094	0.0046
39	MF	0.0120	0.0050	0.0021
40	MF	0.0161	0.0078	0.0031
41	MF	0.0136	0.0066	0.0030

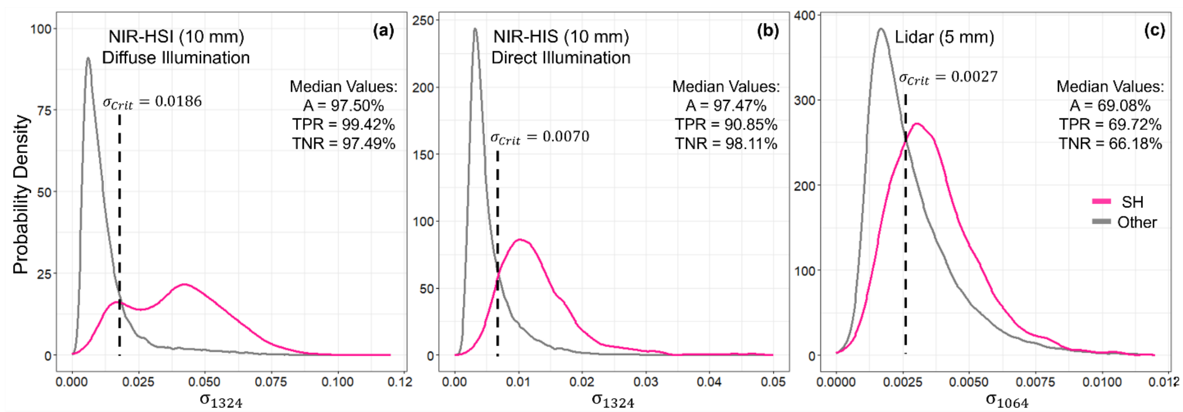

Figure A4: Probability density functions juxtaposing the σ distributions of surface hoar, Samples 24 – 26 (fuchsia curves), with those of all other samples (grey curves). Dotted vertical reference lines represent the distribution intersection, where the critical threshold values (σ_{crit}) were extracted. Median accuracy metrics from the resulting samplewise binary classifications are also listed for each scenario.

Table A2: Samplewise and median accuracy (A) values for each instrument/illumination combination. Results for surface hoar, Samples 24 – 26, are shaded in fuchsia.

	NIR-HSI Diffuse	NIR-HSI Direct	Lidar
		σ_{crit}	
	0.0186	0.0070	0.0027
Sample #		A (%)	
1	72.33	89.15	24.87
2	83.13	95.96	45.25
3	99.68	99.23	73.05
4	50.38	83.29	24.38
5	85.71	90.30	18.42
6	99.23	90.61	63.28
7	99.27	100.00	77.67
8	97.84	97.86	78.17
9	99.68	100.00	70.60
10	85.13	93.71	51.30
11	100.00	99.47	69.40
12	92.55	99.79	58.86
13	97.50	99.52	69.08
14	100.00	99.86	92.61
15	95.87	98.92	62.54
16	99.56	93.72	75.56
17	99.71	98.88	83.55
18	99.72	100.00	87.14
19	100.00	99.03	59.43
20	97.30	92.69	71.82
21	26.82	59.36	31.77
22	91.19	65.53	49.82
23	54.19	49.75	54.41
24	45.48	75.78	69.72
25	99.49	97.47	71.65
26	99.42	90.85	68.56
27	97.47	98.79	61.54
28	99.85	99.88	89.85
29	100.00	99.75	90.04
30	100.00	99.55	95.68
31	96.31	23.70	31.09
32	99.32	100.00	90.86
33	100.00	98.37	86.31
34	98.78	99.88	98.35
35	97.39	85.68	41.07
36	100.00	98.51	77.23
37	96.67	49.46	46.40
38	91.27	25.22	14.25
39	97.30	86.69	77.62
40	61.71	41.65	39.89
41	81.77	56.24	39.62
Median	**97.50**	**97.47**	**69.08**

7 Acknowledgements

This work was funded by the Transportation Avalanche Research Pooled Fund Program (TARP), administered through the Colorado Department of Transportation (CDOT), and by NASA Grant 80NSSC22K0694 from the Terrestrial Hydrology Program. We acknowledge the services and equipment (Riegl VZ-6000) provided by the GAGE Facility, operated by UNAVCO, Inc., with support from the National Science Foundation and the National Aeronautics and Space Administration under NSF Cooperative Agreement EAR-1724794. We acknowledge the use of the Subzero Research Laboratory in the Department of Civil Engineering at Montana State University and thank Ladean McKittrick for laboratory assistance. We would like to thank Resonon, Inc. for providing us with a hyperspectral imager and technical assistance. Last, we thank Joseph Shaw, Nathaniel Field and Riley Logan for technical guidance regarding optical data acquisition.

8 References

Abe, O., & Kosugi, K. (2019). Twenty-year operation of the Cryospheric Environment Simulator. Bulletin of Glaciological Research, 37, 53-65.

Bhandari, A., Hamre, B., Frette, Ø., Zhao, L., Stamnes, J. J., & Kildemo, M. (2011). Bidirectional reflectance distribution function of Spectralon white reflectance standard illuminated by incoherent unpolarized and plane-polarized light. Applied Optics, 50(16), 2431-2442.

Birkeland, K. W. (1998). Terminology and predominant processes associated with the formation of weak layers of near-surface faceted crystals in the mountain snowpack. Arctic and Alpine Research, 30(2), 193-199.

Bühler, Y., Meier, L., & Ginzler, C. (2014). Potential of operational high spatial resolution near-infrared remote sensing instruments for snow surface type mapping. IEEE Geoscience and Remote Sensing Letters, 12(4), 821-825.

Champollion, N., Picard, G., Arnaud, L., Lefebvre, E., & Fily, M. (2013). Hoar crystal development and disappearance at Dome C, Antarctica: observation by near-infrared photography and passive microwave satellite. The Cryosphere, 7(4), 1247-1262.

Domine, F., Taillandier, A. S., & Simpson, W. R. (2007). A parameterization of the specific surface area of seasonal snow for field use and for models of snowpack evolution. Journal of Geophysical Research: Earth Surface, 112(F2).

Donahue, C., Skiles, S. M., & Hammonds, K. (2021). In situ effective snow grain size mapping using a compact hyperspectral imager. Journal of Glaciology, 67(261), 49-57.

Donahue, C., Skiles, S. M., & Hammonds, K. (2022). Mapping liquid water content in snow at the millimeter scale: an intercomparison of mixed-phase optical property models using hyperspectral imaging and in situ measurements. The Cryosphere, 16(1), 43-59.

Fierz, C. R. L. A., Armstrong, R. L., Durand, Y., Etchevers, P., Greene, E., McClung, D. M., ... & Sokratov, S. A. (2009). The international classification for seasonal snow on the ground.

Grenfell, T. C., & Warren, S. G. (1999). Representation of a nonspherical ice particle by a collection of independent spheres for scattering and absorption of radiation. Journal of Geophysical Research: Atmospheres, 104(D24), 31697-31709.

Horton, S., & Jamieson, B. (2017). Spectral measurements of surface hoar crystals. Journal of Glaciology, 63(239), 477-486.

Jamieson, J. B., & Schweizer, J. (2000). Texture and strength changes of buried surface-hoar layers with implications for dry snow-slab avalanche release. Journal of Glaciology, 46(152), 151-160.

Kokhanovsky, A. A., & Zege, E. P. (2004). Scattering optics of snow. Applied optics, 43(7), 1589-1602.

Li, D., Zhao, M. H., Garra, J., Kolpak, A. M., Rappe, A. M., Bonnell, D. A., & Vohs, J. M. (2008). Direct in situ determination of the polarization dependence of physisorption on ferroelectric surfaces. nature materials, 7(6), 473-477.

Lorensen, W. E., & Cline, H. E. (1987). Marching cubes: A high resolution 3D surface construction algorithm. ACM siggraph computer graphics, 21(4), 163-169.

Lutz, E. R., & Birkeland, K. W. (2011). Spatial patterns of surface hoar properties and incoming radiation on an inclined forest opening. Journal of Glaciology, 57(202), 355-366.

Matzl, M., & Schneebeli, M. (2010). Stereological measurement of the specific surface area of seasonal snow types: Comparison to other methods, and implications for mm-scale vertical profiling. Cold Regions Science and Technology, 64(1), 1-8.

Murphy, A. B. (2006). Modified Kubelka–Munk model for calculation of the reflectance of coatings with optically-rough surfaces. Journal of Physics D: Applied Physics, 39(16), 3571.

Nolin, A. W., & Dozier, J. (2000). A hyperspectral method for remotely sensing the grain size of snow. Remote sensing of Environment, 74(2), 207-216.

Sassen, K. (2005). Polarization in lidar. In Lidar (pp. 19-42). Springer, New York, NY.

Schleef, S., Jaggi, M., Löwe, H., & Schneebeli, M. (2014). An improved machine to produce nature-identical snow in the laboratory. Journal of Glaciology, 60(219), 94-102.

Skiles, S. M., Donahue, C. P., Hunsaker, A. G., & Jacobs, J. M. (2023). UAV hyperspectral imaging for multiscale assessment of Landsat 9 snow grain size and albedo. Frontiers in Remote Sensing, 3, 1038287.Stanton, B., Miller, D., Adams, E., & Shaw, J. A. (2016). Bidirectional-reflectance measurements for various snow crystal morphologies. Cold Regions Science and Technology, 124, 110-117.

Walter, B., Brouet, L., Jaggi, M., & Löwe, H. (2023). Automated LiDAR remote sensing for measuring the spatial and temporal evolution of surface hoar formation. In International Snow Science Workshop (ISSW), Bend, OR, USA.

Yang, Y., Marshak, A., Han, M., Palm, S. P., & Harding, D. J. (2017). Snow grain size retrieval over the polar ice sheets with the Ice, Cloud, and land Elevation Satellite (ICESat) observations. Journal of Quantitative Spectroscopy and Radiative Transfer, 188, 159-164.

CHAPTER FIVE

DOCUMENTING, QUANTIFYING, AND MODELING A LARGE GLIDE AVALANCHE IN GLACIER NATIONAL PARK, MONTANA, USA

Contribution of Authors and Co-Authors

Manuscript in Chapter 5

Author: James Dillon

Contributions: Conceptualized the study, collected field data, performed analysis and modeling, and drafted the manuscript.

Co-Author: Erich Peitzsch

Contributions: Provided guidance and local practitioner expertise during data collection and analysis, supplied meteorological and historical avalanche data.

Co-Author: Zach Miller

Contributions: Provided guidance and local practitioner expertise during data collection and analysis, supplied meteorological and historical avalanche data.

Co-Author: Perry Bartelt

Contributions: Provided technical guidance regarding dynamics modeling.

Co-Author: Kevin Hammonds

Contributions: Acquired funding, was responsible for project administration, and supervised throughout.

Manuscript Information

James Dillon, Erich Peitzsch, Zach Miller, Perry Bartelt, Kevin Hammonds

Status of Manuscript:
X Prepared for submission to a peer-reviewed journal
☐ Officially submitted to a peer-reviewed journal
☐ Accepted by a peer-reviewed journal
☐ Published in a peer-reviewed journal

Documenting, quantifying, and modeling a large glide avalanche in Glacier National Park, Montana, USA

James Dillon[1], Erich Peitzsch[2], Zachary Miller[2], Perry Bartelt[3], Kevin Hammonds[1]
[1]Department of Civil Engineering, Montana State University, Bozeman, MT, USA
[2]U.S. Geological Survey Northern Rocky Mountain Science Center, West Glacier, MT, USA
[3]WSL, Swiss Federal Institute for Snow and Avalanche Research, SLF, Davos Dorf, Switzerland

Correspondence to: James Dillon (james.dillon013@gmail.com)

Abstract

Glide avalanches pose a significant and repetitive threat to many operational highway and railroad forecasting programs, and they are likely to become more frequent. While the spatial location of glide release areas is extremely consistent, the timing of glide events is notoriously difficult to forecast, and their destructive potential can be immense. Thus, the timing and dynamics of glide avalanches is an important area of study. To better understand these processes, and to improve assessments of risk to roadways, event documentation is key. Here, we survey a large glide avalanche event along the Going-to-the-Sun Road in Glacier National Park, Montana, USA during road opening operations in the spring of 2022. Using three sets of terrestrial lidar data (pre-event, post-event, and snow-off), we quantified key aspects of the avalanche and created powerful visualizations for analysis. Further, we evaluated meteorological data from automated weather stations between the onset of glide cracking and avalanche release. Last, we synthesized lidar data with a numerical dynamics model to replicate the event in a simulated environment. Using the tuned model, we determined the critical release area snow depth necessary to reach the road (4.2 m). Our method may be of particular use for glide avalanches, which tend to release in roughly the same place and time each year at a known interface. This could make the calculated critical depths more consistently reliable and preclude the need for additional tuning in dynamics models. As 1) lidar technology continues to improve and reduce in cost, 2) transportation corridors continue to extend into avalanche terrain, and 3) glide avalanches potentially become increasingly frequent, the synthesis outlined here provides a valuable tool for operational forecasters on roadways threatened by glide events.

1 Introduction and background

As industrial, transportation, and recreational operations continue to extend into challenging mountainous terrain, highway avalanche control and forecasting operations are increasingly vital (Vera Valero et al., 2016). In contrast to ski area and backcountry forecasting, the critical question surrounding highway operations is simple: can an avalanche hit the road today? Road closures and equipment placement attempt to mitigate (sometimes severe) economic losses while protecting human life and infrastructure. During extreme weather conditions, such as heavy snow accumulation and subsequent storm slab formation, this decision is typically a straightforward one (Margreth et al., 2004). More difficult to predict are large wet snow avalanches, which can carry immense destructive potential, but frequently occur during non-extreme conditions. Given their unpredictability, they often create more challenges for forecasters than their better-understood dry snow counterparts (Borgeson et al., 2010; Stimberis et al., 2011). Glide avalanches, in particular, pose a significant and repetitive threat to many operational highway and railroad forecasting programs (Clarke and McClung, 1999; Reardon and Lundy, 2004; Simenhois and Birkeland, 2010; Stimberis and Rubin, 2004).

The term “glide” describes the progression of a snowpack slipping downhill along the ground-snow interface (Jones, 2004). It is understood that this happens when an accumulation of liquid water at the ground-snow interface, resulting from either water percolation through the snowpack or proximate melt, reduces friction and shear strength at the interface (Mitterer and Schweizer, 2012). When glide rates vary on a slope, a tensile fracture called a glide crack opens, and often a full-depth avalanche follows (hence, glide avalanche) (Clarke and McClung, 1999), thus these avalanches can be exceptionally large (Peitzsch et al., 2012). While the spatial location of glide release areas is extremely consistent, the timing of this critical friction reduction is notoriously difficult to forecast. Not all glide cracks produce avalanches (Reardon et al., 2006), and for those that do, the time to failure can range from hours to weeks or even months (Fees et al., 2023; McClung and Schaerer, 2006). Given their severity and unpredictability, recent research efforts have focused on linking glide avalanche timing to meteorological parameters (Peitzsch et al., 2012; Wever et al., 2018). To further complicate matters, glide avalanche release areas are difficult to mitigate with explosives (Clarke and McClung, 1999; Simenhois and Birkeland, 2010).

In addition to being difficult to forecast and control, glide avalanche occurrence is likely to vary in a changing climate. Currently glide avalanches are most common in maritime snow climates, although they have been observed elsewhere under warmer conditions (McClung and Schaerer, 2006). As global mean temperatures increase, there will likely be changes to the regional distribution of wet snow avalanches (Eckert et al, 2024; Mayer et al. 2024), including a potentially higher frequency of glide avalanches. Thus, the timing and dynamics of glide avalanches is an important area of study. To better understand these processes, and to improve assessments of risk to roadways, event documentation is key (Bründl & Margreth, 2021).

Recent advances in remote sensing technologies, including Terrestrial Laser Scanning (TLS), have provided the opportunity for high-spatiotemporal resolution snowpack surveys over several km^2, while sampling non-destructively and reducing practitioner exposure to hazardous terrain by acquiring data from a safe distance (e.g., Deems et al., 2015; Egli et al., 2012;

Maggioni et al., 2013; Prokop et al., 2015). Differencing temporally disparate TLS datasets can produce high-resolution maps of snow depth or changes therein. An excellent summary of snow depth mapping via TLS is provided by Deems et al. (2013). Maps of snow depth or change in depth derived from TLS can be used for a variety of purposes, such as observing release area and potentially erodible depth prior to avalanche, or quantifying the debris volume and extent of runout after failure.

Depth maps alone cannot be used to effectively predict runout distance, and thus, whether road impacts are imminent. For this purpose, the preferred tool is an avalanche dynamics model. Dynamics models solve the mass and momentum balance in a simulated avalanche using a terrain model and typically calculate flow depth, velocity, impact pressure, and runout distance based on a depth-averaged hydrodynamic system of equations (e.g., Christen et al., 2010; Vera Valero et al., 2016). Given that glide avalanches feature similar release areas that fail semi-annually under comparable snowpack conditions, they present themselves as excellent candidates for dynamics modeling. Detailed event documentation, such as with TLS surveying, can provide precise dynamics model inputs and is necessary for model calibration (Christen et al., 2010; Miller et al., 2023; Prokop et al., 2015), and thus the two technologies naturally complement one another.

Here, we document a large glide avalanche event along the Going-to-the-Sun Road (GTSR) in Glacier National Park (GNP), Montana, USA during road opening operations in the spring of 2022. We collected three sets of TLS data: pre-event, post-event, and snow-off. By differencing these elevation data we produced maps of the total height of snow (HS) and the difference in the height of snow (dHS), the latter of which we used to perform a mass/volume analysis of the avalanche. Further, we evaluated meteorological data from automated weather stations (AWSs) and one United States Department of Agriculture National Resource Conservation Service (NRCS) Snowpack Telemetry (SNOTEL) site between the onset of glide cracking and avalanche release. We performed a synthesis of TLS and the RApid Mass Movements (RAMMS) Extended avalanche dynamics model (Bartelt et al., 2016; Vera Valero et al., 2016) to replicate the event in a simulated environment. We used precise TLS-derived parameters that better capture real-world complexities relative to traditional estimates, such as average release area depth, as key inputs during model initialization, then leveraged the observed runout extent and volume for model calibration. Last, using the calibrated model, we determined the critical release area depth necessary for a road impact. A summary of this workflow is detailed by the flowchart in Fig. 1. As glide avalanches potentially become more frequent and transportation corridors extend further into alpine regions, the work presented here may help to better anticipate future hazards.

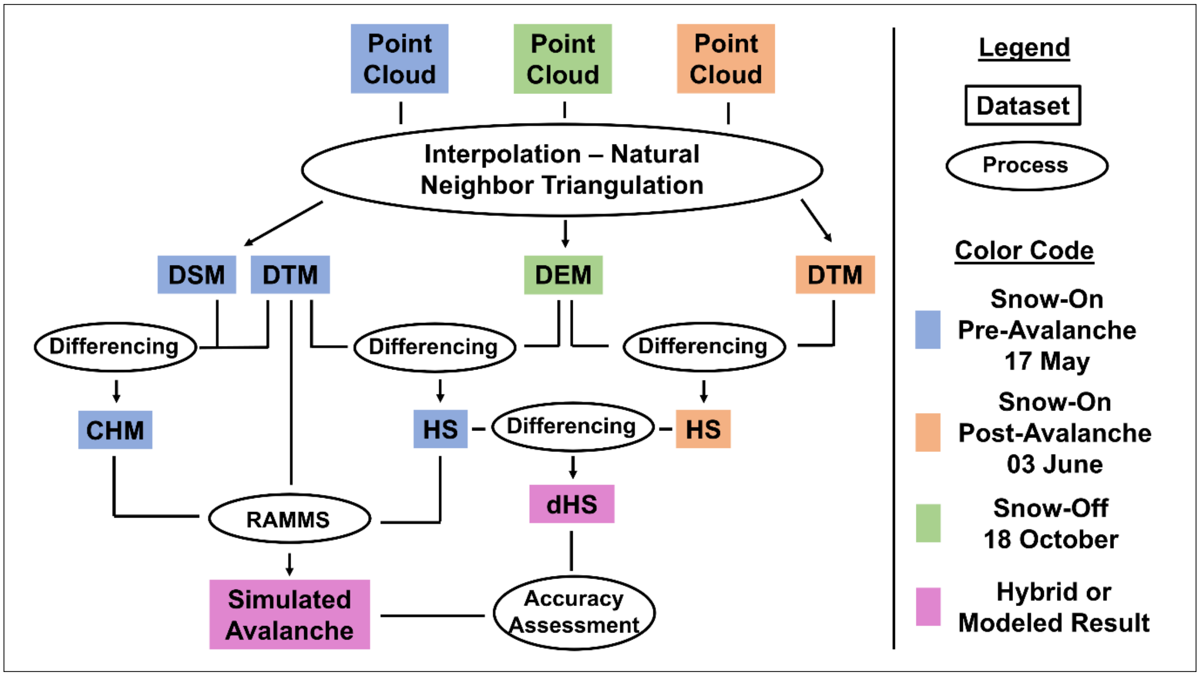

Fig. 1. Flowchart summarizing the methodology. Point clouds are interpolated into digital elevation models (DEMs, snow-off), digital terrain models (DTMs, snow-on) or digital surface models (DSMs, including vegetation). Differencing DTMs from DSMs yields canopy height models (CHMs), while differencing DTMs from DEMs produces maps of the total height of snow (HS). Further differencing these generates maps of the difference in snow height (dHS) across the glide avalanche event. Maps are used as both inputs and validation for the dynamics model RAMMS-Extended.

2 Methods

2.1 Study area

We conducted our survey in conjunction with the annual spring GTSR opening operations. Numerous avalanche paths threaten the GTSR, an 80 km roadway that serves as the primary traffic artery of GNP. Each spring, the 56 km section of the road that is closed during winter must be cleared. To ensure the safety of road workers, GNP and the U.S. Geological Survey (USGS) have partnered to provide an operational forecasting program for the annual spring opening. As a result, there is a reliable record of historical avalanche observations, and the area is well-instrumented.

The avalanche climate exhibits a generally maritime precipitation regime contrasted by continental temperature characteristics, a discrepancy resulting from proximity to the Continental Divide (Mock and Birkeland, 2000; Reardon et al., 2006). Slopes along the GTSR feature high-alpine terrain with a variety of avalanche problems. Considerable spatial variability in snow depth and other physical snowpack characteristics is common. Glide avalanches are prevalent in the spring, and Peitzsch et al. (2012) reported a remarkable 182 glide avalanches observed from the GTSR over an 8-year period (2003 – 2011). Many of these reoccurring glide avalanches have maximal runout extents that transect the GTSR, with the frequency of road impact depending on the path.

Our study Area-of-Interest (AOI) was a 0.78 km^2 region of the Haystack Creek drainage beneath Mount Gould, located in the Lewis Range of GNP at the center of the park (Fig. 2a). We selected this site to observe the Haystack glide avalanche, a pseudo-annual glide event that occurs in late spring. The area allowed for relatively safe access to TLS scan positions from the GTSR (Fig. 2b and 2d). Additionally, the high-elevation (2240 m) Garden Wall AWS (GWWX, Fig. 2b) is positioned less than 1 km from the release area.

