



The Castle Watershed
a new water resource assessment framework for a
sentinel system at the crown of the continent

Alberta Innovates File E323726

Public Final Report

Submitted on: June 1st 2020

Prepared for:

Alberta Innovates, Brett Purdy

Prepared by:

University of Lethbridge,

Chris Hopkinson, Professor & Research Chair,

403 332 4586, c.hopkinson@uleth.ca

With contributions from Professors Stefan Kienzle and Stewart Rood, University of Lethbridge
and Professor David Sauchyn, University of Regina.

University of
Lethbridge



Alberta Innovates (AI) and Her Majesty the Queen in right of Alberta make no warranty, express or implied, nor assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information contained in this publication, nor that use thereof infringe on privately owned rights. The views and opinions of the author expressed herein do not necessarily reflect those of AI or Her Majesty the Queen in right of Alberta. The directors, officers, employees, agents and consultants of AI and the Government of Alberta are exempted, excluded and absolved from all liability for damage or injury, howsoever caused, to any person in connection with or arising out of the use by that person for any purpose of this publication or its contents.

1. TABLE OF CONTENTS

| | |
|--|-----|
| 1. Table of Contents..... | 1 |
| 2. Summary..... | 2 |
| 2.1 Introduction..... | 2 |
| 2.2 Objectives..... | 2 |
| 2.3 Results..... | 2 |
| 2.4 Primary Outcomes..... | 4 |
| 2.5 Knowledge Dissemination..... | 7 |
| 3 Appendices..... | 16 |
| A3.1 Assessing Fractional Cover in the Alpine Treeline Ecotone Using the WSL Monoplotting Tool and Airborne Lidar..... | 16 |
| A3.2 Repeat Oblique Photography shows Terrain and Fire-Exposure Controls on Century-Scale Canopy Cover Change in the Alpine Treeline Ecotone..... | 29 |
| A3.3 Simulating Historic Change in the Alpine Treeline Ecotone using a Random Forest Model | 56 |
| A3.4 Evaluating Lidar-estimated snow depth drivers and their consistency during mid winter and melt onset for a Rocky Mountain watershed | 87 |
| A 3.5 Spatial and Temporal Consistency of machine learning-based imputation of snowpack depth within an operational lidar sampling framework | 119 |
| A3.6 In-Situ LED-Based Observation of Snow Surface and Depth Transects..... | 154 |
| A3.7 Airborne LiDAR Flight Path Sampling Optimization for Watershed-wide Snowpack Depth Monitoring..... | 180 |
| A3.8 Watershed-wide Snow Density and SWE Estimation using Spatially Distributed Time Series Meteorological Observations, Airborne LiDAR, and Field Measurements..... | 182 |
| A3.9 A Lightweight Leddar Optical Fusion Scanning System (FSS) for Canopy Foliage Monitoring..... | 188 |
| A3.10 Deriving Climate Change Projections from Regional Climate Models..... | 217 |
| A3.11 ACRU water resources model setup, simulation and soil mapping for the Castle Headwaters..... | 221 |

2. SUMMARY

2.1 Introduction

Climate change and land surface disturbances are changing vegetation covers in the Oldman River Basin (ORB) Headwaters. The impact of these changes on future water supply is unknown. The project objectives were to evaluate contemporary land surface conditions, historical land cover changes within the Castle Watershed headwaters of the ORB and the combined impact of these plus climatic changes on long term river (2040 to 2070) runoff. Field data collection, remote sensing and modeling studies were conducted to map hydro-meteorological conditions, snowpack, soils and forest cover properties in this Rocky Mountain watershed and how they influence hydrology at a scale suitable for operational water resource planning.

2.2 Objectives

Appendices of articles and reports tied to these objectives shown in parentheses (A3.#).