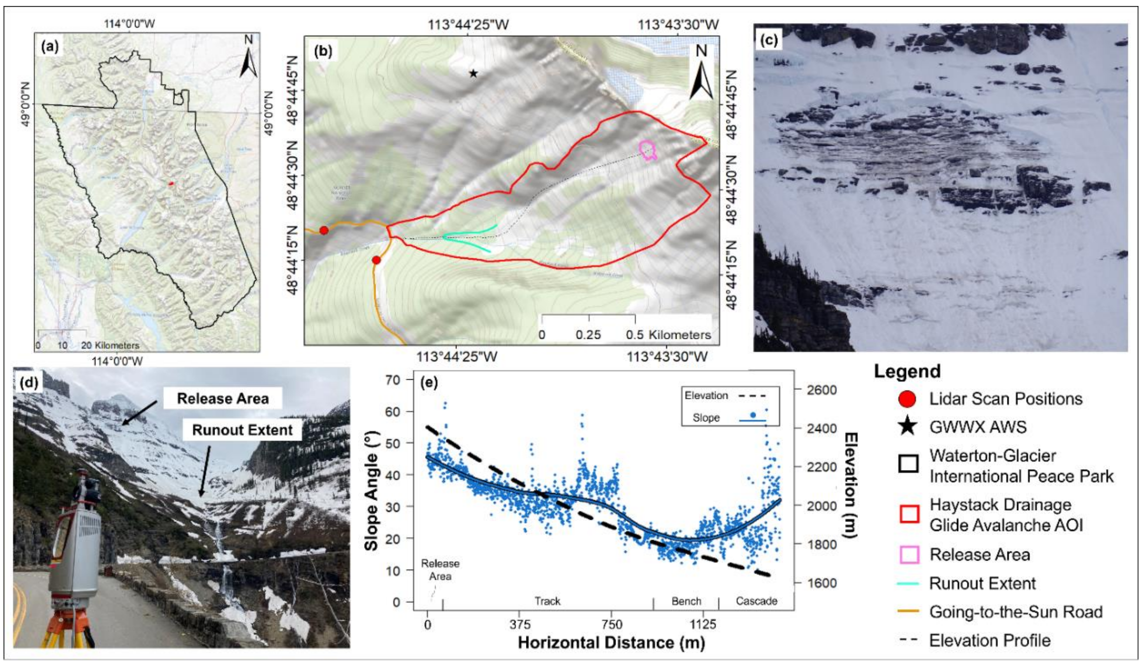

Fig. 2. General location of the study AOI within the Waterton-Glacier International Peace Park (a), topographic view including the extent of release area and runout from the 2022 event (b), photographs of the release area (courtesy of the USGS) (c) and the northwestern scanning position along the GTSR (d), and a slope profile along the avalanche path transect using TLS data (e).

From the release area (Fig. 2c), which sits at an elevation of 2430 m and an aspect of 270°, avalanche debris is funneled through one or more steep gullies to a bench roughly 400 m long, before spilling into a steep descent toward the GTSR along a cascading portion of Haystack Creek. Whether or not the glide avalanche impacts the GTSR largely depends on if the avalanche debris flow had sufficient momentum to clear the bench, where the slope is consistently less than 20°. This pattern of variable steepness is evident in the slope profile (Fig. 2e) along the transect pictured in Fig. 2b. Thus, the Haystack glide avalanche can be innocuous or a severe danger, depending on the size of the avalanche, which is difficult to predict. All available records of this glide avalanche occurrence, colloquially referred to as the “Show-Me” slide, are presented in Table 1. Observations of crown characteristics and runout distance are visual estimates from either adjacent ridges or the GTSR. In the spring of 2022, the slide stopped just before the outlet of the bench (Fig. 2b and 2d). The avalanche was classified as "GS – N – R4 – D3 – G" by professional observers (see Greene et al., 2022 for classification standards).

Table 1. Historical records of the "Show-Me" glide avalanche in the Haystack drainage.

ID (#)	Date	Road Hit?	R	D	Max Crown Depth (m)	Avg. Crown Depth (m)	Vertical Run (m)
111	5/4/2004	No	3.0	3.0	1.8	1.8	670
249	5/5/2005	No	1.5	1.5	1.5	1.5	580
233	5/21/2006	No	NA	2.0	2.0	2.0	595
406	5/17/2010	No	1.0	2.0	2.5	2.5	395
593	5/28/2013	No	2.0	2.0	2.0	2.0	730
678	5/22/2015	No	3.0	2.5	NA	NA	580
878	5/17/2020	Yes	NA	3.0	4.0	4.0	820
1063	6/1/2022	No	4.0	3.0	4.0	3.5	700

2.2 Meteorological data

We evaluated meteorological data from two AWSs and one SNOTEL site. GWWX sits at 2240 m and is less than 1 km from the Haystack AOI. The station is located on a southwest-facing slope just west of the Garden Wall, a rocky ridge that forms the Continental Divide. The other AWS is located at the Logan Pass Visitor Center at an elevation of 2035 m in broad, low-angle terrain around tree line, and is about 5 km from the AOI. Hourly temperature averages are reported at each station. The nearest SNOTEL site is located on Flattop Mountain (1810 m), a broad plateau roughly 12 km from the AOI. We extracted daily HS values from this site to examine trends in melt and settlement. Further, we used these data to leverage the classification tree put forth by Peitzsch et al. (2012). The model uses meteorological variables to predict which days will be "avalanche days", defined as a day where at least one wet slab or glide avalanche occurs on the GTSR, and which will not. We analyzed a time series from 17 May to 3 June, a range that brackets our snow-on TLS data acquisition and the glide event, discussed further in Sect. 2.3.

2.3 Terrestrial laser scanning

2.3.1 Scanning and registration

We acquired TLS data using a tripod-mounted Riegl VZ-6000 long-range terrestrial scanner (Fig. 2d), operating at 1064 nm, during each of three outings surrounding the glide event of spring 2022. For each outing, we obtained data from two scan positions along the GTSR (Fig. 2b), which reduced terrain shadowing, once aligned. We collected two "snow-on" datasets: pre-event, just after the onset of glide crack emergence (17 May) and post-event (03 June). The slide released on the evening of 01 June between 20:00 and 22:00, with no precipitation occurring between then and our data acquisition the morning of 03 June. We also obtained a snow-off dataset (18 October) later in the year. The snow conditions during both snow-on scan outings were typical of a late spring snowpack, with a prevalence of wet, rounded snow grains and melt formations.

Regarding scan parameters, we set the pulse repetition frequency to 30 kHz to maximize range and selected a 0.02° vertical/horizontal step increment to optimally straddle spatial resolution (point density) and collection time, with scan durations of about 20 – 30 minutes. We mounted small circular reflectors to trees throughout the scan field-of-view for initial scan alignment, then further refined registration with manual coarse alignment, and finally with Riegl's semi-automatic, iterative Multi-Station Adjustment (MSA) workflow, careful to avoid alignment based on snow surfaces between dates (Deems et al., 2015; Riegl, 2015a). Alignment (both coarse and MSA) across scanning dates centered on exposed cliff faces and large vegetation, as well as permanent GTSR infrastructure.

For depth mapping via differencing, relative alignment between scans is the most important matter. However, we also underwent global registration to incorporate external datasets into our data visualization. To accomplish this, we leveraged in-situ collected GPS points, acquired via Emlid Reach RS2 systems, corresponding to each tie point and scan position. Once relative and global registration was complete, we exported point clouds from Riegl's proprietary RiScan software into R and ESRI's ArcMap for further processing. Resulting clouds had point spacing of ~11 – 14 pts/m^2, depending on the date.

2.3.2 Mapping and analysis

We converted all registered point clouds (two from each collection date: including and excluding vegetation points) to raster format using a natural neighbor triangulation. Digital surface models (DSMs) are those which include vegetation and other overground points. Digital terrain models (DTMs) exclude vegetation and represent the elevation of the snow surface. Digital elevation models (DEMs) also exclude vegetation but are acquired in snow-off conditions (18 October, in this case), and thus represent the true elevation of the ground surface. Given our point spacing we used a 1 m raster grid size that minimized smoothing and preserved the integrity of terrain variability, while still permitting reasonable processing times. Additionally, we generated a 3D triangulated irregular network (TIN) on which to drape maps or imagery for 3D visualizations (Fig. 3). We differenced DTMs from the DEM, and from one another, to create three corresponding maps: total snow depth on 17 May (pre-event), total snow depth 03 June (post-event), and change in snow depth between snow-on scans (03 June depth – 17 May depth) to visualize the effects of the glide avalanche. Further, we extracted volumes of released snow, entrained snow, and deposition, and performed a subsequent mass balance calculation to estimate the released and entrained snow densities. Finally, we differenced the DTM from the DSM on 17 May to create a canopy height model (CHM) for use in RAMMS-Extended (to be discussed in Sect. 2.4).

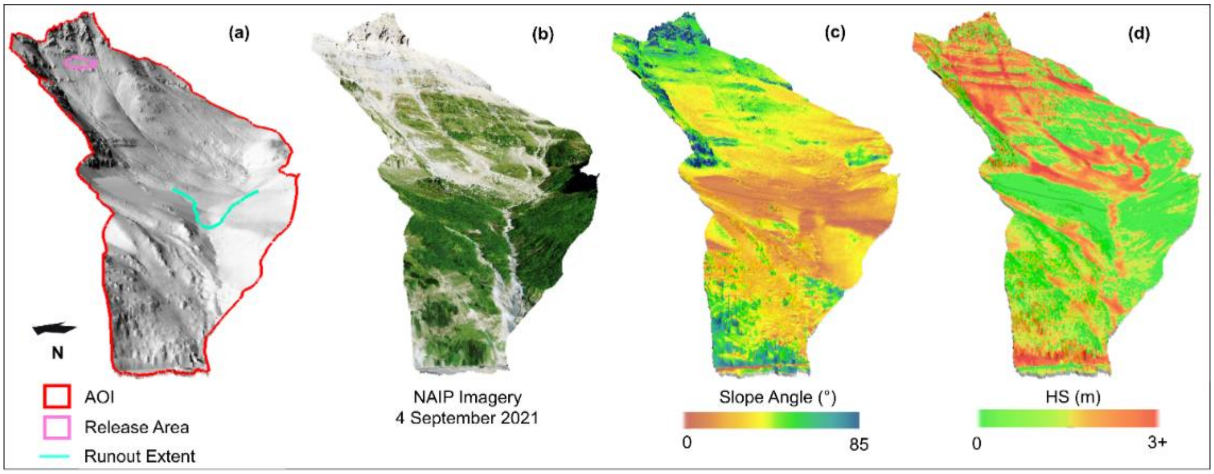

Fig. 3. A hillshade model with the same symbology as Fig. 2 (a), National Agriculture Imagery Project (NAIP) imagery of snow-off conditions (b), a slope angle map created by differentiating the DTM with respect to elevation (c), and a HS map (d). Excluding the NAIP imagery, all images were created from the 17 May TLS data. The extent of the central bench is evident in (c), as is extreme spatial variability of snow depth in (d).

2.4 Dynamics modeling

To better understand the dynamics of the Haystack glide event, we simulated the avalanche in the RAMMS model, originally created by Christen et al. (2010a). More recently, the model has been “extended” by Bartelt et al. (2016) and Vera Valero et al. (2016) to better consider factors including snow temperature, liquid water content (LWC), phase changes, entrainment, cohesion, and fluidization, in part towards improved modeling of wet snow avalanches. In a typical dynamics application, a user will 1) input a digital terrain model (DTM) of the underlying ground surface, 2) assign an avalanche release area extent based on local expertise or a GIS-based terrain analysis (e.g., Bühler et al., 2013; Maggioni and Gruber, 2003), 3) provide an estimate of release area snow depth (to the assumed failure plane) and density using in situ and/or automated weather station observations, and 4) delineate vegetated areas if necessary, typically from photographs or a separate database.

Since its inception, it has been well-documented that fine-tuning of input parameters is imperative to obtaining an accurate avalanche simulation in RAMMS (Christen et al., 2010b). For example, Bühler et al. (2011) demonstrated significantly different outputs from RAMMS by simply varying the spatial resolution of the input DTM. In addition to the quality and spatial resolution of the input DTM, the temporal accuracy (i.e., snow-on vs. snow-off) is likely a relevant factor, as Veitinger and Sovilla (2016) found that snow-covered winter terrain can significantly differ from its underlying, snow-free terrain due to variable snow accumulation. Miller et al. (2022) further highlighted variability related to the input elevation model and resolution. Stoffel et al. (2018) observed that RAMMS-Extended results, particularly whether slides present a critical scenario for roads, depends heavily on release depth, entrainment, and snow temperature, and discussed the considerable uncertainty associated with extrapolating release area snow depth from study plot measurements. Similarly, Glaus et al. (2023)

conducted a sensitivity analysis in RAMMS-Extended, and found that release depth and temperature exerted the strongest influence on avalanche runout independent of terrain.

We hypothesize that using precise TLS inputs may reduce uncertainty and allow for accurate results with minimal tuning (Deems et al., 2015; Prokop et al., 2015). Therefore, we began with the most recent (17 May) snow-on DTM for the sliding surface and corresponding CHM for vegetation. We delineated the release area extent from TLS scans and, for perhaps the most important parameter, the release area HS, we used the TLS-derived average depth from 17 May (3.29 m, discussed in Sect. 3.2). Further, we assigned the potentially erodible depth to the average observed loss in the entrainment zone based on dHS maps, following RAMMS team recommendations. Using the results of our mass balance analysis (Sect. 3.2), we estimated both release area and entrained snow densities.

Regarding snowpack temperature and initial liquid water content (LWC_0), we used forecaster recommendations: isothermal, 0° C and 10 mm m^{-2}, respectively, values typical of a late spring snowpack. Following Miller et al. (2022), who also modeled a wet snow avalanche, we assigned the initial friction parameters: coulomb friction, $\mu_0 = 0.55$; turbulent friction, $\xi_0 =$ 1500 m s^{-2}; cohesion, $N_0 = 200$ Pa. We assume that after LWC exceeds 500 mm m^{-2}, then the friction μ reduces to 0.12 (Miller et al., 2023; Vera Valero et al., 2016, 2018). For the generation of turbulent core energy parameter (α), which describes fluidization as a function of terrain (Vera Valero et al., 2016), we selected a value of 0.05 used elsewhere (e.g., Bartelt et al., 2015; Vera Valero et al., 2018) and typical of wet snow in smoother terrain. We ran simulations with a 5 m grid cell size, as this is recommended in the RAMMS manual, and previous studies have found the optimal grid size to be 2 – 5 m (Miller et al., 2022; Veitinger et al, 2014; Vera Valero et al., 2018). We used the default stopping criteria of 5% moving momentum.

3 Results and discussion

3.1 Meteorological analysis and event timing

The meteorological data from GWWX AWS and the Flattop Mountain SNOTEL site (FT) are presented in Fig. 4. As mentioned in Sect. 2.3.1, we first collected TLS data on 17 May, shortly after the appearance of a glide crack in the release area (left orange vertical reference line in Fig. 4). On 18 May, FT recorded 31 cm in snow accumulation, followed by a period of settlement and colder temperatures prevailing at GWWX. From 22 – 25 May, temperatures warmed with maximum daytime temperatures at GWWX of 4° C, but overnight temperatures consistently dipped below freezing, resulting in a melt-freeze cycle. A brief but substantial warmup ensued from 26 May until the early hours of 28 May, with temperatures remaining above 0° C the evening of 27 May and a maximum temperature of 7.7° C. These conditions likely caused substantial destabilization, but not enough to initiate the Haystack glide event (Fees et al., 2023).

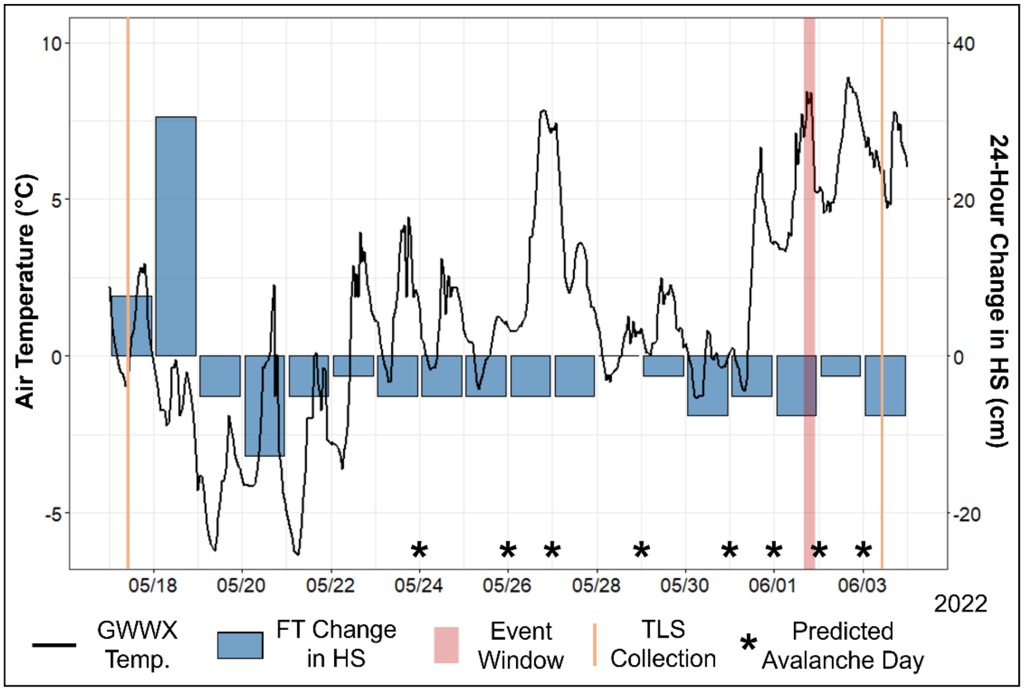

Fig. 4. Hourly temperature averages from GWWX and daily change in snow height at FT. Orange vertical reference lines denote TLS data collection and the red shaded region mark the avalanche event. Asterisks mark predicted avalanche days where wet slab or glide avalanches may be expected based on the classification tree in Peitzsch et al. (2012).

After 27 May, temperatures returned to melt-freeze conditions for several days, with overnight temperatures at or below 0° C, before a final warmup catalyzed the glide release. Beginning 31 May, temperatures increased rapidly, with a daytime maximum of 6.7° C and an overnight low of 3.3° C, the highest overnight temperatures since the glide cracks emerged. The warming trend continued on 01 June, when GWWX temperatures reached 8.4° C late in the day, the highest recorded temperatures since the appearance of the glide crack, and the avalanche released. Asterisks at the bottom of Fig. 4 denote “avalanche days” following the classification tree described in Peitzsch et al. (2012), based on temperature data from the Logan Pass AWS (not pictured but highly correlated with GWWX) and FT settlement data. Indeed, 01 June was considered an avalanche day. Note that the classification tree attempts to predict days when a wet slab or glide avalanche occurs on at least one of the many paths along GTSR, and thus the other avalanche days besides 01 June are not necessarily false positives; the model is not specific to the Haystack path.

3.2 Depth mapping and mass balance

Prior to release, the glide crack could be observed in the 17 May point cloud, with an average width of 1.25 m and a maximum width of 2.25 m. We found extreme spatial variability in HS, even across small spatial scales (e.g., Fig. 3d). On both dates, much of the AOI had snow depths < 1 m, while the deepest areas were > 7 m. Most snow was packed into start zones between cliff bands and drainage funnels (Fig. 3d). The maximum depth of snow in the release area on 17 May was 7.23 m, with an average of 3.29 m (σ = 1.56 m) and a total release volume (V_r) of 25,462 m^3. After the event, we found the average crown height after release (03 June) to be 3.0 m with maximum values of 4.0 m, in good agreement with the visual observations in Table 1. The average release area width was 95 m.

We observe chronological stages of the slide in Fig. 5a, which illustrates the change in snow depth (dHS) before versus after the avalanche. In addition to the extreme loss of snow cover in the release area, a substantial entrainment zone existed (resulting in a reduction of snow depth, red) along the path before transitioning to a deposition area at the bottom of the gully and into the bench, represented as an increase in snow depth (blue). Outside of the avalanche path, we calculated a small reduction in snow depth (7 cm on average) due to both the 18 May storm accumulation and consistent melt and settlement thereafter (Fig. 4). Fig. 5b illustrates dHS along A – A`, which transects the release area and transitions into the onset of entrainment. We can see that initially, in the exposed cliff bands above the start zone, dHS along the transect is essentially equal to zero, before plummeting to large negative values in the midst of the release area. After, the dHS plot reconverges at the lower cliff band that acts as the stauchwall, before entering an entrainment zone which oscillates in magnitude between rock features. Similarly, Fig. 5c plots dHS along transect B – B`, showing the transition from entrainment to deposition. Values are initially negative as entrainment concludes, before rapidly shifting to positive values with irregularity highlighting spatial variability in the debris area.

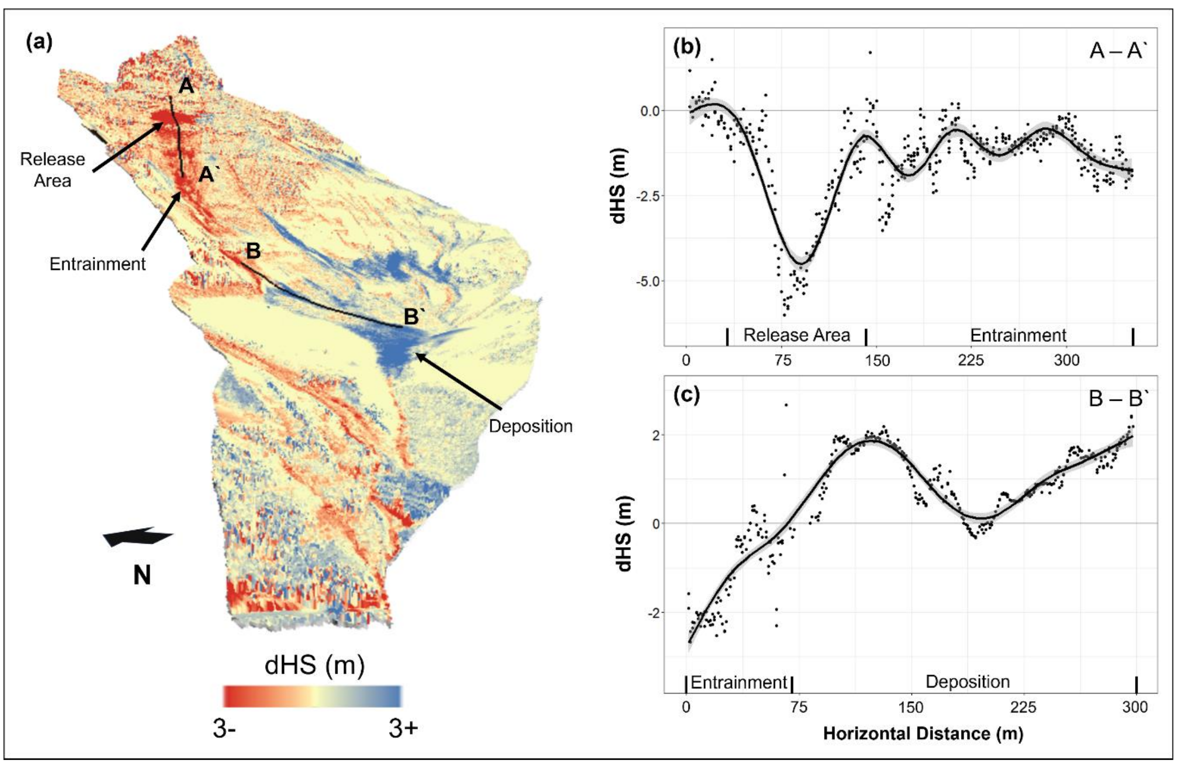

Fig. 5. A map of change in snow height (dHS) before versus after the avalanche event (a) with plots along transects A – A` (b) and B – B` (c). In (b) and (c), points represents raw data values, with a smoothed best-fit line, and grey shading that outlines 95% confidence intervals.

We calculated an entrained snow volume (V_e) of 65,870 m^3 with an average erosion depth of 0.69 m (σ = 0.58 m) and a maximum of 3.31 m, highlighting the substantial contribution of entrained snow to overall avalanche volume. The ratio of entrained snow volume to release area snow volume was 2.59. In the deposition area, the volume of deposited snow (V_d) was 62,501 m^3 with an average height of 0.75 m (σ = 0.67 m) and a maximum height of 4.66 m. Because density measurements from the Haystack AOI were unavailable in spring 2022, we attempted to back-calculate density via a mass balance analysis, with a few assumptions. First, we assumed that the deposition snow density, ρ_d, was 450 kg m^{-3}, as this is the recommended value for RAMMS-Extended. Next, we assume that the entrained snow density (ρ_e), hence snow in the top 1 m of the snowpack, is 50 kg m^{-3} less than the release area snow density (ρ_r). The latter is equivalent to bulk snow density, given that the avalanche released at the ground-snow interface. The mass balance calculation is then summarized by the following equations:

$$V_r\rho_r + V_e\rho_e = V_d\rho_d \tag{1}$$

$$\rho_r = \rho_e + 50 \tag{2}$$

We calculated the resulting densities ρ_e = 291 kg m^{-3} and therefore ρ_r = 341 kg m^{-3}. These estimates are not uncommon for late spring snowpacks and were in good agreement with snowpits elsewhere in GNP close to the time of the Haystack glide event.

Last, we used these density values to determine the masses of snow released, entrained, and deposited (M_r, M_e, and M_d, respectively): M_r = 8,682,542 kg, M_e = 19,168,170 kg, and M_d = 28,125,450 kg. A deep, dense snowpack and substantial entrainment contributed to these large masses despite a relatively small release area,. Using these masses, we calculated the growth index, I_g, introduced by Sovilla (2004) using Eq. 3:

$$I_g = \frac{M_d}{M_r} \tag{3}$$

The resulting growth index of the Haystack avalanche was I_g = 3.2. This result fits nicely into the range of I_g values observed by Sovilla (2004): 1.8 – 8.8. Our results reinforce the conclusions of Sovilla (2004) and Prokop et al. (2015), who noted that entrainment contributes significantly to avalanche mass and volume. Our results indicate that entrainment also contributes substantially to overall mass in a dense, wet snow regime.

3.3 RAMMS-Extended

We sought to accurately reproduce the Haystack glide event in RAMMS-Extended with minimal tuning, leveraging a combination of TLS-derived inputs, wet snow friction values found in the literature, and local forecaster observations/expertise. Using the parameter set described in Sect. 2.4, the simulated avalanche overran the observed runout extent and reached the GTSR. We noted that this incorrect result was related to over-erosion; the simulated avalanche entrained 139,346 m^3 of snow, more than doubling the observed volume of 65,870 m^3. RAMMS-Extended no longer has an “erodibility” parameter. However, an indirect means of reducing the erosion rate is to increase the avalanche yield stress. By tuning the yield stress from its default value of 300 Pa to 2000 Pa, we obtained a simulated avalanche that was in good agreement with both the observed runout extent and entrained volume (Fig. 6a). The simulated entrainment volume was 66,420 m^3, and the avalanche overran the observed runout extent by just 6 m. The maximum deposition height was also in good agreement (simulated 4.30 m compared to observed 4.66 m), and simulated average deposition height (0.75 m, σ = 0.93 m) was identical to the measured value. The simulated deposition is slightly wider than the observed extent, but the pattern is otherwise similar. Input parameters for the optimal simulation are listed in Table 2.

Table 2. Input parameters from the optimal RAMMS-Extended simulation. The top set of parameters were produced or derived from TLS data, the middle set based on input from GTSR forecasters, and the bottom set are based on literature and RAMMS team recommendations.

Surface	17 May DTM
Release HS (m)	3.3
Forest	17 May CHM
Erosion HS (m)	0.7
Release Density, ρ_r (kg m^{-3})	341
Entrained Density, ρ_e (kg m^{-3})	291
Initial Liquid Water Content, LWC_0 (mm m^{-2})	10
Snow Temperature (°C)	0
Initial Coulomb Friction, μ_0	0.55
Transitional Coulomb Friction, μ_{wet}	0.12
Initial Turbulent Friction, ξ_0 (m s^{-2})	1500
Initial Cohesion, N_0 (Pa)	200
Turbulent Core Energy, α	0.05
Yield Stress (Pa)	2000

Towards a forecasting application, we experimented with incrementally increasing the release area HS while keeping all other parameters constant to determine the critical depth necessary for a road impact. A release area HS of 4.2 m resulted in a simulated avalanche with sufficient momentum to clear the bench and impact the road (Fig. 6b). This result aligns with the estimated in-situ crown height from the 2020 glide event (4.0 m), when the avalanche impacted the road (Table 1). While encouraging, this is not thorough validation since in-situ crown height estimates were made visually from nearby ridges. Future surveys of this glide path could provide more thorough validation of the simulated critical slab depth. Specifically, collecting high-accuracy measurements of release area HS (from TLS or other technologies) during a year when a glide avalanche reaches the road would be very useful.

We believe that our proposed workflow of surveying, calibrating a dynamics model, and then extrapolating to estimate a critical depth, may be of particular use for forecasters dealing with: 1) path-by-path hazard assessments, especially above roads, 2) repeat-offender avalanches, and/or 3) glide or wet slab events. All avalanche events have specific snow conditions that require individual event model recalibration and make extrapolation of our specific results inappropriate. However, the consistency of glide avalanches presents a unique case, given that glide avalanches tend to have very consistent release area boundaries, occur at a known interface (ground-snow), and, at this location, release at generally the same time each year. Furthermore, even in regions where glide release areas are not well-understood, promising modeling efforts provide optimism for identifying the spatial extent of worst-case start zones (Fees et al., 2024); one could even envision an eventual coupling of modeled release area extent, TLS-measured snow depth, and dynamics modeling. This consistency could make the calculated critical depths more consistently reliable and preclude the need for additional tuning of dynamics models, and thus forecasters could assess hazard risk based purely on release area HS as it compares to the pre-simulated critical depth. But again, this presumption requires additional study and validation.

The timing of glide avalanche release is difficult to predict. Some progress has been made using time-lapse photography (Fees et al., 2023), physics-based snowpack models (e.g., Wever et al., 2018) and statistical models leveraging meteorological parameters (e.g., Peitzsch et al., 2012), but predicting slope-specific event timing remains very challenging. Our results indicate that if a forecaster had reasonable confidence that an avalanche is incapable of reaching the road, then accurately predicting timing becomes less imperative. Conversely, if there is a high probability that a glide event will result in road impact, then days when meteorological or other parameters raise the probability of occurrence would elicit heightened concern.

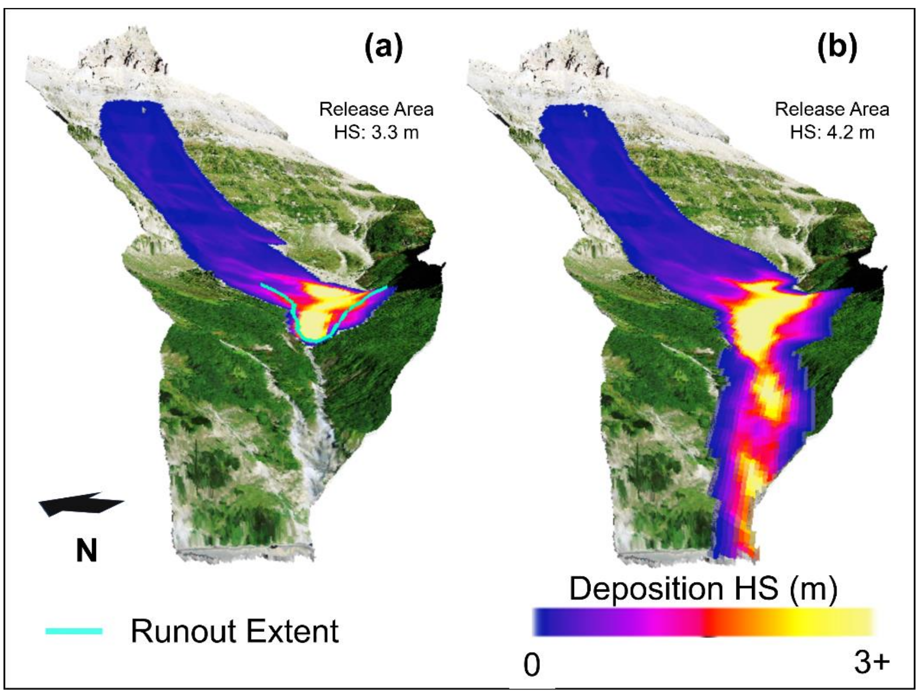

Fig. 6. Simulated deposition snow heights from RAMMS-Extended using the actual release area HS of 3.3 m (a) compared to the critical HS of 4.2 m (b).

4 Conclusions

We thoroughly evaluated a large, road-threatening glide avalanche in Glacier National Park, Montana, using terrestrial lidar and demonstrated a novel synthesis of TLS data with avalanche dynamics modeling for both initialization and validation. We believe this marks the first study to leverage a TLS spatiotemporal analysis toward surveying a glide avalanche. We show the efficiency of TLS in creating powerful visualizations for avalanche forecasting and decision-making. Further, we quantified a variety of parameters from TLS-derived maps, including glide crack width, crown height, runout extent, and the volume released, entrained, and deposited by the avalanche. Perhaps most importantly, using TLS-derived parameters, DTMs and CHMs, we were able to simulate the avalanche in RAMMS-Extended with minimal tuning. Using this tuned model, we calculated the critical depth necessary for road impact (4.2 m). It is likely that our TLS-derived inputs enhance dynamics model results relative to traditional inputs by improving key input parameter accuracy and better representing real-world complexities. The method presented here may be of particular use to avalanche operations that contend with glide avalanches, which often release in generally the same location each year at a known interface (ground-snow). This would make the calculated critical depths more consistently reliable and preclude the need for additional tuning in dynamics models. As 1) TLS technology continues to improve and reduce in cost, 2) transportation corridors continue to extend into avalanche terrain, and 3) glide avalanches potentially become increasingly frequent, the synthesis outlined here provides a valuable tool for operational forecasters on roadways threatened by glide events.

5 References

Bartelt, P., Valero, C. V., Feistl, T., Christen, M., Bühler, Y., & Buser, O. (2015). Modelling cohesion in snow avalanche flow. *Journal of Glaciology*, *61*(229), 837-850.

Bartelt, P., Buser, O., Valero, C. V., & Bühler, Y. (2016). Configurational energy and the formation of mixed flowing/powder snow and ice avalanches. *Annals of Glaciology*, *57*(71), 179-188.

Borgeson, L. E., & Hartman, H. (2010). Predicting wet snow avalanches at the Arapahoe basin ski area. In *Proceedings: ISSW*.

Bründl, M., & Margreth, S. (2021). Integrative risk management: the example of snow avalanches. In *Snow and ice-related hazards, risks, and disasters* (pp. 259-296). Elsevier.

Bühler, Y., Christen, M., Kowalski, J., & Bartelt, P. (2011). Sensitivity of snow avalanche simulations to digital elevation model quality and resolution. *Annals of Glaciology*, *52*(58), 72-80.

Bühler, Y., Kumar, S., Veitinger, J., Christen, M., Stoffel, A., & Snehmani. (2013). Automated identification of potential snow avalanche release areas based on digital elevation models. *Natural Hazards and Earth System Sciences*, *13*(5), 1321-1335.

Christen, M., Kowalski, J., & Bartelt, P. (2010). RAMMS: Numerical simulation of dense snow avalanches in three-dimensional terrain. *Cold Regions Science and Technology*, *63*(1-2), 1-14.

Christen, M., Bartelt, P., & Kowalski, J. (2010). Back calculation of the In den Arelen avalanche with RAMMS: interpretation of model results. *Annals of Glaciology*, *51*(54), 161-168.

Clarke, J., & McClung, D. (1999). Full-depth avalanche occurrences caused by snow gliding, Coquihalla, British Columbia, Canada. *Journal of Glaciology*, *45*(151), 539-546.

Deems, J. S., Painter, T. H., & Finnegan, D. C. (2013). Lidar measurement of snow depth: a review. *Journal of Glaciology*, *59*(215), 467-479.

Deems, J. S., Gadomski, P. J., Vellone, D., Evanczyk, R., LeWinter, A. L., Birkeland, K. W., & Finnegan, D. C. (2015). Mapping starting zone snow depth with a ground-based lidar to assist avalanche control and forecasting. *Cold Regions Science and Technology*, *120*, 197-204.

Eckert, N., Corona, C., Giacona, F., Gaume, J., Mayer, S., van Herwijnen, A., ... & Stoffel, M. (2024). Climate change impacts on snow avalanche activity and related risks. *Nature Reviews Earth & Environment*, 1-21.

Egli, L., Jonas, T., Grünewald, T., Schirmer, M., & Burlando, P. (2012). Dynamics of snow ablation in a small Alpine catchment observed by repeated terrestrial laser scans. *Hydrological Processes*, *26*(10), 1574-1585.

Fees, A., van Herwijnen, A., Altenbach, M., Lombardo, M., & Schweizer, J. (2023). Glide-snow avalanche characteristics at different timescales extracted from time-lapse photography. *Annals of Glaciology*, 1-12.

Fees, A., van Herwijnen, A., Lombardo, M., Schweizer, J., & Lehmann, P. (2024). Glide-snow avalanches: A mechanical, threshold-based release area model. *Natural Hazards and Earth System Sciences Discussions*, *2024*, 1-19.

Glaus, J., Jones, K. W., Bühler, Y., Christen, M., Ruttner-Jansen, P., Gaume, J., & Bartelt, P. (2023). RAMMS:: Extended-sensitivity analysis of numerical fluidized powder avalanche simulation in three-dimensional terrain. In *International Snow Science Workshop Proceedings*.

Jones, A. (2004). Review of glide processes and glide avalanche release. *Avalanche News*, *69*(7), 53-60.

Maggioni, M., & Gruber, U. (2003). The influence of topographic parameters on avalanche release dimension and frequency. *Cold Regions Science and Technology*, *37*(3), 407-419.

Maggioni, M., Freppaz, M., Ceaglio, E., Godone, D., Viglietti, D., Zanini, E., ... & Pallara, O. (2013). A new experimental snow avalanche test site at Seehore peak in Aosta Valley (NW Italian Alps)—part I: conception and logistics. *Cold Regions Science and Technology*, *85*, 175-182.

Margreth, S., & Ammann, W. J. (2004). Hazard scenarios for avalanche actions on bridges. *Annals of Glaciology*, *38*, 89-96.

Mayer, S., Hendrick, M., Michel, A., Richter, B., Schweizer, J., Wernli, H., & van Herwijnen, A. (2024). Changes in snow avalanche activity in response to climate warming in the Swiss Alps. *EGUsphere*, *2024*, 1-32.

McClung, D., & Schaerer, P. A. (2006). *The avalanche handbook*. The Mountaineers Books.

Miller, A., Sirguey, P., Morris, S., Bartelt, P., Cullen, N., Redpath, T., ... & Bühler, Y. (2022). The impact of terrain model source and resolution on snow avalanche modeling. *Natural Hazards and Earth System Sciences*, *22*(8), 2673-2701.

Miller, A. D., Redpath, T. A., Sirguey, P., Cox, S. C., Bartelt, P., Bogie, D., ... & Bühler, Y. (2023). Unprecedented winter rainfall initiates large snow avalanche and mass movement cycle in New Zealand's Southern Alps/Kā Tiritiri o te Moana. *Geophysical Research Letters*, *50*(8), e2022GL102105.

Mitterer, C., & Schweizer, J. (2012). Glide snow avalanches revisited. *The avalanche journal*, *101*, p68-71.

Mock, C. J., & Birkeland, K. W. (2000). Snow avalanche climatology of the western United States mountain ranges. *Bulletin of the American Meteorological Society*, *81*(10), 2367-2392.

Peitzsch, E. H., Hendrikx, J., & Fagre, D. B. (2012). Timing of wet snow avalanche activity: an analysis from Glacier National Park, Montana, USA. In *Proceedings of the International Snow Science Workshop, September 17–21, 2012, Anchorage, Alaska*.

Prokop, A., Schön, P., Singer, F., Pulfer, G., Naaim, M., Thibert, E., & Soruco, A. (2015). Merging terrestrial laser scanning technology with photogrammetric and total station data for the determination of avalanche modeling parameters. *Cold Regions Science and Technology*, *110*, 223-230.

Reardon, B., & Lundy, C. (2004, September). Forecasting for natural avalanches during spring opening of the Going-to-the-Sun Road, Glacier National Park, USA. In *Proceedings of the 2004 International Snow Science Workshop* (pp. 566-581). Jackson Wyoming, USA.

Reardon, B. A., Fagre, D. B., Dundas, M., & Lundy, C. (2006, October). Natural glide slab avalanches, Glacier National Park, USA: a unique hazard and forecasting challenge. In *Proceedings of the 2006 International Snow Science Workshop, Telluride, CO, USA* (pp. 778-785).

Riegl Laser Measurement Systems, GmbH, 2015a. RiVLib. http://www.riegl.com/index.php?id=224

Simenhois, R., & Birkeland, K. (2010, October). Meteorological and environmental observations from three glide avalanche cycles and the resulting hazard management technique. In *International Snow Science Workshop ISSW, Lake Tahoe CA, USA* (pp. 846-853).

Sovilla, B. (2004). *Field experiments and numerical modelling of mass entrainment and deposition processes in snow avalanches* (Doctoral dissertation, ETH Zurich).

Stimberis, J., & Rubin, C. (2004, September). Glide avalanche detection on a smooth rock slope, Snoqualmie Pass, Washington. In *Proceedings ISSW* (pp. 608-610).

Stimberis, J., & Rubin, C. M. (2011). Glide avalanche response to an extreme rain-on-snow event, Snoqualmie Pass, Washington, USA. *Journal of Glaciology*, *57*(203), 468-474.

Stoffel, L., Bartelt, P., Margreth, S., & Schweizer, J. (2018, October). Can scenario-based avalanche dynamics calculations help in the decision making process for road closures? In *Proceedings of the International Snow Science Workshop* (pp. 07-12).

Veitinger, J., Sovilla, B., & Purves, R. S. (2014). Influence of snow depth distribution on surface roughness in alpine terrain: a multi-scale approach. *The Cryosphere*, *8*(2), 547-569.

Veitinger, J., & Sovilla, B. (2016). Linking snow depth to avalanche release area size: measurements from the Vallée de la Sionne field site. *Natural Hazards and Earth System Sciences*, *16*(8), 1953-1965.

Vera Valero, C., Wever, N., Bühler, Y., Stoffel, L., Margreth, S., & Bartelt, P. (2016). Modelling wet snow avalanche runout to assess road safety at a high-altitude mine in the central Andes. *Natural Hazards and Earth System Sciences*, *16*(11), 2303-2323.

Vera Valero, C., Wever, N., Christen, M., & Bartelt, P. (2018). Modeling the influence of snow cover temperature and water content on wet-snow avalanche runout. *Natural Hazards and Earth System Sciences*, *18*(3), 869-887.

Wever, N., Vera Valero, C., & Techel, F. (2018). Coupled snow cover and avalanche dynamics simulations to evaluate wet snow avalanche activity. *Journal of Geophysical Research: Earth Surface*, *123*(8), 1772-1796.

CHAPTER SIX

DETERMINING THE FLOW STATE OF CHANNELS UNDER VEGETATION WITH AIRBORNE LIDAR

Contribution of Authors and Co-Authors

Manuscript in Chapter 6

Author: James Dillon

Contributions: Conceptualized the study, acquired data, analyzed results, and drafted the manuscript.

Co-Author: Rick Lawrence

Contributions: Provided guidance and advised during conceptualization and analysis.

Co-Author: Kevin Hammonds

Contributions: Acquired funding, was responsible for project administration, and supervised throughout.

Manuscript Information

James Dillon, Rick Lawrence, Kevin Hammonds

Water Resources Research

Status of Manuscript:
☐ Prepared for submission to a peer-reviewed journal
☐ Officially submitted to a peer-reviewed journal
☐ Accepted by a peer-reviewed journal
X Published in a peer-reviewed journal

Publisher: Wiley
Submitted: 23 June 2022
Published: 6 June 2023
Volume: 59
DOI: https://doi.org/10.1029/2022WR033071

Determining the Flow State of Channels under Vegetation with Airborne Lidar

J. W. Dillon[1], R. L. Lawrence[2], and K. D. Hammonds[1]

[1]Department of Civil Engineering, Montana State University, Bozeman, MT 59715.

[2] Department of Land Resources and Environmental Sciences, Montana State University, Bozeman, MT 59715.

Corresponding authors: James Dillon (james.dillon013@gmail.com) & Kevin Hammonds (kevin.hammonds@montana.edu)

Key Points:

- A reduction in the lidar return intensity of wet stream reaches under dense vegetation relative to dry channel reaches is statistically validated.
- Optimal intensity thresholds for delineating wet from dry channel reaches obscured by vegetation are derived from probability density functions.
- A classification workflow is presented to assign and map the flow state of obscured channel networks based on this novel phenomenon, with a median accuracy of 93.0%.

Abstract

Headwater channels are vital to ecological health, water quality, and watershed connectivity. However, the geographic extent and the temporal wetting and drying dynamics of headwater catchments are not thoroughly understood, primarily due to occlusive vegetation. Light Detection and Ranging (lidar) data can be utilized to address this knowledge gap. Airborne lidar is capable of penetrating vegetation and has been used as a spatial and spectral tool for water body delineation. In the work presented here we: 1) develop a novel means of normalizing lidar ground return intensity to allow for comparison across datasets; 2) demonstrate a statistically significant reduction in median return intensity of wet stream reaches under dense vegetation relative to dry channel reaches; and 3) leverage this reduction to create classified maps of wet and dry channel networks in densely vegetated drainages. Across four study basins spanning over 100 km^2 we observed an average reduction in median intensity of 41.7% and 72.2% for areas with and without dense vegetation, respectively. Optimal thresholds for delineating wet from dry stream reaches were determined using probability density functions. Resulting classified maps yielded overall accuracies ranging from 87.3 – 95.3% when compared to the National Hydrography Dataset via stratified random sampling. This study demonstrates that remote delineation of channel flow states in densely vegetated areas is possible, thus allowing for better consideration, designation, and conservation of headwater channels.

1 Introduction and Background

Wetted channel networks shrink and swell naturally, both seasonally, and in direct response to precipitation and snowmelt runoff events (Buttle et al., 2012; Hooshyar et al., 2015; Leopold & Miller, 1956; Uys & O'Keeffe, 1997). Flow regimes of these networks are altered by climate change, varying land-use, and diversion (Assendelft & van Meerveldt, 2019; Jacobson, 2004; Larned et al., 2010). Monitoring the temporal and spatial dynamics of channel flow states is vital in the fields of hydrology, geomorphology, ecology, and biochemistry (McDonough et al., 2011). For instance, observing wet versus dry reaches in a stream network can inform periodicity designations (i.e., perennial, intermittent, or ephemeral), as well as determinations of channel hydrological states (i.e., flowing, disjointed pools, or dry). Additionally, non-perennial streams provide biodiverse habitats (Meyer et al., 2007; Stehr & Branson, 1938) and migration corridors (Colvin et al., 2009; Erman & Leidy, 1975; Hartman & Brown, 1987), and also serve as biochemical reaction hotspots (McClain et al., 2003). Furthermore, an understanding of a watershed's hydrologic processes is intertwined with the temporal dynamics of its stream networks. Maps of wet channel networks help to improve surface-subsurface modeling (Camporese et al., 2010) and can be used for estimation of downstream flow rates (Godsey & Kirchner, 2014) and water quality (Romaní et al., 2006). Therefore, a high spatiotemporal resolution understanding of channel flow states is essential for models, management, and conservation practices. However, in many cases there are few available data on the wetting and drying dynamics of a drainage network. For example, stream gauges are often only located on relatively large, perennial rivers.

Remote sensing applications have contributed to a broader understanding of channel flow dynamics and periodicity by complementing conventional methods and addressing the limitations of discrete point measurements. For instance, near-infrared (NIR) satellite imagery produced by passive remote sensors has been used to map water bodies for decades, owing to the spectral signature of water (Work & Gilmer, 1976). At NIR wavelengths, water is highly absorptive relative to surrounding soil and vegetation, a contrast that is leveraged to differentiate surfaces in data products such as the Normalized Difference Water Index (NDWI) (McFeeters, 1996; Xu, 2006). However, mapping with passive remote sensing is generally restricted to large, open water bodies, limited by spatial resolution and a general inability to observe surfaces under the cover of forest vegetation (Assenfeldt & van Meerveldt, 2019). Therefore, headwater catchments, where there is often dense vegetation cover and a prevalence of narrow and/or temporarily flowing channels, are often excluded. As a result, the geographic extent and temporal wetting dynamics of headwater streams, generally defined as the uppermost channels in a watershed, are not thoroughly understood (Hooshyar et al., 2015; Leopold, 1994; Meyer & Wallace, 2001; Nadeau & Rains, 2007). This is an unfortunate limitation, as headwater channels are the most abundant streams in both number and length in a drainage network (Horton, 1945; Leopold et al., 1964). Further, according to Colvin et al. (2019), they are thought to comprise 79% of total stream length in the U.S. Headwater streams are vital to ecological health, water quality, and watershed connectivity (Haigh et al., 1998; Wipfli et al., 2007).