- 1) Map forest cover and historical century-scale change in Westcastle watershed. Relate these forest cover changes to changing hydroclimatic conditions. (A3.1; 3.2; 3.3)
- 2) Map seasonal snowpack distributions across the Westcastle watershed. Quantify the spatially variable land surface attribute controls on snowpack depth and place the existing snow monitoring stations into the watershed context; i.e. do existing monitoring stations over- or under-represent headwater snowpack volume? (A3.4; 3.5; 3.6; 3.7; 3.8)
- 3) Evaluate and refine new geomatics sampling (lidar) and machine learning routines to spatially impute time-variant land surface attributes of forest cover & snowpack. (A3.3; 3.5)
- 4) Develop new and innovative low power (LED-based) lidar for in situ 1D water level, 2D snow depth transect and 3D forest canopy phenology monitoring. (A3.6; 3.9)
- 5) Develop a satellite synthetic aperture radar-based snowmelt observation and tracking routine for mountain environments. (Objective unmet)
- 6) Map soils at a resolution sufficient to parameterize a hydrological model. (A3.11)
- 7) Evaluate paleo-historical and simulate regional climate change forecasts across the Castle headwaters (A3.10)
- 8) Use above land surface and climate data to simulate water supply changes (A3.11)

See “Notable Outcomes” (2.4) for a list of public outputs related to these objectives.

2.3 Results

Almost all the above objectives were met (see section 2.4) and are summarized in detail in the journal papers listed in section 2.5 and presented in the appendices. The SAR snowmelt tracking showed promise in early stages of analysis but the associated grad student withdrew from his program mid way through his thesis in the last year of the

WIP, so this sub-project was not completed. Also, while it was initially planned that snowpack and forest cover distribution and control results would be used in the Hydrological HRU-simulations, this objective was not achieved so hydro simulations used pre-existing landcover data in combination with new high resolution soil cover maps. Consequently, there is scope for improving model simulations using the newly obtained data and knowledge on historical forest cover changes and snowpack distribution controls.

Two Ph.D. and Six MSc theses were directly supported through our WIP project. A further 17 HQP graduate students were involved either as short term research assistants (13) or as incomplete thesis students (4). Thus far the project has directly supported the publication of 23 Peer-reviewed journal (19) and proceedings papers (4), as well as 52 invited seminars, conference papers or media appearances. This project has been leveraged to provide baseline support to one NSERC Discovery Grant (Hopkinson), Two NSERC Engage Grants (Hopkinson) and one grant to the University of Lethbridge from Alberta Environment and Protection. Furthermore, the results of this project have provided the basis for a new project examining the forest regeneration and water resources implications following the Kenow Wildfire (contribution funding from Parks Canada). After the last unsuccessful continuation WIP proposal, Hopkinson has started the process of removing all the hydrometeorological infrastructure (one stream gauge and 8 meteorological stations) from Westcastle that were installed to support the WIP.

The WIP enabled new and expanded partnerships with Alberta Environment and Parks (snowpack monitoring and new Parkland planning discussions), Clean Harbors Airborne Imaging (lidar monitoring market development), Castle Mountain Resort (lidar and climate data to support ski hill development planning and day-to-day operations), Tough Country Communications (development of a remote real time weather station telemetry system used by Castle Park visitors and residents), Leddartech, Quebec (development of LED-based snowpack and forest canopy monitoring sensors).

Outputs have been disseminated via traditional meetings, seminars, conference presentations, journal articles. We have also engaged directly with our industry and government partners both in data collection and in discussing next steps in application development. Much is on hold pending funding, as resources have been and continue to be dedicated to removing infrastructure from Westcastle. We have had early discussions with Leddartech and Campbell Scientific about potential commercialization of our leddar prototype snow and canopy monitoring sensors. We continue to discuss lidar snow and vegetation monitoring opportunities with Clean Harbors Airborne Imaging, while also engaging with AEP and CCMEQ (Canada Centre for Mapping and Earth Observation) on the use of remote sensing for monitoring purposes. However, with environmental monitoring being of lower provincial priority now than it was when our WIP started, both the commercial and funding opportunities have diminished for the time being.

2.4 Notable outcomes

- 1) To support several of the research objectives (1,2,4,5,7,8 from section 2.2) for this WIP project a hydrometeorological array was established based out of our field station in the Westcastle. For four of the stations, weather and image data were collected via telemetry and uploaded to a web site in real time: <http://artemis.uleth.ca:8080/CastleMet/index.html> (See landing page graphic below). Remaining weather stations were located at different slope aspects around the Castle Mountain Resort (CMR) at treeline and in Hague cirque south of the ski hill (See previous interim WIP reports for detail on installations and data records).

| | | | | | | | |
|----------------|----------|-----------|-----------|------------|-------------|-------------|------------|
| Site Locations | CMR Cams | UofL Cams | Ridge met | Valley met | Station met | Comparisons | Historical |
|----------------|----------|-----------|-----------|------------|-------------|-------------|------------|