Light Detection and Ranging (lidar) via airborne laser scanning (ALS) provides an opportunity to determine the flow status of channels in headwater catchments. As an active remote sensing system, lidar units emit rapid pulses of light (typically at a NIR wavelength) and record both the relative strength of backscattered signal after reflecting off a target, as well as the duration of travel. Each data point, or "return", has spatial coordinates (X, Y and Z) and a return intensity value which provides a relative measure of target reflectance. Thus, lidar generates both a spatial and a spectral data product. Spatial lidar data are used extensively for topographic mapping, specifically towards the generation of Digital Elevation Models (DEMs). Lidar DEMs have been processed to delineate drainage networks (Lashermes et al., 2007; Orlandini et al., 2011; Passalacqua et al., 2010), hence using spatial, topographic data to outline systems of channels. In the context of this study, a drainage network determines where wet channels *could* be, but does not provide any information regarding which stream reaches are wet or dry. Furthermore, researchers have proven that lidar can achieve sufficient vegetation penetration to produce drainage network data products in certain forested areas. For example, James et al. (2007) and James & Hunt (2010) successfully circumvented canopy cover when mapping drainages in the Piedmont region of South Carolina, USA. Though part of the incoming radiation is intercepted by vegetation, enough energy is transmitted to provide spatial information on the ground beneath.

Complex hydrological pathways cannot by accurately portrayed with a sole reliance on topographic information, thus a variety of factors and information sources should be considered (Devito et al., 2005). For example, it has long been recognized that channel classification schemes that consider vegetation community structure and soil moisture patterns provide improved representation of water and sediment conveyance processes (Hack and Goodlett, 1960;

Montgomery 1999). More recently, measurements of surface roughness and texture derived from lidar have proven useful for detecting specific stream channel components including large woody debris (Abalharth et al., 2015). This study builds on the recent work utilizing lidar intensity data to improve drainage network maps by leveraging spectral lidar intensity data to detect the presence of water in channels, much like passive NIR satellite imagery. As mentioned above, water is relatively absorptive at the NIR wavelengths commonly used in lidar ALS systems. Because of this absorption, related studies found that water bodies are characterized by low return intensity (Brzank et al., 2008; Hofle et al., 2009), as more of the incoming radiation is absorbed by water relative to adjacent land areas. The resulting regions of distinctly low intensity have been shown to effectively outline water surfaces and inundated areas (Antonarakis et al., 2008; Lang & McCarty, 2009). Similarly, interactions between NIR lidar and water tend to yield laser shot dropouts, where the reflected signal is so weak that it does not generate a return. The resulting regions of low point density can also be used to delineate water body boundaries (Worstell et al., 2014). However, most studies have focused on large and/or open water bodies, rather than explicitly addressing the problem of small channels under dense vegetation.

Forested areas present a challenge for delineating water bodies with lidar intensity. This is because vegetation occlusion tends to yield ground returns with lower intensity than those of open ground, similar to the low intensity of water targets. Figure 1 illustrates a simplified schematic of four possible lidar return scenarios, and demonstrates this problem. Figure 1a and 1b depict the case of an open area without any obstruction from vegetation, where the lidar beam encounters dry ground and water, respectively. As discussed above, water is absorptive in the NIR spectral region, and so the return intensity from water is markedly lower. In the 2nd row, dry ground and water are obscured by vegetation (Figure 1c and 1d, respectively). Vegetation scattering results in multiple returns from a single pulse that can be used to create maps of vegetation height. However, because only a portion of the outgoing pulse reaches the ground, less energy is available for scattering at the ground level, resulting in distinctly lower ground return intensity relative to open areas, even when dry ground is encountered (e.g., Figure 1a versus 1c). This makes it difficult to rely solely on the spatial distribution of return intensity as an indicator of water when vegetation is present, as low-intensity returns from the scenarios in Figure 1b, 1c, and 1d could easily be conflated.

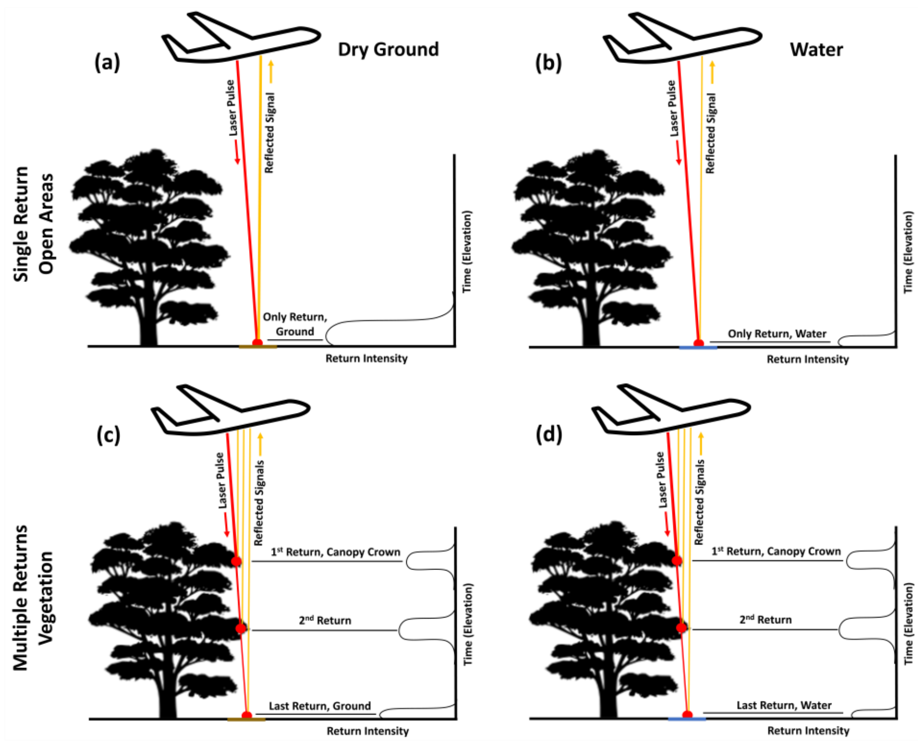

Figure 1. Simplified schematic of four lidar scattering scenarios. Single returns in an open area from dry ground and water (a and b, respectively) are compared to multiple returns from vegetation scattering (c and d) with the same ground cover.

Previous researchers have attempted to circumvent the issue of vegetation obscuring airborne lidar. In Hooshyar et al. (2015), a study focused on headwater catchments with considerable vegetation, the researchers actively masked nearby dense vegetation in their intensity maps to prevent the erroneous classification of wet surfaces. As a result, they acknowledged that their method is limited to partially covered channels. In the case of stream channels largely occluded by vegetation, however, this tactic would result in masking of the water returns we seek to map. Thus, when dense vegetation is prevalent this method becomes impracticable. Figure 2 illustrates this limitation. The left column demonstrates successful masking on a sparsely covered stream, while the more densely forested area in the right column contains a well-hidden channel which complicates masking. To begin, National Hydrography Dataset flowlines confirm the presence of perennial channels at two different locations (Figure 2a and 2b). The presence of low-intensity noise from vegetated areas around the channels is evident in the raw intensity maps for both areas (Figure 2c and 2d). Using vegetation maps (Figure 2e and 2f), pixels coincident with dense vegetation are excluded, resulting in masked ground intensity maps (Figure 2g and 2h). Successful masking on the larger channel in the left column allows for easy delineation. This is not the case in the narrower, more occluded perennial channel in the right column. Because much of the channel is under canopy cover, 89% of the pixels coincident with the channel are caught in the filter and masked, eliminating their intensity data.

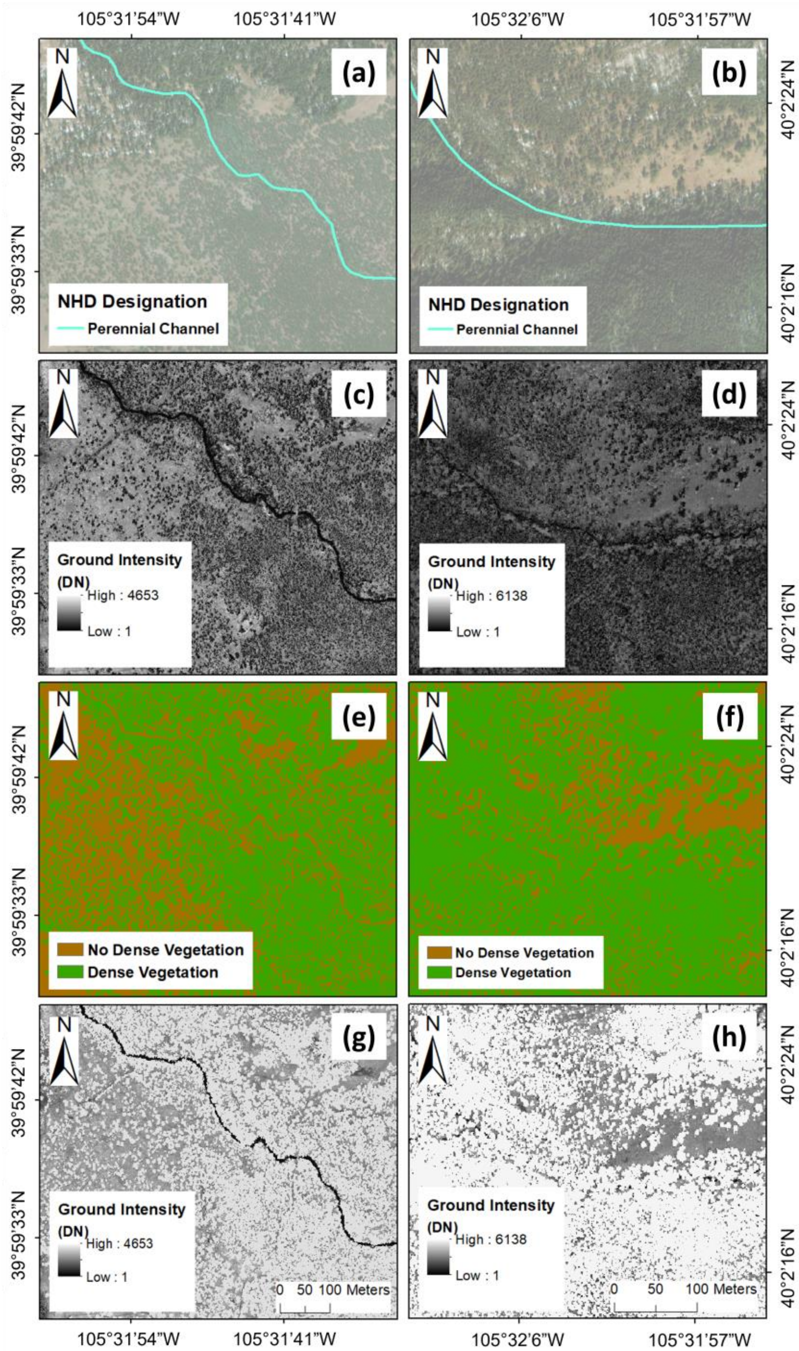

Figure 2. Pilot study results from the study area with aerial imagery (a and b), raw ground return intensity maps (c and d), the extent of dense vegetation (e and f) and resulting ground intensity maps after masking densely vegetated areas (g and h). The left column demonstrates successful masking allowing for delineation of a larger, sparsely vegetated channel, while the right column illustrates how vegetation masking is impracticable on smaller, occluded channels. Dense vegetation refers to locations where the elevation difference between the top of the forest canopy and the ground is greater than 2 m, following Hooshyar et al. (2015), while no dense vegetation essentially implies an open area.

In summary, the capacity of lidar to map drainage networks under dense vegetation using spatial data has been demonstrated, and spectral lidar intensity data have been used to outline water bodies. However, determining which channels are wet versus dry in headwater catchments continues to present a challenge, particularly in forested regions or areas with dense riparian vegetation. Though airborne lidar can penetrate canopy, the use of lidar intensity data to locate wetted channel networks under vegetation is an unvalidated method. In the work presented here, we attempt to address this knowledge gap by presenting a novel method for classifying wet and dry stream reaches in forested areas. Intuitively we note that, in the case of identical vegetation conditions, liquid water under canopy should produce a lower return intensity signal when compared to dry ground, due to its lower NIR reflectance. Therefore, *we hypothesize that a quantifiable reduction in ground return intensity exists when a lidar signal encounters both dense vegetation and liquid water at the ground surface, relative to vegetation with dry ground beneath*. To explore this prospect, we selected study basins in California and Colorado with lidar, hydrometric data, and dense riparian forests. We first use spatial lidar data to delineate drainage networks and identify the presence of occlusive vegetation. Next, frequency distributions of ground return intensity under dense vegetation are evaluated to statistically compare wet versus dry occluded channels. Per our hypothesis, an intensity reduction in occluded wet channels relative to dry channels is expected. A quantifiable intensity reduction can then be leveraged to create a classified map product depicting wetted channel connectivity in headwater catchments with dense vegetation. Lidar-derived maps of wet channel networks could be used to determine headwater stream periodicity in regions lacking reference data, to observe intraannual wetting and drying dynamics at high spatiotemporal resolution, or to detect changes in stream permanence due to changing climate and land use. Specifically, there are no existing approaches for broad scale tracking of the potential expansion of intermittent and ephemeral channel extent in response to changes in frequency and intensity of drought.

2 Study Basins and Data

2.1 Study Basins

Three study areas (Basins 1 – 3) along the western shore of Lake Tahoe, CA, also utilized by Hooshyar et al. (2015), are revisited (Figure 3a). These basins benefit from overlapping lidar and hydrograph data. The boundaries of Basins 1 – 3 are delineations of complete catchments, each with a United States Geological Survey (USGS) stream gage at the outlet. Basin 1 features an additional gage along Ward Creek several kilometers upstream of Lake Tahoe. Basin 4 includes most of the Middle Boulder Creek catchment, except for an area lacking lidar data availability (Figure 3b). Basin 4 also drains to a continuous stream gage, managed by the Colorado Department of Water Resources, before flowing into the Barker Meadow Reservoir. Relevant historic hydrometric statistics and flow at the time of lidar acquisitions were obtained for these stations. Furthermore, Basin 4 lidar data were acquired in May and August, allowing for a temporal comparison. The western side of Lake Tahoe receives a mean annual precipitation of 140 cm, while Nederland, CO on the eastern end of Basin 4 reports 72 cm. Precipitation at all basins primarily falls between November and May, with most runoff occurring during spring snowmelt, typically peaking in May and June. Basins are largely covered by dense, intact forest, with minimal disruption from human development and urbanization. Study basin areas ranged from 18.2 – 31.9 km^2 (Table 1).

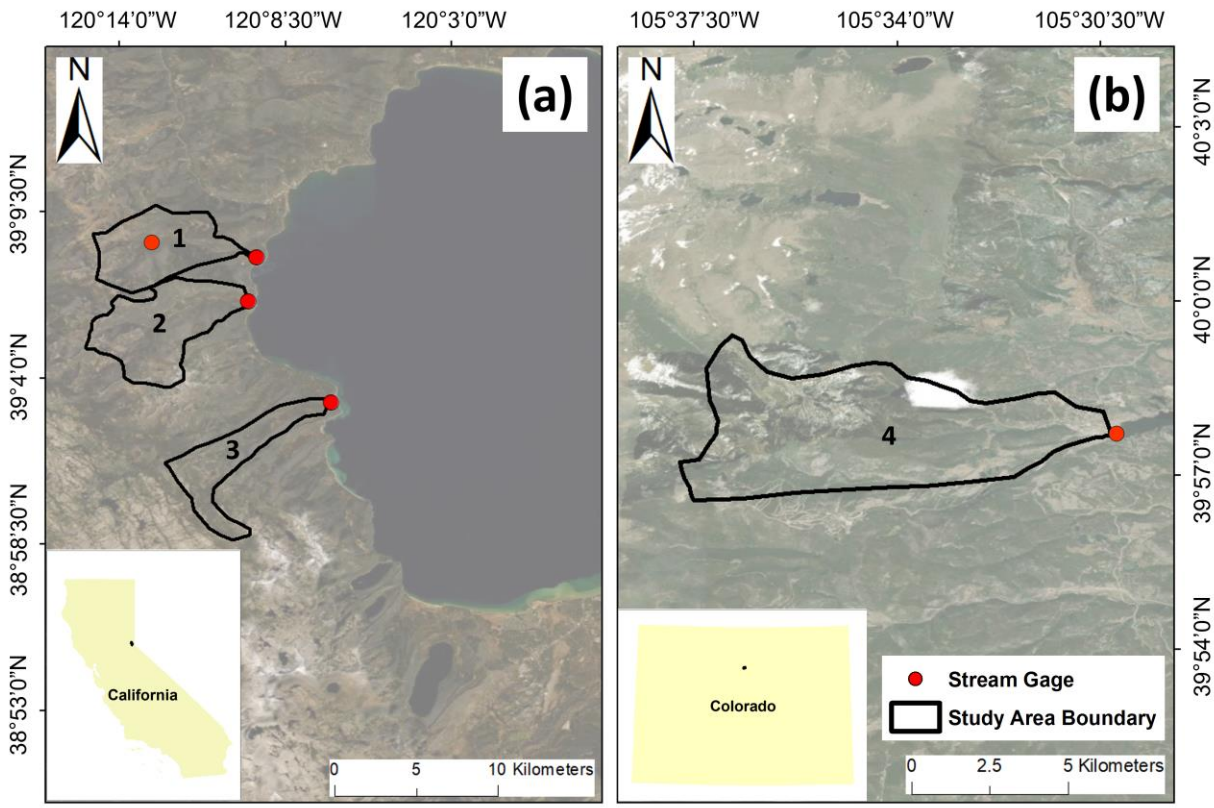

Figure 3. Map of study basin locations.

Table 1. Study basin statistics. Acquisition flow rates refer to the mean measured flow at the stream gages during periods of lidar scanning. Mean annual, 30-day low, and peak flow rate metrics represent the 2010 water year for each gage. Values are also presented as a percentage of the long-term mean annual discharge (% LT MAD) consistent with Tennant (1976) where flows >200 % represent flushing flows, and flows < 10 % represent severe degradation for fish resources.

Site #	Primary Perennial Stream	Area (km^2)	Lidar Acquisition Date(s), 2010	Gage (#)	Mean Acquisition Flow Rate (m^3/s)	30-Day Low Flow Rate (m^3/s)	Mean Annual Flow Rate (m^3/s)	Peak Flow Rate (m^3/s)
1	Ward Creek	23.1	14 Aug	USGS 10336676	0.06 (8.3%)	0.02 (2.8%)	0.68 (93.7%)	8.63 (1189.0%)
				USGS 10336674	0.07 (15.3%)	0.03 (6.6%)	0.40 (87.5%)	5.26 (1151.0%)
2	Blackwood Creek	28.4	20 - 23 Aug	USGS 10336660	0.11 (11.0%)	0.04 (4.0%)	0.89 (88.6%)	10.61 (1056.8%)
3	General Creek	18.2	20 - 23 Aug	USGS 10336645	0.02 (4.3%)	0.02 (4.3%)	0.45 (97.0%)	6.11 (1316.8%)
4	Middle Boulder Creek	31.9	20 - 21 May 21 - 26 Aug	CO DWR BOCMIDCO	2.26 (138.9%) 0.82 (50.4%)	0.09 (5.5%)	1.44 (88.5%)	15.48 (951.4%)

2.2 Data

Lidar data for Basins 1 – 3 were acquired from the Lake Tahoe Basin LiDAR dataset, collected by Watershed Sciences, Inc. in conjunction with the Tahoe Regional Planning Agency. Data for Basin 4 was from the Boulder Creek Critical Zone Observatory LiDAR Survey collected by the National Center for Airborne Laser Mapping (NCALM). All data were accessed through the Open Topography public repository. Researchers at Basins 1 – 3 utilized a 1064 nm laser, while Basin 4 researchers used 1047 nm, both NIR wavelengths. Other relevant statistics from the data collection campaigns are listed in Appendix Table A1. To reiterate, each lidar return has spatial coordinates (X, Y, and Z) and a range-calibrated return intensity value. A single pulse can generate multiple returns when vegetation is present, as previously illustrated in Figure 1c and 1d. This allows for elevation and intensity data to be segregated via return number, such that sub-datasets can be determined for both the ground and canopy crown layers.

In the Tahoe area, lidar data at Basins 1 and 2 were collected while discharge approached 30-day low flow rates, while Basin 3 lidar data were collected at the 30-day low flow rate (Figure 4a). Therefore, scans coincided with flows expected to cause dry channels in some locations. Weather data procured from the National Climate Data Center (NCDC) revealed no precipitation during lidar data collection at the nearest station in Tahoe City, CA. Wet and dry season lidar data are available for Basin 4, as aerial scanning was conducted in both May and August of 2010 (Table 1). Both outings coincided with transitional periods (Figure 4b). May data collection in Basin 4 took place after the watershed had surpassed the mean annual flow rate (139% LT MAD), 17 days before peak flows were reached. In contrast, the August data collection in Basin 4 occurred during hydrograph recession, with flows at 50% LT MAD.

Hereafter, the May and August lidar datasets for Basin 4 will be referred to as Basins 4M and 4A, respectively. Once again, weather data provided by the NCDC showed no precipitation at the Nederland, CO station during either data collection campaign.

The National Hydrography Dataset (NHD), accessed via the USGS National Map, was used as reference data. The NHD is considered to be the most comprehensive and up-to-date hydrography dataset for the United States, used extensively from local to national scales (Hafen et al., 2020; Nadeau & Rains, 2007; Simley, 2006). The positions of NHD flowline features, including streams, were manually surveyed by USGS field crews using the standards of the day for Digital Line Graphs (https://pubs.usgs.gov/circ/1984/0895g/report.pdf). During topographic inspection, stream periodicity classifications (i.e. perennial, intermittent, or ephemeral) were determined based on several factors, including previous observations, feedback from local residents, water containment, soil type, and vegetation at the time of the site visit (Personal communication from D. Anderson, USGS Hydrography Program, to J. Dillon, November 17, 2022). Periodicity determinations were meant to represent stream permanence across a normal range of climactic conditions (Hafen et al., 2020). Furthermore, many NHD flowline geometries and periodicities have since been updated by partnering state agencies as part of a stewardship program, leveraging local knowledge. All flowlines in Basins 1 – 3 were updated between 2012 and 2016, while flowlines in Basin 4 were recompiled in 2012. Thus, position and periodicity determinations for streams in all study basins were recently assigned by state agencies based on field observations. The USGS defines perennial streams as those which normally have water in their channel year-round, except for infrequent periods of severe drought (https://nhd.usgs.gov/userguide.html). Intermittent streams tend to flow seasonally, only when receiving water from rainfall runoff, springs, or other surface sources such as snowmelt. Last, ephemeral streams flow only in direct response to precipitation. They receive little to no water from springs, snowmelt, or other sources, and their channels remain above the water table at all times. Based on these definitions and the data in Table 1 and Figure 4, the following classification rules were created for reference data:

Basins 1 – 3, 4A

- Perennial streams were assumed to be wet
- Intermittent and ephemeral reaches were presumed dry

Basin 4M

- Perennial streams were assumed to be wet
- Intermittent reaches were presumed wet; at 139% LT MAD, flows were likely high enough to activate intermittent channels, with snowmelt runoff driving the exponential rise of the hydrograph
- Ephemeral streams were still assumed to be dry, as flows at only 15% of peak discharge likely were not high enough to activate ephemeral channels

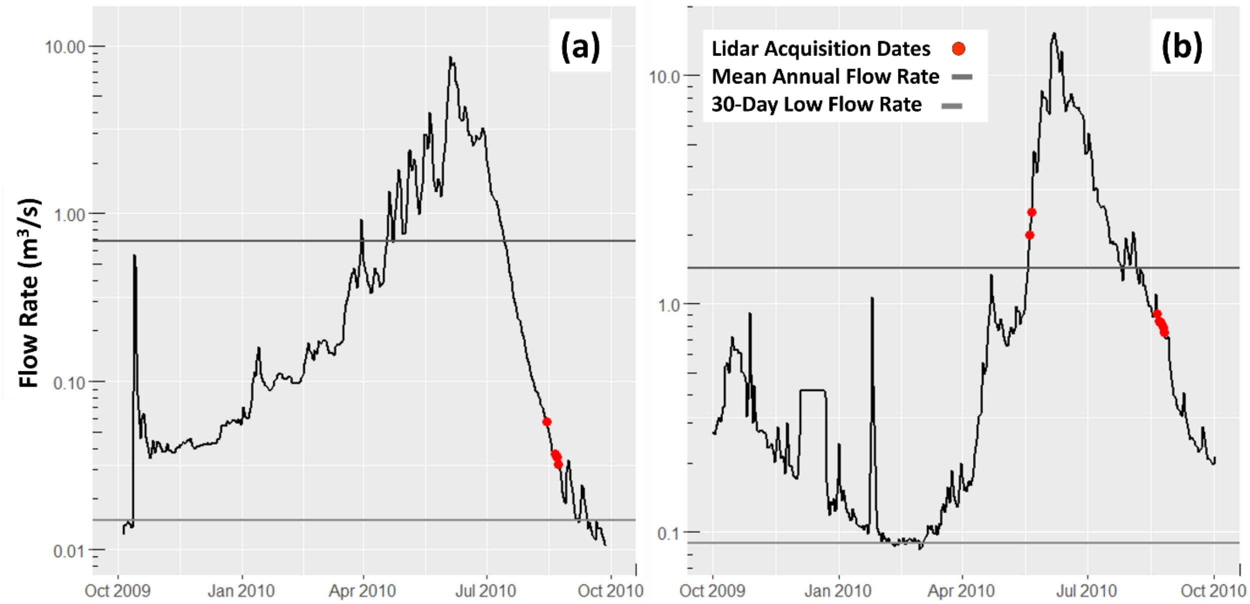

Figure 4. Hydrographs and lidar acquisition dates for Basin 1 (a) and Basin 4 (b). The hydrograph for Basin 1 uses data from the downstream gage (# 10336676) at the mouth of Ward Creek. The Basin 1 hydrograph is representative of Basins 2 and 3 (as well as the upstream gage at Basin 1), and includes lidar acquisition dates for all Lake Tahoe study basins. Horizontal lines depict the mean annual flow and the 30-day low flow rate for the 2010 water year.