ARTEMIS University of Lethbridge West Castle watershed meteorological station network
Oldman River headwaters, Alberta, Canada



Weather stations installed & maintained by The ARTEMIS Laboratory, University of Lethbridge (www.uleth.ca)

Contact Dr. Chris Hopkinson: c.hopkinson@uleth.ca
www.artemis-strikingly.com

Support provided by Castle Mountain Resort (www.skicastle.ca)
Alberta Environment & Parks, Tough Country Communications (www.toughcountryinternet.ca)






- 2) A new technique was devised for integrating airborne lidar canopy cover data with oblique imagery to quantify and map contemporary and historical canopy fractional

cover in the mountainous headwaters of the Westcastle Watershed (McCaffrey and Hopkinson, 2017; [A3.1](#)). This approach utilized the Mountain Legacy Project data but could be applied to any situation where lidar and oblique imagery co-exist. This represents a useful approach for interpolating over gaps in forest cover data either in space or time, where one or the other sampling data exist. ([see objectives 1 and 3, section 2.2](#))

- 3) In applying and extending the above noted new methodology, we were able to quantify both natural and wildfire-induced forest treeline changes across much of the Westcastle watershed (McCaffrey and Hopkinson, 2020; [A3.2](#)). Of hydrological importance, it was found that treeline was least impacted by wildfire on north-facing slopes, with evidence for climatic treeline advance also on north-facing slopes. These observations suggest that treeline in this region is limited by moisture conditions, and by extension that enhanced snowmelt and evaporative losses on south-facing slopes provide limits to vegetation growth. Therefore illustrating the important feedback between forest cover, slope aspect, snowpack and evaporation that requires further study. ([see objective 1, section 2.2](#))
- 4) A machine learning approach was developed and implemented to fill the gaps in or spatially impute the contemporary and historical map of tree cover and treeline to ensure a spatially continuous and historically reasonable map of forest cover for the entire watershed (McCaffrey and Hopkinson, Accepted; [A3.3](#)). This analysis is useful as it illustrates how landcover change can be simulated from a fusion of aerial and oblique image sample data and then imputed using machine learning approaches. This is particularly useful in a situation where we need continuous spatial and historical landcover data for hydrological simulations (not implemented in this study). ([see objectives 1 and 3, section 2.2](#))
- 5) A spatially-explicit assessment of snowdepth enabled a quantification of snowdepth drivers and their consistency across the entire Westcastle watershed (Cartwright et al. In Review 1; [A3.4](#)). Elevation, aspect and localized depressions were the highest ranked controls during both accumulation and ablation phases but the landcover feature of most significant overall influence on watershed-scale snowpack volume was treeline. Combined with evidence from the above objectives, this suggests that with variations in treeline, there will also be a feedback to snowpack accumulation, melt and its timing. The feedback between long-term change in treeline, snowpack distribution and melt timing and its implication to future water resources needs to be further evaluated. ([see objective 2, section 2.2](#))
- 6) Machine learning imputation of snowpack depth and volume was evaluated using lidar and a random forest routine (Cartwright et al. In Review 2; [A3.5](#)). This proof of concept demonstrated that lidar was a viable training dataset for an operational snow volume monitoring framework. AEP and Airborne Imaging were our govt and industry partners on this project and we are still working with these partners on operational implementation. ([see objectives 2 and 3, section 2.2](#))
- 7) The WIP was supplemented by an NSERC ENGAGE grant in partnership with Leddartech (QC) to develop and test a leddar (LED lidar) sensor for water surface and snowpack depth transect monitoring. Two graduate students worked on the

water level monitoring and we had some success but, overall, the sensor was found unreliable and unstable over water surfaces due to specular reflection and insufficient signal gain control (unpublished results). However, the snowpack experiments and sensor development were highly successful (Parsons and Hopkinson, 2016; Barnes et al. 2020; A3.6), with our customized leddar sensor prototype operating successfully at our Westcastle field site for two full winters. We have recently added telemetry capability for remote control and monitoring. Initial discussions were held with Campbell Scientific about further testing and development. With the end of our WIP funding, we cannot pursue further refinement or commercialization. (see objective 4, section 2.2)