3 Analysis and Methodology

Water surfaces are characterized by low return intensity relative to dry ground, a spectral signature that is used to map wet channels via an intensity threshold (Antonarakris et al., 2008; Brzank et al., 2008; Hofle et al., 2009). However, dense vegetation generates low-intensity noise that complicates water body classification. We hypothesized that, even when under dense vegetation, wet stream reaches will produce a significantly lower return intensity relative to dry ground under vegetation. In the work presented here we statistically validate that such a reduction is consistently quantifiable and reproduceable. Further we leverage this phenomenon, along with spatial lidar data, to wet from dry channels in densely vegetated areas. Our analyses consisted of four major tasks: 1) development of relevant map layers, 2) normalization of intensity data, 3) statistical analysis and 4) generation of classified maps. All work was performed using ESRI's ArcMap 10.7.1 (ESRI, 2022) and the statistical programming software R (R Core Team, 2022).

3.1 Map Layer Development

Several pertinent map layers were produced from raw lidar data (Figure 5). To begin, subgroups of ground returns, as classified by the data collectors (Watershed Sciences, Inc. and NCALM; Section 2.2) using TerraSolid's TerraScan software, were separated from the remainder of the lidar dataset. The intensity values from ground returns were interpolated to create a map of ground return intensity (Figure 5a). Similarly, elevation values from ground returns were interpolated to yield a DEM. Using the workflow of Omran et al. (2016), the DEMs were processed to produce maps of drainage networks (Figure 5b). Next, first returns and only returns were separated, and this subgroup's elevation data were interpolated to produce a Digital Surface Model (DSM). The DSM is identical to the DEM when vegetation is absent but represents the elevation of the top of the canopy when it is present. Subtracting the DEM from the DSM results in a map of canopy height. This canopy height model is then binarized; pixels with a vegetation height value greater than 2 m were categorized as "dense vegetation", while all other pixels were classified as "no dense vegetation" (Figure 5c), following Hooshyar et al. (2015). The spatial resolution of all maps was 1 m. Last, maps of NHD stream classifications served as reference data (Figure 5d).

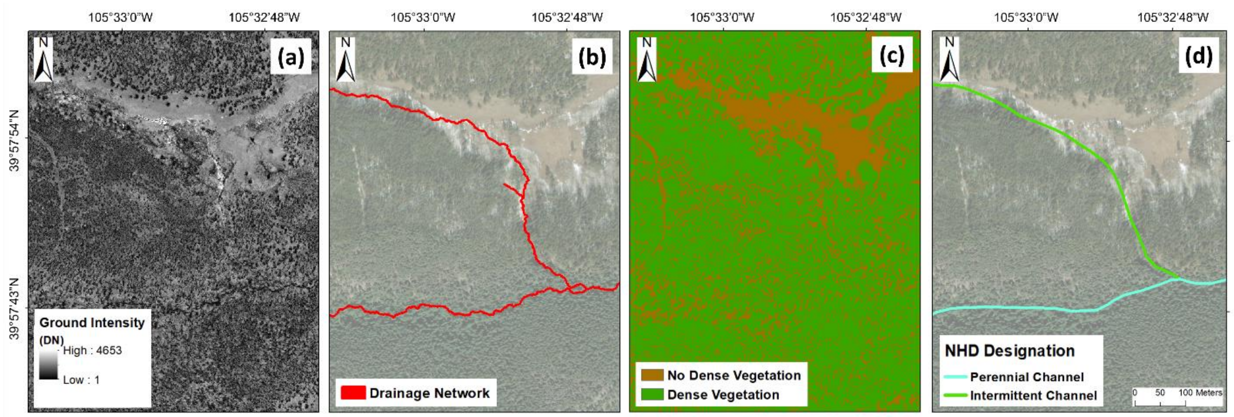

Figure 5. Core map layers to be used in the following analyses: ground intensity (a), drainage network delineation (b), binarized vegetation (c), and National Hydrography Dataset channel designations as reference data (d). The perennial channel shown is North Beaver Creek in Basin 4.

3.2 Intensity Normalization

To compare values across different basins and datasets, we normalized all intensity maps using a common land cover feature. Given our interest in canopy obstruction, we examined the average intensity of dry pixels under dense vegetation. For each dataset, a series of overlays were used to isolate ground intensity pixels coincident with dense vegetation and outside of the drainage network. The mean values of these intensity subsets were calculated, and termed $\bar{I}_{ND,V}$; the mean intensity of pixels both outside the drainage network, or non-drainage (ND), and under dense vegetation (V). This value was used as the normalizing factor for each dataset via Equation 1:

$$I^{nor} = {}^{I}/_{\bar{I}_{ND,V}} \tag{1}$$

where I is the raw intensity and I^{nor} is the unitless normalized result for a given pixel. Hereafter, the superscript "nor" implies that an intensity value has been normalized, while subscripts denote the data subsets that the value is associated with. By applying Equation 1, we normalized by the *mean effect of canopy occlusion* for a given basin/dataset, allowing for comparison across datasets spatially and temporally. For example, a value of I^{nor} less than 1 denotes a pixel with lower intensity than the average value of dry ground under dense vegetation.

3.3 Statistical Analysis

The goal of statistical analysis was to investigate our hypothesis: that a significant reduction in return intensity occurs in wet channel pixels relative to their dry channel counterparts, *even when under dense vegetation*. Further, we sought to determine the optimal thresholds of normalized intensity (I^{nor}) to map wet channels within a drainage network, whether open or occluded by dense vegetation.

For each dataset, pixel values of I^{nor} were first divided into four subgroups: 1) wet channels/dense vegetation, 2) dry channels/dense vegetation, 3) wet channels/no dense vegetation, and 4) dry channels/no dense vegetation. Figure 6 describes this segregation workflow; recall subscripts correspond to dataset subgroups. To begin, maps of I^{nor} were truncated to only include pixels within the drainage network (D), termed I_D^{nor}. Next, I_D^{nor} values were further sorted by the presence (V) or absence (NV) of dense vegetation ($I_{D,V}^{nor}$ and $I_{D,NV}^{nor}$, respectively). Last, pixels were separated into bins coincident with wet streams versus dry channels, using the description and reference data rules in Section 2.2. For instance, following the convention discussed above, $I_{D,V,Wet}^{nor}$ would refer to the normalized intensity of pixels within the drainage network (D), under dense vegetation (V), and coincident with a known wet stream segment (Wet) as determined by the NHD and hydrograph data. To reiterate, NHD-designated perennial channels were assumed wet across all datasets, and all other channels presumed dry with the exception of intermittent streams in dataset 4M (Section 2.2).

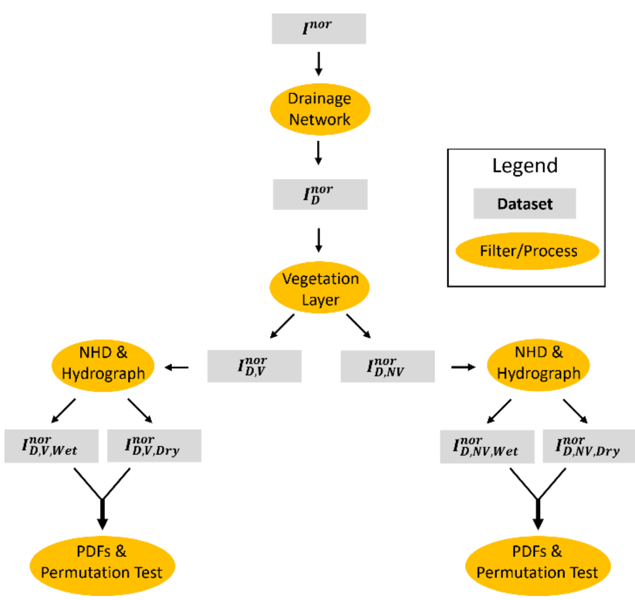

Figure 6. Workflow for subdividing pixels of Normalized Intensity (I^{nor}) maps.

For each dataset, a pair of significance tests were conducted to evaluate the difference in I^{nor} distributions between wet and dry channel pixels; one for the case of dense vegetation and the other for no dense vegetation. In other words, $I^{nor}_{D,V,Wet}$ was compared to $I^{nor}_{D,V,Dry}$, while $I^{nor}_{D,NV,Wet}$ was compared to $I^{nor}_{D,NV,Wet}$ (Figure 6). To restate our hypothesis, we expect that $I^{nor}_{D,V,Wet}$ and $I^{nor}_{D,NV,Wet}$ values will be significantly lower than their "Dry" counterparts based on previous work and radiative transfer theory. When comparing median values, our hypotheses can be summarized as follows:

$$H_0\text{: } M(I^{nor}_{D,V,Wet}) = M(I^{nor}_{D,V,Dry}) \ \ and \ \ M(I^{nor}_{D,NV,Wet}) = M(I^{nor}_{D,NV,Dry})$$

$$H_A\text{: } M(I^{nor}_{D,V,Wet}) < M(I^{nor}_{D,V,Dry}) \ \ and \ \ M(I^{nor}_{D,NV,Wet}) < M(I^{nor}_{D,NV,Dry})\ ; \ \ \alpha = 0.05$$

Each of the four subgroups were randomly sampled with replacement to a sample size of 3198 pixels, the size of the smallest dataset. Doing so ensured equal sample sizes and addressed the rather large range in pixel count values across datasets. A one-sided permutation test was preferred to a traditional t – test, as assumptions of equal variance and normality appeared to be violated in some cases. Further, the permutation test was conducted to compare median I^{nor} values between datasets, rather than means. Although a comparison of means is a typical default, a preliminary examination of intensity distributions revealed large outliers in some datasets. As outliers can produce mean values that poorly represent the overall distribution, medians were analyzed to mitigate the presence of outliers. Resulting p-values represent the probability that a difference in median I^{nor} values between wet and dry channel pixels is due to random variation. Therefore, low p-values would provide evidence for our hypothesis. P-values for each comparison were recorded, as well as median and standard deviation information for each dataset.

Last, frequency analyses were performed on intensity maps to create probability density functions (PDFs), used to visually contrast the distributions of wet versus dry stream channels for both the case of dense vegetation and no dense vegetation. For each dataset and vegetation state, the intersection of wet and dry PDFs was deemed the optimal threshold to delineate wet from dry pixels in a drainage network (Figure 7). These optimal cutoff points, or critical thresholds, were termed $I^{nor}_{Crit,V}$ and $I^{nor}_{Crit,NV}$, ideal delineation values for dense vegetation (V) and no dense vegetation (NV), respectively.

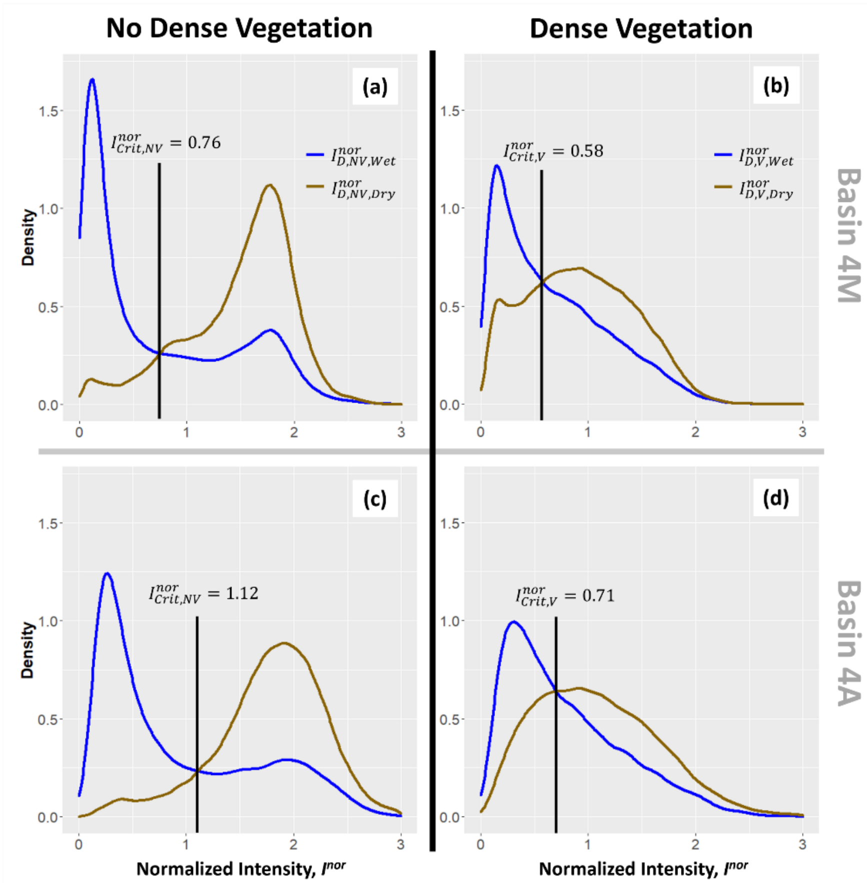

Figure 7. Example probability density functions to contrast intensity distributions of wet and dry pixels. The distributions shown are from datasets 4M and 4A. Vertical black lines at distribution intersections represent critical thresholds for delineating wet from dry streams for each dataset and vegetation state.

3.4 Wet Channel Classification

The statistical analyses presented in Section 3.3, specifically values of $I^{nor}_{Crit,V}$ and $I^{nor}_{Crit,NV}$, were used to produce a classified map product. Each pixel in the drainage network was classified as wet or dry. An outline of this workflow is presented in Figure 8. Once again, maps of I^{nor} (Figure 8a) were truncated using the drainage network (Figure 8b) to include only pixels within the network, I^{nor}_{D} (Figure 8c). Next, depending on whether a drainage pixel was coincident with dense vegetation (Figure 8d), the appropriate critical threshold from the PDF (Figure 8e) was applied to classify pixels as wet or dry. In the example presented, drainage pixels under dense canopy with an I^{nor}_{D} value less than 0.63 would be classified as wet, and vice versa. Last, following previous models (e.g., Hooshyar et al., 2012), a final smoothing step was performed to minimize the influence of isolated misclassified pixels on overall accuracy. All pixels in a given stream link were reclassified to the majority assignment, resulting in the final data product (Figure 8f). A stream link is defined as a section of channel between two successive junctions.

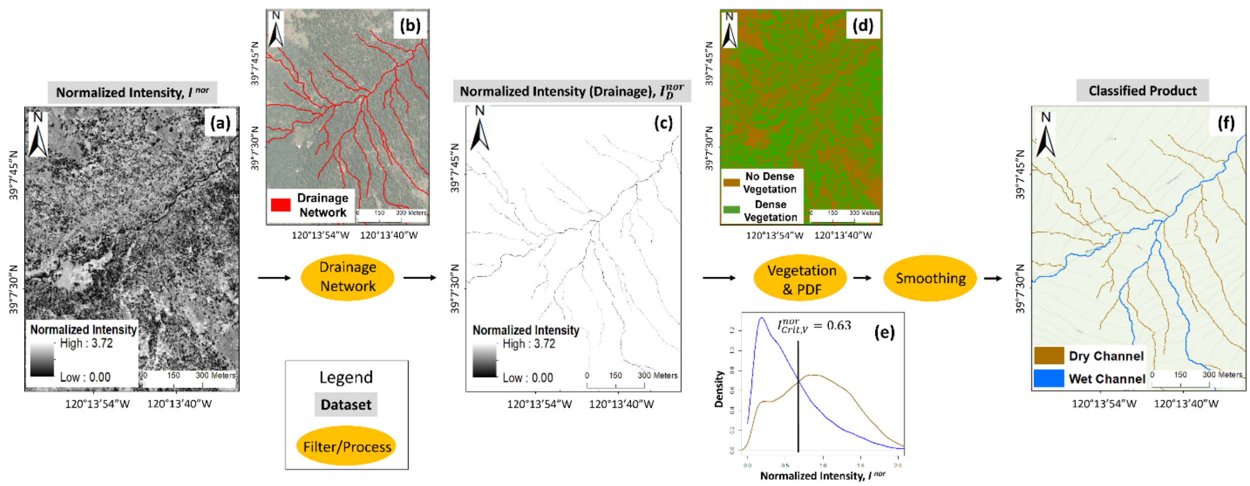

Figure 8. Demonstration of classification workflow. A normalized ground intensity map, I^{nor} (a) is trimmed by the coincident drainage network (b) to only include pixels within the network, I_D^{nor} (c). Using a binarized vegetation map (d) and a probability density function analysis (e), a classified product is produced (f). The channel shown is Ward Creek in Basin 1.

The accuracy of each classified map was assessed against the NHD with a stratified random sample of 300 points. Following Foody (2009), a minimum total sample size of 200 points was determined based on an expected accuracy of 85% and allowable sampling error of 5%. This value was then increased to 300 points to ensure a minimum of 50 points per each stratum, considering the popular heuristic metric dating back to Hay (1979). Confusion matrices of the form depicted in Figure 9 were produced to display classification results. Perhaps the most widely utilized accuracy assessment approach in remote sensing, a confusion matrix is a table used to clearly and concisely summarize the performance of a classification algorithm (Foody, 2002). For example, in Basins 1 – 3 where low-flow conditions were present at the time of lidar data acquisition, it was assumed that only perennial channels were flowing (see description and reference data rules in Section 2.2). Therefore, reaches classified as wet that were coincident with NHD-designated perennial channels would qualify as a TP, while wet classifications on all other channels were FPs. Accuracy was calculated with Equation 2 from the seminal work of Story and Congalton (1986). Furthermore, modeled values of total wet channel length from classified maps were compared to those of the NHD.

		Modeled	
		Wet	*Dry*
Reference	*Wet*	TP	FN
	Dry	FP	TN

Figure 9. Example confusion matrix. In this case true positives (TPs) refer to the number of wet pixels correctly classified as wet, false negatives (FNs) to the number of pixels classified as dry that were actually wet, and so on. Reference designations of wet and dry regions were based off of the National Hydrography Dataset according to the rules in Section 2.2

$$Accuracy, A = \frac{TP + TN}{TP + TN + FP + FN} \qquad (2)$$

4 Results

4.1 Summary of Basic Statistics and Significance Testing

Frequency distributions for samples of I_D^{nor} had significantly lower median values in wet channels when compared to dry channels (Table 2). A reduction was observed both when dense vegetation was present (a novel result) and in its absence (an expected result). Under dense vegetation the average reduction in median $I_{D,V}^{nor}$ values was 41.7%, compared to a 72.2% intensity reduction in open areas, $I_{D,NV}^{nor}$. In each case, one-sided permutation tests produced values of $p < 0.001$ across all datasets, and thus intensity reductions in wet relative to dry channels were highly significant. In other words, the probability that these reductions were caused by random variation is very low. Therefore, our primary hypothesis was confirmed: even when dense vegetation created low-intensity noise in dry ground returns, *the median intensity of occluded wet channels was quantifiably lower.*

Table 2. Probability density function and classification results. Asterisks (*) adjacent to wet intensity statistics denote statistically significant reductions relative to dry pixels at α=0.05.

Dataset	Median (sd) $I^{nor}_{D,V,Wet}$	$I^{nor}_{Crit,V}$	Median (sd) $I^{nor}_{D,V,Dry}$	Median (sd) $I^{nor}_{D,NV,Wet}$	$I^{nor}_{Crit,NV}$	Median (sd) $I^{nor}_{D,NV,Dry}$	Modeled Wetted Length (km)
1	0.41 (0.59)*	0.63	0.97 (0.68)	0.21 (0.61)*	0.75	1.52 (0.61)	19.0
2	0.43 (0.56)*	0.53	0.76 (0.59)	0.27 (0.61)*	0.76	1.46 (0.54)	23.7
3	0.76 (0.55)*	0.92	1.10 (0.56)	0.75 (0.60)*	1.14	1.55 (0.46)	22.7
4M	0.53 (0.70)*	0.58	0.89 (1.31)	0.40 (0.81)*	0.76	1.62 (1.48)	32.0
4A	0.66 (0.78)*	0.71	1.03 (0.61)	0.62 (1.77)*	1.12	1.84 (1.22)	29.5

4.2 Classified Network Maps

Reference wet channel lengths from the NHD are compared to modeled values from classified maps in Figure 10. Modeled data was in reasonably good agreement with the NHD, although a tendency to underestimate wet channel length is observed, particularly in dataset 4M. Classified maps for Basins 1 – 3 and their coincident NHD designations are illustrated in Figure 11, along with Accuracy, *A* values from stratified random sampling. The same can be found for Basin 4 in Figure 12, with a temporal comparison between datasets 4M and 4A. Accuracy ranged from 87.3 – 95.4% with a median of 93.0%. In Basins 1 – 3, ephemeral and undesignated channels were almost entirely rejected from the wet channel classification, and much of the perennial channel length was correctly classified. These basins performed slightly better than 4M and 4A, with a minimum accuracy of 93.0% ranging to 95.4%. In both 4M and 4A most perennial channels were correctly designated as wet and ephemeral as dry. Further, all intermittent reaches were classified as dry in 4A, while several intermittent stretches were classified as wet channels in 4M, as we would expect. Therefore, True Negative (TN) classifications were abundant in 4M and 4A, while False Positives (FPs) were almost nonexistent. False Negatives (FNs) were the primary source of error in each case, consistent with the low bias seen in Figure 10.

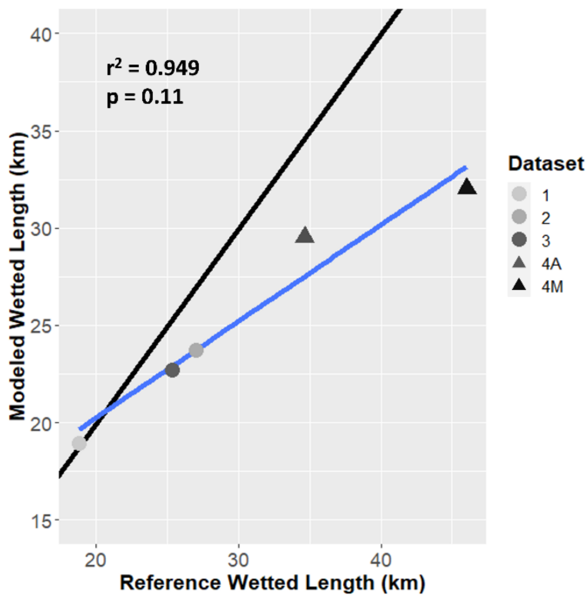

Figure 10. Modeled wet stream length values from the classification workflow compared to National Hydrography Dataset data. The blue line is a linear best fit while a 1:1 reference line is in black. Values are colored by dataset, while shapes correspond to general locations. Hence Tahoe Basins 1 – 3 are circular and both datasets for Basin 4 in Colorado are triangular.

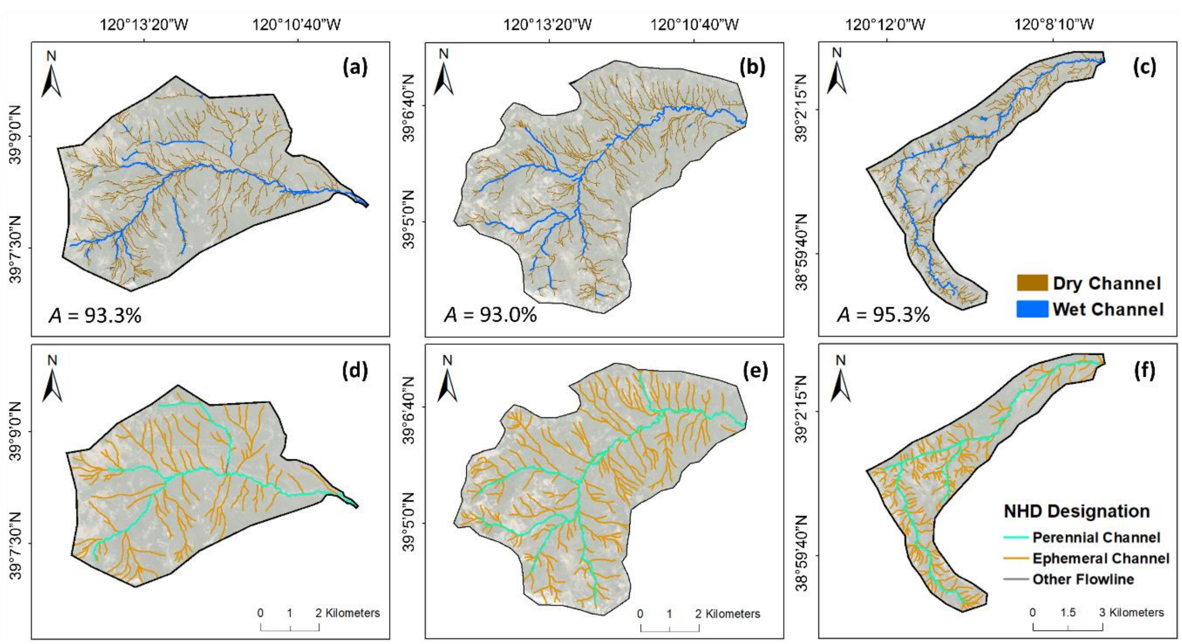

Figure 11. Classified maps depicting the extent of modeled wet and dry channel lengths (a – c) compared to NHD channel designations (d – f) for Basins 1 – 3 overlaid on aerial imagery.

We observe that a portion of classification error can be attributed to improper network delineation, rather than a failure of our ground intensity reduction hypothesis. The delineation workflow of Omran et al. (2016) was primarily selected for drainage network delineation due to its demonstrated capability, ArcMap compatibility, and computational efficiency. Though resulting drainage network maps exhibited good agreement with NHD flowlines, some artifacts were visibly observable, namely due to interference from roads. For example, the perennial channel in the northeast quadrant of Figure 12 (North Beaver Creek, red rectangle) was sporadically misclassified as dry in both 4M and 4A, despite being a relatively large, continuously wetted stream. In these cases, the drainage network maps occasionally deviated from the actual stream channel. Thus, intensity data were incongruent with spatial network delineations, and misclassification ensued. As a result, we note the importance of site-specific network segregation prior to integrating lidar intensity data.

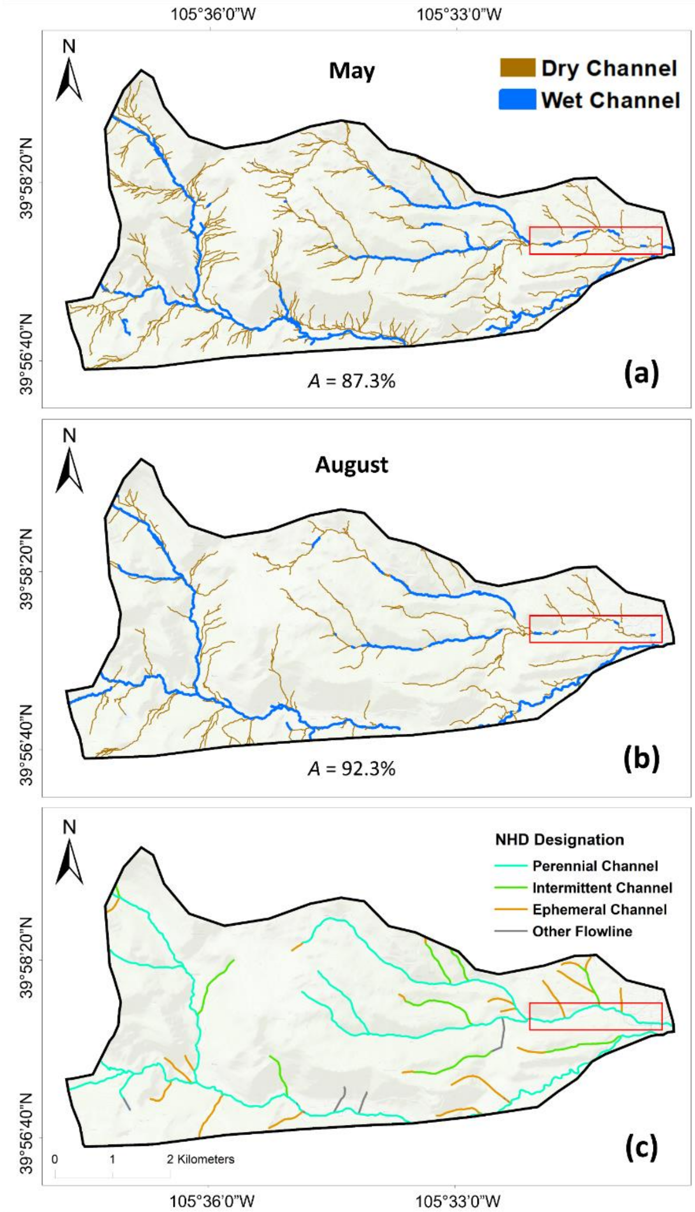

Figure 12. Classified maps depicting the extent of modeled wet and dry channel lengths (a and b) compared temporally and to National Hydrography Dataset channel designations (c) for Basin 4 atop topographic relief. The red rectangle highlights an area where inaccurate spatial drainage delineation caused misclassification.

5 Discussion

In the work presented here we 1) developed a novel means of normalizing lidar ground return intensity to allow for comparison across datasets, 2) demonstrated a statistically significant reduction in median return intensity of wet channels under dense vegetation relative to dry channels, and 3) leveraged this reduction to create maps of wet channel networks in densely vegetated drainages. In practice, any representative area where hydrologic conditions are known could be used to calibrate, or train the model by finding the optimal thresholds $I_{Crit,V}^{nor}$ and $I_{Crit,NV}^{nor}$. Using these threshold values, the model could then be spatially extrapolated to delineate wet channels in the associated catchment across a broader area. Data for this representative calibration area could come from ground-based reference observations, a dataset such as the NHD, or coincident aerial imagery, depending on availability. For example, Figure 13a depicts a small, randomly selected area from Basin 2, where a perennial stream (Blackwood Creek) can be seen weaving in and out of canopy cover. A probability density analysis of intensity data from this area (Figure 13b) produced an $I_{Crit,V}^{nor}$ value of 0.56, only slightly greater than the result for the whole of Basin 2 ($I_{Crit,V}^{nor}$ = 0.53). Therefore, using only this ~ 0.15 km^2 area for calibration, the model could have been extrapolated over the entire Basin 2 drainage area with comparable results to those presented in Section 4. Alternatively, if no ground truth, imagery, or other reference data is available, a practitioner could utilize gaussian mixture models, as in Hooshyar et al. (2015). However, as we have demonstrated, data coincident with dense vegetation should be evaluated separately, even with this approach. Last, although spatial extrapolation appears practical when evaluated carefully, it is important to note that temporal extrapolation is likely imprudent, especially over seasonal time scales. Changes in ground cover, the state of vegetation (e.g., pre- versus post-leaf out), the weather conditions, or the sensor used, for example, could shift optimal thresholds between data acquisitions. This is evident when we compare thresholds between Basin 4M and 4A (Table 2).

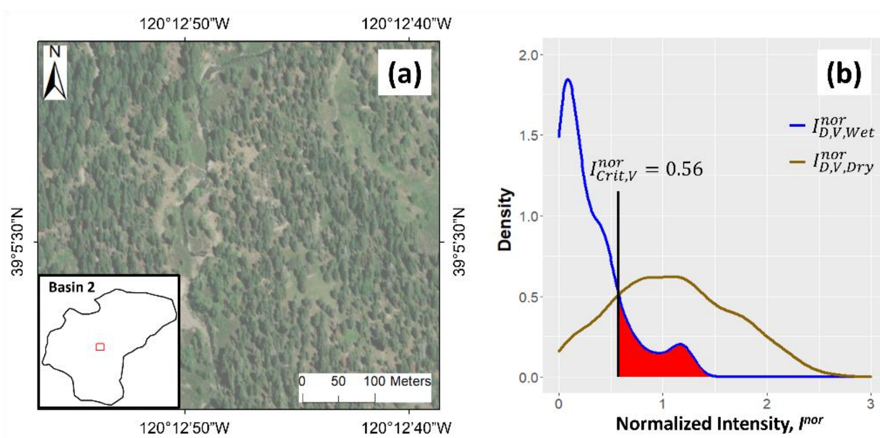

Figure 13. Randomly selected, representative area along Blackwood Creek in Basin 2 (a) used to demonstrate model calibration in practice. Probability density functions from this calibration area (b) produced nearly the same critical threshold as that of the entire basin. The red highlighted area under the curve represents the probability of misclassification, used to quantify model uncertainty.

The classification workflow demonstrated a median accuracy of 93.0% across all datasets, though with notable strengths and weaknesses. For instance, the model was exceptional at excluding NHD-designated intermittent and ephemeral channels from the wet channel classification during low flow conditions, meaning there were very few FPs. However, we qualitatively note that our model misclassifies perennial channels as dry more frequently as proximity to channel heads increases, where stream widths tend to be at a minimum. This likely implies a stream width limitation in our classified product related to lidar ground return point spacing, where an insufficient density of low-intensity pixels results in misclassification. Misclassification near perennial channel heads contributed to our classification workflow's tendency to underestimate wetted channel lengths when compared to the NHD, as noted in Section 4. Model uncertainty can be quantified by the area under probability density curves, as demonstrated in Figure 13b. The area under the blue $I_{D,V,Wet}^{nor}$ curve beyond the critical $I_{Crit,V}^{nor}$ threshold value (highlighted in red) is 0.14. Thus, in this case we would anticipate wet pixels under dense vegetation to be misclassified as dry (hence false negatives, FN) about 14% of the time.

We note that there is a certain degree of unquantifiable uncertainty involved in our validation procedure, namely from the lack of ground truth. Though careful site/dataset selection, NHD periodicity classifications, hydrograph, and precipitation data provided an excellent understanding of the hydrologic conditions during lidar data acquisition, it is impossible to know for certain which streams were wet versus dry in the absence of ground-based observations. Simply put, reliance on static NHD permanence designations is an inherent source of uncertainty. Though the NHD is the most comprehensive hydrography dataset in the United States, its accuracy continues to evolve and improve. For instance, Hafen et al. (2020) found that it is not uncommon for in situ streamflow observations to disagree with NHD periodicity classifications, and noted the dynamic influence of climate variability on this disagreement. Further, the coarse spatial resolution of NHD periodicity classifications (relative to the 1 m resolution used here) presents a limitation. Complex oscillations between perennial and temporary streams over relatively short spatial scales are frequent in nature, due to localized elevated water tables, for example. Unfortunately, extensive in situ observations are almost never coincident with historical lidar datasets.

It is apparent that smoothing each stream link segment by the majority pixel classification resulted in occasional discontinuity in the wet channel network (Figures 11 and 12). Although perhaps aesthetically unappealing, this discontinuity likely provides useful insight. As mentioned above, disjointed sections may indicate a need to re-evaluate or increase the spatial resolution of stream periodicity assignments. Furthermore, temporary (intermittent and ephemeral) streams oscillate between three main hydrological states: dry, disjointed standing pools, and flowing water, varying seasonally and in response to rainfall/snowmelt events (Assendelft & van Meerveldt, 2019; Buttle et al., 2012). The timing of these dynamics is of great importance to the fields of ecology and biochemistry (Leigh et al., 2016; McClain et al., 2003; Meyer et al., 2007); for example, temporary channels and their associated vegetative communities play a key role as migration corridors (Goodrich et al., 2018). Therefore, disconnected segments may well reflect vital ecological processes. Additionally, temporary stream dynamics reflect subsurface storage patterns (Camporese et al., 2010), affect downstream discharge flow rates (Godsey & Kirchner, 2014), along with water quality and nutrient availability (Romaní et al., 2006). Our model demonstrated its capacity to capture this temporal variation of intermittent channels, as is evident in Basin 4 (Figure 14).

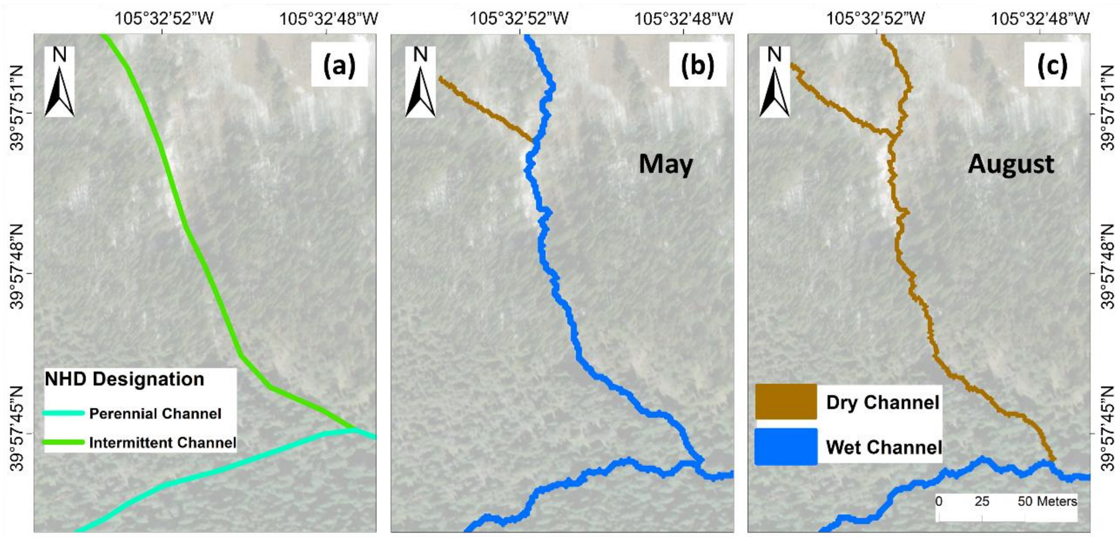

Figure 14. National Hydrography Dataset channel designations (a) compared with classified maps (b and c) depicting an intermittent channel along North Beaver Creek in Basin 4. The channel flows with snowmelt runoff in May as the catchment approaches peak flows (b) then runs dry as the drainage nears baseflow (c).

Calculations of intensity reduction in wet channels relative to dry channels were in good agreement with comparable literature. Across drainage basins, we observed an average reduction in median Normalized Intensity of 41.7% under dense vegetation ($I_{D,V,Wet}^{nor}$ versus $I_{D,V,Dry}^{nor}$) and 72.2% for areas with no dense vegetation ($I_{D,NV,Wet}^{nor}$ versus $I_{D,NV,Dry}^{nor}$). A smaller reduction in the presence of dense vegetation is to be expected due to the multiple return phenomenon described in Section 1. Song et al. (2002) demonstrated a reduction of about 80%, while Antonarakis et al. (2008) calculated an optimal threshold for water delineation that was 44% less than that of grass and low vegetation. Lang and McCarty (2009), who did encounter considerable canopy cover in portions of their study area while mapping inundated wetlands, recorded a mean reduction of 85%, although they did not segregate pixels under dense vegetation in the manner presented here. Hooshyar et al. (2015), who also focused on narrow channels but masked dense vegetation, found reductions in optimal delineation thresholds between wet and dry channels from 46% to 73%. Therefore, the intensity results calculated here were within the range of values recorded in similar studies. The work presented here represents one of the most thorough statistical analyses of wet channel intensity reduction to date, and to our knowledge the only study to demonstrate this phenomenon strictly under dense vegetation.

Accuracy of the classified data products were also comparable to that of related research efforts. Classification accuracy in the work presented here ranged from 87.3% to 95.3%, with a median of 93.0%. Seminal studies focusing on large, continuous water bodies without significant vegetation often boast accuracy upwards of 95% (Antonarakis et al., 2008; Hofle et al., 2009), although the problem of headwater channels under dense vegetation is undoubtedly more difficult. Lang and McCarty (2009) performed, to our knowledge, the only other study that focused on intensity signatures under vegetation to map water bodies, though the inundated wetlands in their research area were also large and continuous. They present an impressive accuracy of nearly 97%, however they note a considerable decline in the presence of evergreen forests, such as those in our study sites. Hooshyar et al. (2015), with whom we shared Basins 1 – 3, did not perform an accuracy assessment and instead centered their validation on correlating estimates of wetted length to streamflow at downstream gages. Qualitatively our results appear comparable. That said, when comparing modeled wet stream channel lengths to NHD reference data (Figure 10) for Basins 1 – 3, our results demonstrate a mean error of 7.8% while Hooshyar et al. (2015) register 34.7%. In other words, the classified map products presented here are in better agreement with NHD-designated perennial channel lengths during low-flow conditions, the only streams assumed to be flowing as the basins neared baseflow. This improvement indicates the benefit of considering intensity data under dense vegetation rather than excluding it. Therefore, the accuracy of our workflow is on par with similar studies focused only on larger and/or open water bodies, despite confronting the important challenge of small, headwater streams under vegetation.

6 Conclusion and Future Work

We have demonstrated that water under dense vegetation exhibits a distinctly reduced lidar intensity signature compared to dry channels. We then utilized this phenomenon to classify stream networks as wet or dry despite occlusive vegetation. While our findings can be operationalized in their current form, the need for a more complete study is evident. Future studies should combine lidar data acquisition coincident with ground-based data collection to afford more robust reference data than the NHD. Reference data could be acquired manually, or with a ground-based monitoring system like that of Assendelft and van Meerveldt (2019). This would provide an opportunity to better evaluate potential limitations, such as channel width or water depth. Further, the consistency of the observed intensity reduction across varying landscapes should be investigated. In new locations, the statistical significance of differences in wet channel and dry channel intensity under vegetation should be confirmed, as in the work presented here. Geographic factors such as topography, type and extent of vegetation, ground cover, climate, and other parameters will certainly influence the intensity distributions and shift optimal classification thresholds. As lidar data acquisition continues to decrease in expense and increase in frequency, this methodology can be used as a means of remotely mapping occluded wet channel networks and observing temporal dynamics, particularly in challenging landscapes where field observations are difficult. Pairing lidar sensors with unmanned aerial systems, like that of Spence and Mengistu (2015), will assist in achieving sufficient spatial and temporal resolution for the broader implementation of our method.

7 Appendix

Table A1. Summary of lidar collection statistics.

Dataset	Scanner	Wavelength (nm)	Scan Angle	Flight Altitude (m)	Point Density (points/m^2)	Pulse Rate (kHz)
Lake Tahoe Basin	Leica ALS50	1064	14°	900 - 1300	13.20	83 - 106
Boulder Creek Critical Zone Observatory	Optech Gemini	1047	14°	600	10.28	33 - 167

8 Acknowledgments

The authors would like to thank Dr. Kathryn Plymesser, Dr. Scott Powell, and three anonymous reviewers for their invaluable commentary and insights.

9 Open Research

Lidar data used in this study were acquired from the public Open Topography repository (https://www.opentopography.org/). NHD data were accessed through the USGS via The National Map interface (https://www.usgs.gov/the-national-map-data-delivery), while stream gage data for Basins 1 – 3 were also procured from the USGS (https://waterdata.usgs.gov/nwis/rt). Stream gage data for Basin 4 was provided by the Colorado Department of Water Resources (https://dwr.state.co.us/tools/stations). Last, precipitation data were acquired from the National Climate Data Center (https://www.ncdc.noaa.gov/cdo-web/search).

10 References

Abalharth, M., Hassan, M. A., Klinkenberg, B., Leung, V., & McCleary, R. (2015). Using LiDAR to characterize logjams in lowland rivers. *Geomorphology*, *246*, 531-541.

Antonarakis, A. S., Richards, K. S., & Brasington, J. (2008). Object-based land cover classification using airborne LiDAR. *Remote Sensing of environment*, *112*(6), 2988-2998.

Assendelft, R. S., & van Meerveld, H. J. (2019). A low-cost, multi-sensor system to monitor temporary stream dynamics in mountainous headwater catchments. *Sensors*, *19*(21), 4645.

Brzank, A., Heipke, C., Goepfert, J., & Soergel, U. (2008). Aspects of generating precise digital terrain models in the Wadden Sea from lidar–water classification and structure line extraction. *ISPRS Journal of Photogrammetry and Remote Sensing*, *63*(5), 510-528.

Buttle, J. M., Boon, S., Peters, D. L., Spence, C., Van Meerveld, H. J., & Whitfield, P. H. (2012). An overview of temporary stream hydrology in Canada. *Canadian Water Resources Journal/Revue Canadienne Des Ressources Hydriques*, *37*(4), 279-310.

Camporese, M., Paniconi, C., Putti, M., & Orlandini, S. (2010). Surface-subsurface flow modeling with path-based runoff routing, boundary condition-based coupling, and assimilation of multisource observation data. *Water Resources Research*, *46*(2).

Colvin, R., Giannico, G. R., Li, J., Boyer, K. L., & Gerth, W. J. (2009). Fish use of intermittent watercourses draining agricultural lands in the Upper Willamette River Valley, Oregon. *Transactions of the American Fisheries Society*, *138*(6), 1302-1313.

Colvin, S. A., Sullivan, S. M. P., Shirey, P. D., Colvin, R. W., Winemiller, K. O., Hughes, R. M., ... & Eby, L. (2019). Headwater streams and wetlands are critical for sustaining fish, fisheries, and ecosystem services. *Fisheries*, *44*(2), 73-91.

Devito, K., Creed, I., Gan, T., Mendoza, C., Petrone, R., Silins, U., & Smerdon, B. (2005). A framework for broad-scale classification of hydrologic response units on the Boreal Plain: is topography the last thing to consider?. *Hydrological Processes: An International Journal*, *19*(8), 1705-1714.

Erman, D. C., & Leidy, G. R. (1975). Downstream movement of rainbow trout fry in a tributary Sagehen Creek, under permanent and intermittent flow. *Transactions of the American Fisheries Society*, *104*(3), 467-473.

Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote sensing of environment*, *80*(1), 185-201.

Foody, G. M. (2009). Sample size determination for image classification accuracy assessment and comparison. *International Journal of Remote Sensing*, *30*(20), 5273-5291.

Godsey, S. E., & Kirchner, J. W. (2014). Dynamic, discontinuous stream networks: hydrologically driven variations in active drainage density, flowing channels and stream order. *Hydrological Processes*, *28*(23), 5791-5803.

Goodrich, D. C., Kepner, W. G., Levick, L. R., & Wigington Jr, P. J. (2018). Southwestern intermittent and ephemeral stream connectivity. *JAWRA Journal of the American Water Resources Association*, *54*(2), 400-422.

Hack, J. T., & Goodlett, J. C. (1960). *Geomorphology and forest ecology of a mountain region in the central Appalachians* (No. 347). United States Government Printing Office.

Hafen, K. C., Blasch, K. W., Rea, A., Sando, R., & Gessler, P. E. (2020). The influence of climate variability on the accuracy of NHD perennial and nonperennial stream classifications. *JAWRA Journal of the American Water Resources Association*, *56*(5), 903-916.

Haigh, M. J., Singh, R. B., & Krecek, J. (1998). Headwater control: matters arising. *Headwaters: water resources and soil conservation. Edited by MJ Haigh, J. Krecek, GS Rajwar, and MP Kilmartin. AA Balkema, Rotterdam, Netherlands*, *3*, 24.

Hartman, G. F., & Brown, T. G. (1987). Use of small, temporary, floodplain tributaries by juvenile salmonids in a west coast rain-forest drainage basin, Carnation Creek, British Columbia. *Canadian Journal of Fisheries and Aquatic Sciences*, *44*(2), 262-270.

Hay, A. M. (1979). Sampling designs to test land-use map accuracy. *Photogrammetric Engineering and Remote Sensing*, *45*(4), 529-533.

Höfle, B., Vetter, M., Pfeifer, N., Mandlburger, G., & Stötter, J. (2009). Water surface mapping from airborne laser scanning using signal intensity and elevation data. *Earth Surface Processes and Landforms*, *34*(12), 1635-1649.

Hooshyar, M., Kim, S., Wang, D., & Medeiros, S. C. (2015). Wet channel network extraction by integrating LiDAR intensity and elevation data. *Water Resources Research*, *51*(12), 10029-10046.

Horton, R. E. (1945). Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology. *Geological society of America bulletin*, *56*(3), 275-370.

Jacobson, P. M. (2004). *Connecticut River ecological study (1965-1973) revisited*. American Fisheries Society.

James, L. A., & Hunt, K. J. (2010). The LiDAR-side of headwater streams: mapping channel networks with high-resolution topographic data. *southeastern geographer*, *50*(4), 523-539.

James, L. A., Watson, D. G., & Hansen, W. F. (2007). Using LiDAR data to map gullies and headwater streams under forest canopy: South Carolina, USA. *Catena*, *71*(1), 132-144.

Lang, M. W., & McCarty, G. W. (2009). Lidar intensity for improved detection of inundation below the forest canopy. *Wetlands*, *29*(4), 1166-1178.

Larned, S. T., Datry, T., Arscott, D. B., & Tockner, K. (2010). Emerging concepts in temporary-river ecology. *Freshwater Biology*, *55*(4), 717-738.

Lashermes, B., Foufoula-Georgiou, E., & Dietrich, W. E. (2007). Channel network extraction from high resolution topography using wavelets. *Geophysical Research Letters*, *34*(23).

Leigh, C., Boulton, A. J., Courtwright, J. L., Fritz, K., May, C. L., Walker, R. H., & Datry, T. (2016). Ecological research and management of intermittent rivers: an historical review and future directions. *Freshwater Biology*, *61*(8), 1181-1199.

Leopold, L.B., 1994. A View of the River. Harvard University Press, Cambridge, Massachusetts, USA.

Leopold, L. B., & Miller, J. P. (1956). *Ephemeral streams: Hydraulic factors and their relation to the drainage net* (Vol. 282). US Government Printing Office.

Leopold, L.B., M.G. Wolman and J.P. Miller, 1964. Fluvial Processes in Geomorphology. Dover Publications, New York, New York, USA.

McClain, M. E., Boyer, E. W., Dent, C. L., Gergel, S. E., Grimm, N. B., Groffman, P. M., ... & Pinay, G. (2003). Biogeochemical hot spots and hot moments at the interface of terrestrial and aquatic ecosystems. *Ecosystems*, 301-312.

McDonough, O. T., Hosen, J. D., & Palmer, M. A. (2011). Temporary streams: the hydrology, geography, and ecology of non-perennially flowing waters. *River ecosystems: dynamics, management and conservation*, 259-290.

McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International journal of remote sensing*, *17*(7), 1425-1432.

Meyer, J. L., & Wallace, J. B. (2001). Lost linkages and lotic ecology: rediscovering small streams. In *Ecology: achievement and challenge: the 41st Symposium of the British Ecological Society sponsored by the Ecological Society of America held at Orlando, Florida, USA, 10-13 April 2000* (pp. 295-317). Blackwell Science.

Meyer, J. L., Strayer, D. L., Wallace, J. B., Eggert, S. L., Helfman, G. S., & Leonard, N. E. (2007). The contribution of headwater streams to biodiversity in river networks 1. *JAWRA Journal of the American Water Resources Association*, *43*(1), 86-103.

Montgomery, D. R. (1999). Process domains and the river continuum 1. *JAWRA Journal of the American Water Resources Association*, *35*(2), 397-410.

Nadeau, T. L., & Rains, M. C. (2007). Hydrological connectivity between headwater streams and downstream waters: how science can inform policy 1. *JAWRA Journal of the American Water Resources Association*, *43*(1), 118-133.

Omran, A., Dietrich, S., Abouelmagd, A., & Michael, M. (2016). New ArcGIS tools developed for stream network extraction and basin delineations using Python and java script. *Computers & Geosciences*, *94*, 140-149.

Orlandini, S., Tarolli, P., Moretti, G., & Dalla Fontana, G. (2011). On the prediction of channel heads in a complex alpine terrain using gridded elevation data. *Water Resources Research*, *47*(2).

Passalacqua, P., Do Trung, T., Foufoula-Georgiou, E., Sapiro, G., & Dietrich, W. E. (2010). A geometric framework for channel network extraction from lidar: Nonlinear diffusion and geodesic paths. *Journal of Geophysical Research: Earth Surface*, *115*(F1).

Romaní, A. M., Vázquez, E., & Butturini, A. (2006). Microbial availability and size fractionation of dissolved organic carbon after drought in an intermittent stream: biogeochemical link across the stream–riparian interface. *Microbial ecology*, *52*(3), 501-512.

Simley, J. (2006). The National Hydrography Dataset: Introduction. *Water Resources IMPACT*, *8*(2), 4-4.

Song, J. H., Han, S. H., Yu, K. Y., & Kim, Y. I. (2002). Assessing the possibility of land-cover classification using lidar intensity data. *International archives of photogrammetry remote sensing and spatial information sciences*, *34*(3/B), 259-262.

Spence, C., & Mengistu, S. (2016). Deployment of an unmanned aerial system to assist in mapping an intermittent stream. *Hydrological Processes*, *30*(3), 493-500.

Stehr, W. C., & Branson, J. W. (1938). An ecological study of an intermittent stream. *Ecology*, 294-310.

Story, M., & Congalton, R. G. (1986). Accuracy assessment: a user's perspective. *Photogrammetric Engineering and remote sensing*, *52*(3), 397-399.

Uys, M. C., & O'keeffe, J. H. (1997). Simple words and fuzzy zones: early directions for temporary river research in South Africa. *Environmental Management*, *21*(4), 517-531.

Wipfli, M. S., Richardson, J. S., & Naiman, R. J. (2007). Ecological linkages between headwaters and downstream ecosystems: Transport of organic matter, invertebrates, and wood down headwater channels 1. *JAWRA Journal of the American Water Resources Association*, *43*(1), 72-85.

Work, E. A., & Gilmer, D. S. (1976). Utilization of satellite data for inventorying prairie ponds and lakes. *Photogrammetric Engineering and Remote Sensing*, *42*(5), 685-694.

Worstell, B. B., Poppenga, S., Evans, G. A., & Prince, S. (2014). Lidar point density analysis: implications for identifying water bodies. *US Geological Survey: Reston, VA, USA*.

Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International journal of remote sensing*, *27*(14), 3025-3033.

CHAPTER SEVEN

CONCLUSION

Snow occupies a large portion of the Earth's surface and represents an imperative hydrological resource, a severe hazard, and a key component in our planetary energy budget. As the most reflective natural surface on earth, snow exerts control on Earth's radiative energy balance, with a direct effect on global and regional climate. Further, seasonal snowmelt constitutes a large natural reservoir, supplying one-sixth of the global population and 75% of the western United States with fresh water. The timing and magnitude of snowmelt accumulating in headwaters is critical to ecological health, water quality, downstream water supply, and watershed connectivity. Last, snow presents a serious hazard to both human life and infrastructure, namely in the form of avalanches. Hundreds of lives have been lost to snow avalanches in the past decade, and transportation corridors continue to extend further and further into mountainous regions. Therefore, accurate measurements of physical snowpack properties and snowmelt runoff are integral to climate models, headwater hydrology, and avalanche forecasts. This is particularly true in an era of rapid change for the cryosphere, as climate change is driving a decrease in annual snowfall, earlier snowmelt, and melting in the polar regions. However, much of the Earth's snow cover exists in remote and isolated terrain, making measurements of snow and snowmelt difficult. Further, snowpack properties, stability, and runoff are highly variable in both space and time, and thus extrapolation is often not viable.

Advances in remote sensing technology offer a revolutionized means of snowpack observation relative to traditional discrete methods. As the era of cryosphere remote sensing flourishes, researchers race to create methodologies to make optimal use of troves of data. In the work presented here, we join in these efforts by devising numerous novel mapping methodologies, aimed at expanding the capacity of climate modeling, avalanche forecasting, and hydrologic practitioners. Our hypothesis was that underutilized instruments and spatial analysis methods may address limitations in current efforts, with a particular focus on lidar and hyperspectral imaging. Using laboratory, terrestrial, and airborne techniques, we executed a three-prong approach to better quantify snow grain size, delineate specific avalanche hazards, and track streamflow in dense forests.

Regarding climate modeling, the optical snow grain size is an all-important parameter. Grain size is the primary controller of snow broadband albedo. In the NIR spectral regions, reflectance is sensitive to snow microstructure, and thus reflectance measurements can be used to “retrieve” snow grain size by matching measured values to simulations from a radiative transfer model. While this procedure has been executed for decades, differing models and retrieval techniques can produce dramatically differing results, with no clear standard method. To address this, we performed a thorough laboratory intercomparison against micro-CT benchmarks to determine optimal methods. We found that performance varied significantly between radiative transfer models, and, to a lesser extent, between retrieval techniques, with hyperspectral methods expectedly outperforming those using fewer spectral data. However, the latter result was not as pronounced as hypothesized, which provides optimism for applications that lack hyperspectral capacity. Further, we quantified tremendous uncertainty related to modeled snow grain shape, and demonstrate an optimization analysis. Last, we performed a novel validation of using lidar for grain size retrieval, a significant result given that lidar holds notable advantages over traditional passive detectors.

Towards improved avalanche forecasting, we introduced two new methodologies. First, in a laboratory study, we leveraged the oft-overlooked concept of reflectance texture to quantify an optical signature specific to surface hoar, a prominent avalanche weak layer responsible for over one third of large avalanches. The physical nature of surface hoar crystals produces large localized deviation in NIR reflectance, which can be quantified in a focal analysis to produce pixelwise maps of texture and, through a classification algorithm, maps delineating the presence versus absence or surface hoar. Next, we took to the field with a long-range terrestrial lidar to document the formation and failure of a large glide avalanche in Glacier National Park. Using lidar data, we performed a mass balance analysis of the event, and used our data to tune an avalanche dynamics model, re-creating the slide in a simulated environment. Finally, using this tuned model, we determined the critical release area depth necessary to pose a threat to the transportation corridor beneath the avalanche path.

From a hydrological perspective, we addressed the knowledge gap of stream channel mapping in forested regions. For years, aerial NIR imagery and lidar have been used to map water bodies and stream channels, utilizing the high NIR absorptivity of water. However, these methods have overlooked forested regions, where streams and their reflectance signatures are occluded from aerial view. This is a substantial oversight, as a large portion of small waterways reside in densely forested areas. We found that, due to the high pulse rate of modern-day lidar technology, lidar units can effectively “see” beneath forest canopy, and that the low-intensity signature of water remains intact. Using this finding, we expanded existing stream mapping techniques to include regions occluded by canopy cover, allowing for continuous mapping both in open and forested areas.

In totality, we have created a suite of methods to benefit practitioners across the cryosphere, further demonstrating the utility of remote sensing in a field with immense variability where discrete measurements are historically prevalent. Further, we have continued to establish that work in controlled laboratory environments can help to narrow in on novel phenomena and rigorously quantify uncertainty towards support of efforts in the field. However, additional work is always required. Future efforts should focus on scaling our findings to broader applications. For example, although promising, our grain size and surface hoar mapping efforts from the laboratory will face a much wider variety of illumination conditions and confounding factors in the field, such as variable solar zenith angles, mixtures of direct and diffuse illumination, and light-absorbing impurities. A small, well-controlled, slope-scale expansion would be a logical next step. A similar point can be made for our glide avalanche and headwater channel findings. Our glide avalanche measurement/modeling synthesis could easily be expanded to several paths along the Going-to-the-Sun road for better validation, while the resilience of our stream channel mapping algorithm might be tested on a wider variety of vegetation with enhanced ground truth. Still, our work represents numerous valuable foundations and a substantial contribution to the science of the cryosphere.

REFERENCES CITED

Abalharth, M., Hassan, M. A., Klinkenberg, B., Leung, V., & McCleary, R. (2015). Using LiDAR to characterize logjams in lowland rivers. *Geomorphology*, *246*, 531-541.

Abe, O., & Kosugi, K. (2019). Twenty-year operation of the Cryospheric Environment Simulator. Bulletin of Glaciological Research, 37, 53-65.

Ackroyd, C. S. (2023). *Airborne Lidar Intensity Correction for Mapping Snow Cover Extent and Grain Size in Mountainous Terrain* (Doctoral dissertation, The University of Utah).

Antonarakis, A. S., Richards, K. S., & Brasington, J. (2008). Object-based land cover classification using airborne LiDAR. *Remote Sensing of environment*, *112*(6), 2988-2998.

Assendelft, R. S., & van Meerveld, H. J. (2019). A low-cost, multi-sensor system to monitor temporary stream dynamics in mountainous headwater catchments. *Sensors*, *19*(21), 4645.

Bair, E. H., Stillinger, T., & Dozier, J. (2020). Snow Property Inversion from Remote Sensing (SPIReS): A generalized multispectral unmixing approach with examples from MODIS and Landsat 8 OLI. *IEEE Transactions on Geoscience and Remote Sensing*, *59*(9), 7270-7284.

Barnett, T.P., J.C. Adam and D.P. Lettenmaier 2005. Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, 438(7066): 303-309.

Bartelt, P., Buser, O., Valero, C. V., & Bühler, Y. (2016). Configurational energy and the formation of mixed flowing/powder snow and ice avalanches. *Annals of Glaciology*, *57*(71), 179-188.

Bartelt, P., Valero, C. V., Feistl, T., Christen, M., Bühler, Y., & Buser, O. (2015). Modelling cohesion in snow avalanche flow. *Journal of Glaciology*, *61*(229), 837-850.

Bhandari, A., Hamre, B., Frette, Ø., Zhao, L., Stamnes, J. J., & Kildemo, M. (2011). Bidirectional reflectance distribution function of Spectralon white reflectance standard illuminated by incoherent unpolarized and plane-polarized light. Applied Optics, 50(16), 2431-2442.

Birkeland, K. W. (1998). Terminology and predominant processes associated with the formation of weak layers of near-surface faceted crystals in the mountain snowpack. Arctic and Alpine Research, 30(2), 193-199.

Bohn, N., Painter, T. H., Thompson, D. R., Carmon, N., Susiluoto, J., Turmon, M. J., ... & Guanter, L. (2021). Optimal estimation of snow and ice surface parameters from imaging spectroscopy measurements. *Remote Sensing of Environment*, *264*, 112613.

Borgeson, L. E., & Hartman, H. (2010). Predicting wet snow avalanches at the Arapahoe basin ski area. In *Proceedings: ISSW*.

Bründl, M., & Margreth, S. (2021). Integrative risk management: the example of snow avalanches. In *Snow and ice-related hazards, risks, and disasters* (pp. 259-296). Elsevier.

Brzank, A., Heipke, C., Goepfert, J., & Soergel, U. (2008). Aspects of generating precise digital terrain models in the Wadden Sea from lidar–water classification and structure line extraction. *ISPRS Journal of Photogrammetry and Remote Sensing*, *63*(5), 510-528.

Bühler, Y., Christen, M., Kowalski, J., & Bartelt, P. (2011). Sensitivity of snow avalanche simulations to digital elevation model quality and resolution. *Annals of Glaciology*, *52*(58), 72-80.

Bühler, Y., Kumar, S., Veitinger, J., Christen, M., Stoffel, A., & Snehmani. (2013). Automated identification of potential snow avalanche release areas based on digital elevation models. *Natural Hazards and Earth System Sciences*, *13*(5), 1321-1335.

Bühler, Y., Meier, L., & Ginzler, C. (2014). Potential of operational high spatial resolution near-infrared remote sensing instruments for snow surface type mapping. *IEEE Geoscience and Remote Sensing Letters*, *12*(4), 821-825.Clark and Roush, 1984

Buttle, J. M., Boon, S., Peters, D. L., Spence, C., Van Meerveld, H. J., & Whitfield, P. H. (2012). An overview of temporary stream hydrology in Canada. *Canadian Water Resources Journal/Revue Canadienne Des Ressources Hydriques*, *37*(4), 279-310.

Camporese, M., Paniconi, C., Putti, M., & Orlandini, S. (2010). Surface-subsurface flow modeling with path-based runoff routing, boundary condition-based coupling, and assimilation of multisource observation data. *Water Resources Research*, *46*(2).

Champollion, N., Picard, G., Arnaud, L., Lefebvre, E., & Fily, M. (2013). Hoar crystal development and disappearance at Dome C, Antarctica: observation by near-infrared photography and passive microwave satellite. The Cryosphere, 7(4), 1247-1262.

Christen, M., Bartelt, P., & Kowalski, J. (2010). Back calculation of the In den Arelen avalanche with RAMMS: interpretation of model results. *Annals of Glaciology*, *51*(54), 161-168.

Christen, M., Kowalski, J., & Bartelt, P. (2010). RAMMS: Numerical simulation of dense snow avalanches in three-dimensional terrain. *Cold Regions Science and Technology*, *63*(1-2), 1-14.

Clarke, J., & McClung, D. (1999). Full-depth avalanche occurrences caused by snow gliding, Coquihalla, British Columbia, Canada. *Journal of Glaciology*, *45*(151), 539-546.

Colvin, R., Giannico, G. R., Li, J., Boyer, K. L., & Gerth, W. J. (2009). Fish use of intermittent watercourses draining agricultural lands in the Upper Willamette River Valley, Oregon. *Transactions of the American Fisheries Society*, *138*(6), 1302-1313.

Colvin, S. A., Sullivan, S. M. P., Shirey, P. D., Colvin, R. W., Winemiller, K. O., Hughes, R. M., ... & Eby, L. (2019). Headwater streams and wetlands are critical for sustaining fish, fisheries, and ecosystem services. *Fisheries*, *44*(2), 73-91.

Deems, J. S., Gadomski, P. J., Vellone, D., Evanczyk, R., LeWinter, A. L., Birkeland, K. W., & Finnegan, D. C. (2015). Mapping starting zone snow depth with a ground-based lidar to assist avalanche control and forecasting. *Cold Regions Science and Technology*, *120*, 197-204.

Deems, J. S., Painter, T. H., & Finnegan, D. C. (2013). Lidar measurement of snow depth: a review. *Journal of Glaciology*, *59*(215), 467-479.

Devito, K., Creed, I., Gan, T., Mendoza, C., Petrone, R., Silins, U., & Smerdon, B. (2005). A framework for broad-scale classification of hydrologic response units on the Boreal Plain: is topography the last thing to consider?. *Hydrological Processes: An International Journal*, *19*(8), 1705-1714.

Dillon, J., Donahue, C., Schehrer, E., Birkeland, K., & Hammonds, K. (2024). Mapping surface hoar from near-infrared texture in a laboratory. *The Cryosphere*, *18*(5), 2557-2582.

Domine, F., Taillandier, A. S., & Simpson, W. R. (2007). A parameterization of the specific surface area of seasonal snow for field use and for models of snowpack evolution. Journal of Geophysical Research: Earth Surface, 112(F2).

Donahue, C. P., Menounos, B., Viner, N., Skiles, S. M., Beffort, S., Denouden, T., ... & Heathfield, D. (2023). Bridging the gap between airborne and spaceborne imaging spectroscopy for mountain glacier surface property retrievals. *Remote Sensing of Environment*, *299*, 113849.

Donahue, C., & Hammonds, K. (2022b). Laboratory observations of preferential flow paths in snow using upward-looking polarimetric radar and hyperspectral imaging. *Remote Sensing*, *14*(10), 2297.

Donahue, C., Skiles, S. M., & Hammonds, K. (2021). In situ effective snow grain size mapping using a compact hyperspectral imager. Journal of Glaciology, 67(261), 49-57.

Donahue, C., Skiles, S. M., & Hammonds, K. (2022a). Mapping liquid water content in snow at the millimeter scale: an intercomparison of mixed-phase optical property models using hyperspectral imaging and in situ measurements. *The Cryosphere*, *16*(1), 43-59.

Dumont, M., Brissaud, O., Picard, G., Schmitt, B., Gallet, J. C., & Arnaud, Y. (2010). High-accuracy measurements of snow Bidirectional Reflectance Distribution Function at visible and NIR wavelengths–comparison with modelling results. *Atmospheric Chemistry and Physics*, *10*(5), 2507-2520.

Dumont, M., Flin, F., Malinka, A., Brissaud, O., Hagenmuller, P., Lapalus, P., ... & Ando, E. (2021). Experimental and model-based investigation of the links between snow bidirectional reflectance and snow microstructure. *The Cryosphere*, *15*(8), 3921-3948.

Eckert, N., Corona, C., Giacona, F., Gaume, J., Mayer, S., van Herwijnen, A., ... & Stoffel, M. (2024). Climate change impacts on snow avalanche activity and related risks. *Nature Reviews Earth & Environment*, 1-21.

Egli, L., Jonas, T., Grünewald, T., Schirmer, M., & Burlando, P. (2012). Dynamics of snow ablation in a small Alpine catchment observed by repeated terrestrial laser scans. *Hydrological Processes*, *26*(10), 1574-1585.

Erman, D. C., & Leidy, G. R. (1975). Downstream movement of rainbow trout fry in a tributary Sagehen Creek, under permanent and intermittent flow. *Transactions of the American Fisheries Society*, *104*(3), 467-473.

Fees, A., van Herwijnen, A., Altenbach, M., Lombardo, M., & Schweizer, J. (2023). Glide-snow avalanche characteristics at different timescales extracted from time-lapse photography. *Annals of Glaciology*, 1-12.

Fees, A., van Herwijnen, A., Lombardo, M., Schweizer, J., & Lehmann, P. (2024). Glide-snow avalanches: A mechanical, threshold-based release area model. *Natural Hazards and Earth System Sciences Discussions*, *2024*, 1-19.

Fierz, C. R. L. A., Armstrong, R. L., Durand, Y., Etchevers, P., Greene, E., McClung, D. M., ... & Sokratov, S. A. (2009). The international classification for seasonal snow on the ground.

Flanner, M. G., & Zender, C. S. (2005). Snowpack radiative heating: Influence on Tibetan Plateau climate. *Geophysical research letters*, *32*(6).

Flanner, M. G., Arnheim, J. B., Cook, J. M., Dang, C., He, C., Huang, X., ... & Zender, C. S. (2021). SNICAR-ADv3: a community tool for modeling spectral snow albedo. *Geoscientific Model Development*, *14*(12), 7673-7704.

Flanner, M. G., Liu, X., Zhou, C., Penner, J. E., & Jiao, C. (2012). Enhanced solar energy absorption by internally-mixed black carbon in snow grains. *Atmospheric Chemistry and Physics*, *12*(10), 4699-4721.

Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote sensing of environment*, *80*(1), 185-201.

Foody, G. M. (2009). Sample size determination for image classification accuracy assessment and comparison. *International Journal of Remote Sensing*, *30*(20), 5273-5291.

Gallet, J. C., Domine, F., Zender, C. S., & Picard, G. (2009). Measurement of the specific surface area of snow using infrared reflectance in an integrating sphere at 1310 and 1550 nm. *The Cryosphere*, *3*(2), 167-182.Gergely et al., 2014

Gergely, M., Wolfsperger, F., & Schneebeli, M. (2013). Simulation and validation of the InfraSnow: An instrument to measure snow optically equivalent grain size. *IEEE transactions on geoscience and remote sensing*, *52*(7), 4236-4247.

Glaus, J., Jones, K. W., Bühler, Y., Christen, M., Ruttner-Jansen, P., Gaume, J., & Bartelt, P. (2023). RAMMS:: Extended-sensitivity analysis of numerical fluidized powder avalanche simulation in three-dimensional terrain. In *International Snow Science Workshop Proceedings*.

Godsey, S. E., & Kirchner, J. W. (2014). Dynamic, discontinuous stream networks: hydrologically driven variations in active drainage density, flowing channels and stream order. *Hydrological Processes*, *28*(23), 5791-5803.

Goodrich, D. C., Kepner, W. G., Levick, L. R., & Wigington Jr, P. J. (2018). Southwestern intermittent and ephemeral stream connectivity. *JAWRA Journal of the American Water Resources Association*, *54*(2), 400-422.

Grenfell, T. C., & Warren, S. G. (1999). Representation of a nonspherical ice particle by a collection of independent spheres for scattering and absorption of radiation. *Journal of Geophysical Research: Atmospheres*, *104*(D24), 31697-31709.

Hack, J. T., & Goodlett, J. C. (1960). *Geomorphology and forest ecology of a mountain region in the central Appalachians* (No. 347). United States Government Printing Office.

Hafen, K. C., Blasch, K. W., Rea, A., Sando, R., & Gessler, P. E. (2020). The influence of climate variability on the accuracy of NHD perennial and nonperennial stream classifications. *JAWRA Journal of the American Water Resources Association*, *56*(5), 903-916.

Haigh, M. J., Singh, R. B., & Krecek, J. (1998). Headwater control: matters arising. *Headwaters: water resources and soil conservation. Edited by MJ Haigh, J. Krecek, GS Rajwar, and MP Kilmartin. AA Balkema, Rotterdam, Netherlands*, *3*, 24.

Hartman, G. F., & Brown, T. G. (1987). Use of small, temporary, floodplain tributaries by juvenile salmonids in a west coast rain-forest drainage basin, Carnation Creek, British Columbia. *Canadian Journal of Fisheries and Aquatic Sciences*, *44*(2), 262-270.

Hay, A. M. (1979). Sampling designs to test land-use map accuracy. *Photogrammetric Engineering and Remote Sensing*, *45*(4), 529-533.

He, C., Takano, Y., Liou, K. N., Yang, P., Li, Q., & Chen, F. (2017). Impact of snow grain shape and black carbon–snow internal mixing on snow optical properties: Parameterizations for climate models. *Journal of Climate*, *30*(24), 10019-10036.

Höfle, B., Vetter, M., Pfeifer, N., Mandlburger, G., & Stötter, J. (2009). Water surface mapping from airborne laser scanning using signal intensity and elevation data. *Earth Surface Processes and Landforms*, *34*(12), 1635-1649.

Hooshyar, M., Kim, S., Wang, D., & Medeiros, S. C. (2015). Wet channel network extraction by integrating LiDAR intensity and elevation data. *Water Resources Research*, *51*(12), 10029-10046.

Horton, R. E. (1945). Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology. *Geological society of America bulletin*, *56*(3), 275-370.

Horton, S., & Jamieson, B. (2017). Spectral measurements of surface hoar crystals. Journal of Glaciology, 63(239), 477-486.

Jacobson, P. M. (2004). *Connecticut River ecological study (1965-1973) revisited*. American Fisheries Society.

James, L. A., & Hunt, K. J. (2010). The LiDAR-side of headwater streams: mapping channel networks with high-resolution topographic data. *southeastern geographer*, *50*(4), 523-539.

James, L. A., Watson, D. G., & Hansen, W. F. (2007). Using LiDAR data to map gullies and headwater streams under forest canopy: South Carolina, USA. *Catena*, *71*(1), 132-144.

Jamieson, J. B., & Schweizer, J. (2000). Texture and strength changes of buried surface-hoar layers with implications for dry snow-slab avalanche release. Journal of Glaciology, 46(152), 151-160.

Jones, A. (2004). Review of glide processes and glide avalanche release. *Avalanche News*, *69*(7), 53-60.

Kokhanovsky, A. A., & Zege, E. P. (2004). Scattering optics of snow. *Applied optics*, *43*(7), 1589-1602.

Lang, M. W., & McCarty, G. W. (2009). Lidar intensity for improved detection of inundation below the forest canopy. *Wetlands*, *29*(4), 1166-1178.

Larned, S. T., Datry, T., Arscott, D. B., & Tockner, K. (2010). Emerging concepts in temporary-river ecology. *Freshwater Biology*, *55*(4), 717-738.

Lashermes, B., Foufoula-Georgiou, E., & Dietrich, W. E. (2007). Channel network extraction from high resolution topography using wavelets. *Geophysical Research Letters*, *34*(23).

Legagneux, L., Cabanes, A., & Dominé, F. (2002). Measurement of the specific surface area of 176 snow samples using methane adsorption at 77 K. *Journal of Geophysical Research: Atmospheres*, *107*(D17), ACH-5.

Leigh, C., Boulton, A. J., Courtwright, J. L., Fritz, K., May, C. L., Walker, R. H., & Datry, T. (2016). Ecological research and management of intermittent rivers: an historical review and future directions. *Freshwater Biology*, *61*(8), 1181-1199.

Leopold, L. B., & Miller, J. P. (1956). *Ephemeral streams: Hydraulic factors and their relation to the drainage net* (Vol. 282). US Government Printing Office.

Leopold, L.B., 1994. A View of the River. Harvard University Press, Cambridge, Massachusetts, USA.

Leopold, L.B., M.G. Wolman and J.P. Miller, 1964. Fluvial Processes in Geomorphology. Dover Publications, New York, New York, USA.

Li, D., Zhao, M. H., Garra, J., Kolpak, A. M., Rappe, A. M., Bonnell, D. A., & Vohs, J. M. (2008). Direct in situ determination of the polarization dependence of physisorption on ferroelectric surfaces. *nature materials*, *7*(6), 473-477.

Libois, Q., Picard, G., Dumont, M., Arnaud, L., Sergent, C., Pougatch, E., ... & Vial, D. (2014). Experimental determination of the absorption enhancement parameter of snow. *Journal of Glaciology*, *60*(222), 714-724.

Libois, Q., Picard, G., France, J. L., Arnaud, L., Dumont, M., Carmagnola, C. M., & King, M. D. (2013). Influence of grain shape on light penetration in snow. *The Cryosphere*, *7*(6), 1803-1818.

Lorensen, W. E., & Cline, H. E. (1987). Marching cubes: A high resolution 3D surface construction algorithm. *ACM siggraph computer graphics*, *21*(4), 163-169.

Lutz, E. R., & Birkeland, K. W. (2011). Spatial patterns of surface hoar properties and incoming radiation on an inclined forest opening. Journal of Glaciology, 57(202), 355-366.

Maggioni, M., & Gruber, U. (2003). The influence of topographic parameters on avalanche release dimension and frequency. *Cold Regions Science and Technology*, *37*(3), 407-419.

Maggioni, M., Freppaz, M., Ceaglio, E., Godone, D., Viglietti, D., Zanini, E., ... & Pallara, O. (2013). A new experimental snow avalanche test site at Seehore peak in Aosta Valley (NW Italian Alps)—part I: conception and logistics. *Cold Regions Science and Technology*, *85*, 175-182.

Maggioni, M., Freppaz, M., Christen, M., Bartelt, P., and Zanini, E.: Back-calculation of small avalanches with the 2D avalanche dynamics model RAMMS: four events artificially triggered at the Seehore test site in Aosta Valley (NW-Italy), in: Proceedings of the International Snow Science Workshop, 16–21 September 2012, Anchorage, Alaska, pp. 591–598, 2012.

Malinka, A. (2023). Stereological approach to radiative transfer in porous materials. Application to the optics of snow. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *295*, 108410.

Malinka, A. V. (2014). Light scattering in porous materials: Geometrical optics and stereological approach. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *141*, 14-23.

Malinka, A., Zege, E., Heygster, G., & Istomina, L. (2016). Reflective properties of white sea ice and snow. *The Cryosphere*, *10*(6), 2541-2557.

Margreth, S., & Ammann, W. J. (2004). Hazard scenarios for avalanche actions on bridges. *Annals of Glaciology*, *38*, 89-96.

Marks, D., & Dozier, J. (1992). Climate and energy exchange at the snow surface in the alpine region of the Sierra Nevada: 2. Snow cover energy balance. *Water Resources Research*, *28*(11), 3043-3054.

Matzl, M., & Schneebeli, M. (2006). Measuring specific surface area of snow by near-infrared photography. *Journal of Glaciology*, *52*(179), 558-564.

Matzl, M., & Schneebeli, M. (2010). Stereological measurement of the specific surface area of seasonal snow types: Comparison to other methods, and implications for mm-scale vertical profiling. Cold Regions Science and Technology, 64(1), 1-8.

Mayer, S., Hendrick, M., Michel, A., Richter, B., Schweizer, J., Wernli, H., & van Herwijnen, A. (2024). Changes in snow avalanche activity in response to climate warming in the Swiss Alps. *EGUsphere*, *2024*, 1-32.

McClain, M. E., Boyer, E. W., Dent, C. L., Gergel, S. E., Grimm, N. B., Groffman, P. M., ... & Pinay, G. (2003). Biogeochemical hot spots and hot moments at the interface of terrestrial and aquatic ecosystems. *Ecosystems*, 301-312.

McClung, D., & Schaerer, P. A. (2006). *The avalanche handbook*. The Mountaineers Books.

McDonough, O. T., Hosen, J. D., & Palmer, M. A. (2011). Temporary streams: the hydrology, geography, and ecology of non-perennially flowing waters. *River ecosystems: dynamics, management and conservation*, 259-290.

McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International journal of remote sensing*, *17*(7), 1425-1432.

Meirold-Mautner, I., & Lehning, M. (2004). Measurements and model calculations of the solar shortwave fluxes in snow on Summit, Greenland. *Annals of Glaciology*, *38*, 279-284.

Meyer, J. L., & Wallace, J. B. (2001). Lost linkages and lotic ecology: rediscovering small streams. In *Ecology: achievement and challenge: the 41st Symposium of the British Ecological Society sponsored by the Ecological Society of America held at Orlando, Florida, USA, 10-13 April 2000* (pp. 295-317). Blackwell Science.

Meyer, J. L., Strayer, D. L., Wallace, J. B., Eggert, S. L., Helfman, G. S., & Leonard, N. E. (2007). The contribution of headwater streams to biodiversity in river networks 1. *JAWRA Journal of the American Water Resources Association*, *43*(1), 86-103.

Miller, A. D., Redpath, T. A., Sirguey, P., Cox, S. C., Bartelt, P., Bogie, D., ... & Bühler, Y. (2023). Unprecedented winter rainfall initiates large snow avalanche and mass movement cycle in New Zealand's Southern Alps/Kā Tiritiri o te Moana. *Geophysical Research Letters*, *50*(8), e2022GL102105.

Miller, A., Sirguey, P., Morris, S., Bartelt, P., Cullen, N., Redpath, T., ... & Bühler, Y. (2022). The impact of terrain model source and resolution on snow avalanche modeling. *Natural Hazards and Earth System Sciences*, *22*(8), 2673-2701.

Mitterer, C., & Schweizer, J. (2012). Glide snow avalanches revisited. *The avalanche journal*, *101*, p68-71.

Mock, C. J., & Birkeland, K. W. (2000). Snow avalanche climatology of the western United States mountain ranges. *Bulletin of the American Meteorological Society*, *81*(10), 2367-2392.

Montgomery, D. R. (1999). Process domains and the river continuum 1. *JAWRA Journal of the American Water Resources Association*, *35*(2), 397-410.

Murphy, A. B. (2006). Modified Kubelka–Munk model for calculation of the reflectance of coatings with optically-rough surfaces. *Journal of Physics D: Applied Physics*, *39*(16), 3571.

Nadeau, T. L., & Rains, M. C. (2007). Hydrological connectivity between headwater streams and downstream waters: how science can inform policy 1. *JAWRA Journal of the American Water Resources Association*, *43*(1), 118-133.

Nolin, A. W., & Dozier, J. (2000). A hyperspectral method for remotely sensing the grain size of snow. Remote sensing of Environment, 74(2), 207-216.

Omran, A., Dietrich, S., Abouelmagd, A., & Michael, M. (2016). New ArcGIS tools developed for stream network extraction and basin delineations using Python and java script. *Computers & Geosciences*, *94*, 140-149.

Orlandini, S., Tarolli, P., Moretti, G., & Dalla Fontana, G. (2011). On the prediction of channel heads in a complex alpine terrain using gridded elevation data. *Water Resources Research*, *47*(2).

Painter, T. H., Barrett, A. P., Landry, C. C., Neff, J. C., Cassidy, M. P., Lawrence, C. R., ... & Farmer, G. L. (2007). Impact of disturbed desert soils on duration of mountain snow cover. *Geophysical Research Letters*, *34*(12).

Painter, T. H., Bryant, A. C., & Skiles, S. M. (2012). Radiative forcing by light absorbing impurities in snow from MODIS surface reflectance data. *Geophysical Research Letters*, *39*(17).

Painter, T. H., Molotch, N. P., Cassidy, M., Flanner, M., & Steffen, K. (2007). Contact spectroscopy for determination of stratigraphy of snow optical grain size. *Journal of Glaciology*, *53*(180), 121-127.

Painter, T. H., Rittger, K., McKenzie, C., Slaughter, P., Davis, R. E., & Dozier, J. (2009). Retrieval of subpixel snow covered area, grain size, and albedo from MODIS. *Remote Sensing of Environment*, *113*(4), 868-879.

Painter, T. H., Skiles, S. M., Deems, J. S., Bryant, A. C., & Landry, C. C. (2012). Dust radiative forcing in snow of the Upper Colorado River Basin: 1. A 6 year record of energy balance, radiation, and dust concentrations. *Water Resources Research*, *48*(7).

Passalacqua, P., Do Trung, T., Foufoula-Georgiou, E., Sapiro, G., & Dietrich, W. E. (2010). A geometric framework for channel network extraction from lidar: Nonlinear diffusion and geodesic paths. *Journal of Geophysical Research: Earth Surface*, *115*(F1).

Peitzsch, E. H., Hendrikx, J., & Fagre, D. B. (2012). Timing of wet snow avalanche activity: an analysis from Glacier National Park, Montana, USA. In *Proceedings of the International Snow Science Workshop, September 17–21, 2012, Anchorage, Alaska*.

Perovich, D. K., & Govoni, J. W. (1992, December). Light reflection from a sea-ice cover during the onset of summer melt. In *Ocean Optics XI* (Vol. 1750, pp. 508-516). SPIE.

Picard, G., Arnaud, L., Domine, F., & Fily, M. (2009). Determining snow specific surface area from near-infrared reflectance measurements: Numerical study of the influence of grain shape. *Cold Regions Science and Technology*, *56*(1), 10-17.

Picard, G., Libois, Q., & Arnaud, L. (2016). Refinement of the ice absorption spectrum in the visible using radiance profile measurements in Antarctic snow. *The Cryosphere*, *10*(6), 2655-2672.

Prokop, A., Schön, P., Singer, F., Pulfer, G., Naaim, M., Thibert, E., & Soruco, A. (2015). Merging terrestrial laser scanning technology with photogrammetric and total station data for the determination of avalanche modeling parameters. *Cold Regions Science and Technology*, *110*, 223-230.

Reardon, B. A., Fagre, D. B., Dundas, M., & Lundy, C. (2006, October). Natural glide slab avalanches, Glacier National Park, USA: a unique hazard and forecasting challenge. In *Proceedings of the 2006 International Snow Science Workshop, Telluride, CO, USA* (pp. 778-785).

Reardon, B., & Lundy, C. (2004, September). Forecasting for natural avalanches during spring opening of the Going-to-the-Sun Road, Glacier National Park, USA. In *Proceedings of the 2004 International Snow Science Workshop* (pp. 566-581). Jackson Wyoming, USA.

Reuter, B., Richter, B., & Schweizer, J. (2016). Snow instability patterns at the scale of a small basin. *Journal of Geophysical Research: Earth Surface*, *121*(2), 257-282.

Riegl Laser Measurement Systems, GmbH, 2015a. RiVLib. http://www.riegl.com/index.php?id=224

Robledano, A., Picard, G., Dumont, M., Flin, F., Arnaud, L., & Libois, Q. (2023). Unraveling the optical shape of snow. *Nature Communications*, *14*(1), 3955.

Romaní, A. M., Vázquez, E., & Butturini, A. (2006). Microbial availability and size fractionation of dissolved organic carbon after drought in an intermittent stream: biogeochemical link across the stream–riparian interface. *Microbial ecology*, *52*(3), 501-512.

Sassen, K. (2005). Polarization in lidar. In *Lidar* (pp. 19-42). Springer, New York, NY.

Schleef, S., Jaggi, M., Löwe, H., & Schneebeli, M. (2014). An improved machine to produce nature-identical snow in the laboratory. Journal of Glaciology, 60(219), 94-102.

Schweizer, J., Kronholm, K., Jamieson, J. B., & Birkeland, K. W. (2008). Review of spatial variability of snowpack properties and its importance for avalanche formation. *Cold Regions Science and Technology*, *51*(2-3), 253-272.

Seidel, F. C., Rittger, K., Skiles, S. M., Molotch, N. P., & Painter, T. H. (2016). Case study of spatial and temporal variability of snow cover, grain size, albedo and radiative forcing in the

Sierra Nevada and Rocky Mountain snowpack derived from imaging spectroscopy. *The Cryosphere*, *10*(3), 1229-1244.

Simenhois, R., & Birkeland, K. (2010, October). Meteorological and environmental observations from three glide avalanche cycles and the resulting hazard management technique. In *International Snow Science Workshop ISSW, Lake Tahoe CA, USA* (pp. 846-853).

Simley, J. (2006). The National Hydrography Dataset: Introduction. *Water Resources IMPACT*, *8*(2), 4-4.

Simpkins, G. 2018. Snow-related water woes. *Nature Climate Change*, 8(11): 945-945.

Skiles, S. M., Donahue, C. P., Hunsaker, A. G., & Jacobs, J. M. (2023a). UAV hyperspectral imaging for multiscale assessment of Landsat 9 snow grain size and albedo. *Frontiers in Remote Sensing*, *3*, 1038287.

Skiles, S. M., Flanner, M., Cook, J. M., Dumont, M., & Painter, T. H. (2018). Radiative forcing by light-absorbing particles in snow. *Nature Climate Change*, *8*(11), 964-971.

Skiles, S. M., Painter, T. H., Deems, J. S., Bryant, A. C., & Landry, C. C. (2012). Dust radiative forcing in snow of the Upper Colorado River Basin: 2. Interannual variability in radiative forcing and snowmelt rates. *Water Resources Research*, *48*(7).

Song, J. H., Han, S. H., Yu, K. Y., & Kim, Y. I. (2002). Assessing the possibility of land-cover classification using lidar intensity data. *International archives of photogrammetry remote sensing and spatial information sciences*, *34*(3/B), 259-262.

Sovilla, B. (2004). *Field experiments and numerical modelling of mass entrainment and deposition processes in snow avalanches* (Doctoral dissertation, ETH Zurich).

Spence, C., & Mengistu, S. (2016). Deployment of an unmanned aerial system to assist in mapping an intermittent stream. *Hydrological Processes*, *30*(3), 493-500.
Stamnes, K., Tsay, S. C., Wiscombe, W., & Jayaweera, K. (1988). Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media. *Applied optics*, *27*(12), 2502-2509.

Stanton, B., Miller, D., Adams, E., & Shaw, J. A. (2016). Bidirectional-reflectance measurements for various snow crystal morphologies. *Cold Regions Science and Technology*, *124*, 110-117.

Stehr, W. C., & Branson, J. W. (1938). An ecological study of an intermittent stream. *Ecology*, 294-310.

Stimberis, J., & Rubin, C. (2004, September). Glide avalanche detection on a smooth rock slope, Snoqualmie Pass, Washington. In *Proceedings ISSW* (pp. 608-610).

Stimberis, J., & Rubin, C. M. (2011). Glide avalanche response to an extreme rain-on-snow event, Snoqualmie Pass, Washington, USA. *Journal of Glaciology*, *57*(203), 468-474.

Stoffel, L., Bartelt, P., Margreth, S., & Schweizer, J. (2018, October). Can scenario-based avalanche dynamics calculations help in the decision making process for road closures? In *Proceedings of the International Snow Science Workshop* (pp. 07-12).

Story, M., & Congalton, R. G. (1986). Accuracy assessment: a user's perspective. *Photogrammetric Engineering and remote sensing, 52*(3), 397-399.

Toon, O. B., McKay, C. P., Ackerman, T. P., & Santhanam, K. (1989). Rapid calculation of radiative heating rates and photodissociation rates in inhomogeneous multiple scattering atmospheres. *Journal of Geophysical Research: Atmospheres*, *94*(D13), 16287-16301.Warren, 1982

Uys, M. C., & O'keeffe, J. H. (1997). Simple words and fuzzy zones: early directions for temporary river research in South Africa. *Environmental Management*, *21*(4), 517-531.

Veitinger, J., & Sovilla, B. (2016). Linking snow depth to avalanche release area size: measurements from the Vallée de la Sionne field site. *Natural Hazards and Earth System Sciences*, *16*(8), 1953-1965.

Veitinger, J., Sovilla, B., & Purves, R. S. (2014). Influence of snow depth distribution on surface roughness in alpine terrain: a multi-scale approach. *The Cryosphere*, *8*(2), 547-569.

Vera Valero, C., Wever, N., Bühler, Y., Stoffel, L., Margreth, S., & Bartelt, P. (2016). Modelling wet snow avalanche runout to assess road safety at a high-altitude mine in the central Andes. *Natural Hazards and Earth System Sciences*, *16*(11), 2303-2323.

Vera Valero, C., Wever, N., Christen, M., & Bartelt, P. (2018). Modeling the influence of snow cover temperature and water content on wet-snow avalanche runout. *Natural Hazards and Earth System Sciences*, *18*(3), 869-887.

Walter, B., Brouet, L., Jaggi, M., & Löwe, H. (2023). Automated LiDAR remote sensing for measuring the spatial and temporal evolution of surface hoar formation. In International Snow Science Workshop (ISSW), Bend, OR, USA.

Warren, S. G. (1982). Optical properties of snow. *Reviews of Geophysics*, *20*(1), 67-89.

Warren, S. G. (1984). Impurities in snow: Effects on albedo and snowmelt. *Annals of Glaciology, 5,* 177-179.

Warren, S. G. (1984). Optical constants of ice from the ultraviolet to the microwave. *Applied optics*, *23*(8), 1206-1225.

Warren, S. G., & Brandt, R. E. (2008). Optical constants of ice from the ultraviolet to the microwave: A revised compilation. *Journal of Geophysical Research: Atmospheres*, *113*(D14).

Wever, N., Vera Valero, C., & Techel, F. (2018). Coupled snow cover and avalanche dynamics simulations to evaluate wet snow avalanche activity. *Journal of Geophysical Research: Earth Surface*, *123*(8), 1772-1796.

Wipfli, M. S., Richardson, J. S., & Naiman, R. J. (2007). Ecological linkages between headwaters and downstream ecosystems: Transport of organic matter, invertebrates, and wood down headwater channels 1. *JAWRA Journal of the American Water Resources Association*, *43*(1), 72-85.

Wiscombe, W. J., & Warren, S. G. (1980). A model for the spectral albedo of snow. I: Pure snow. *Journal of Atmospheric Sciences*, *37*(12), 2712-2733.

Work, E. A., & Gilmer, D. S. (1976). Utilization of satellite data for inventorying prairie ponds and lakes. *Photogrammetric Engineering and Remote Sensing*, *42*(5), 685-694.

Worstell, B. B., Poppenga, S., Evans, G. A., & Prince, S. (2014). Lidar point density analysis: implications for identifying water bodies. *US Geological Survey: Reston, VA, USA*.

Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International journal of remote sensing*, *27*(14), 3025-3033.

Yang, Y., Marshak, A., Han, M., Palm, S. P., & Harding, D. J. (2017). Snow grain size retrieval over the polar ice sheets with the Ice, Cloud, and land Elevation Satellite (ICESat) observations. Journal of Quantitative Spectroscopy and Radiative Transfer, 188, 159-164.