- 8) Following outputs #5 and #6 above, we started to explore operational lidar flight path optimization for snowpack depth sampling and machine learning imputation (proposed in unsuccessful follow up WIP). The purpose being to dramatically reduce the operational cost of lidar acquisition by reducing aerial cover to a ~10% areal sample or less. If successful, this would mean an entire watershed of snow depth could be sampled in less than 30 minutes instead of 3 to 4 hours, or it would mean the entire ORB headwaters could be sampled and modeled within one single flight, which would dramatically reduce aerial monitoring costs (see abstract in A3.7). (see objectives 2 and 3, section 2.2)
- 9) Building on output #5 and #6 above we started to develop a full SWE model from the watershed-level snow depth and spatialized meteorological data collected at varying elevations and slope aspects within Westcastle (an objective in the last WIP proposal). The benefit of this modeling (if completed and successful) was that it could be combined with the lidar sampling and imputation routine in #8 to create a cost-effective spatially explicit full headwater snowpack water equivalent model for water resource monitoring and HRU- or grid-based hydrological model-based runoff forecasting. (See objectives 2 and 3, section 2.2; and summary in A3.8.)
- 10) Building on outputs #2, #3 and #7, we customized our leddar profiling sensor to build a low cost low power robotic scanning system for in-situ canopy leaf area index (LAI) phenology monitoring (Xi et al, 2019; A3.9). The prototype was developed and tested but there is no further funding at this time for refinement of the hardware and software for continuous operation. Note, LAI is a key landcover parameter in hydrological models to calculate the fluxes of precipitation interception, snow melt and evaporation. Consequently, new sensors that can be left operating in-situ beneath canopy to monitor 3D foliage profiles could be used in existing meteorological monitoring networks to aid with runoff forecasting. (see objective 4, section 2.2)
- 11) Researchers from the Prairie Adaptation Research Collaborative (PARC – Sauchyn et al.), at the University of Regina, contributed to evaluating the hydrological response of the montane to alpine transition zones of the Castle headwaters (See Appendix 10). Climate change projections were derived from 10 Regional Climate Models (RCMs) at high resolution (25km) (see Barrow and Sauchyn 2017) and used as inputs to the ACRU hydrological model for runoff simulation (A3.11). (see objective 7, section 2.2)

- 12) The mean annual flow of the Castle River at Beaver Mines was reconstructed from 1150 to 2010 using tree-ring data from multiple species and proxies (i.e. earlywood, latewood, and annual ring-width). The proxy streamflow data were downscaled stochastically to weekly estimates using methods developed by Sauchyn and Ilich (2017). The spectral analysis of this long times series revealed statistically significant modes of variability at inter-annual and decadal scales, associated with the strong influence of coupled ocean-atmosphere oscillations in the Pacific Ocean basin on the hydroclimate of western Canada. (See [A3.10.](#)) (see [objective 7 and 8, section 2.2](#))
- 13) A new high resolution soil map was developed from field sample data for the Westcastle (Kienzle et al. See [A3.11](#)) (see [objective 6, section 2.2](#))
- 14) Water yields upstream of the Oldman Reservoir were simulated for historical (1951-2017) and future (2041-2070) water years using inputs from #11, #12 and #13, above (Kienzle et al. see [A3.11](#)) (see [objective 8, section 2.2](#))

2.5 Knowledge Dissemination (supported HQP underlined)

Ph.D. Theses (2):

- Barnes, Celeste. ABD. INNOVATIONS IN OPERATIONAL SOURCE WATER SNOWPACK MONITORING USING IN SITU LEDDAR, AIRBORNE LIDAR AND MACHINE LEARNING IMPUTATION. Doctor of Philosophy Thesis, Department of Geography and Environment, University of Lethbridge.
- Xi, Zhouxin, 2019. FINE-SCALE INVENTORY OF FOREST BIOMASS WITH GROUND-BASED LIDAR. Doctor of Philosophy Thesis, Department of Geography, University of Lethbridge.

M.Sc. Theses (6):

- Bexte, Kyle. 2019. Dept. of Geography, U. of Lethbridge. A SPATIO-TEMPORAL ANALYSIS OF NEAR-SURFACE AIR TEMPERATURE WITHIN THE WEST CASTLE WATERSHED, ALBERTA. Masters of Science Thesis, Department of Geography, University of Lethbridge. Pp 106.
- Bonifacio, Charmaine. 2017, DEVELOPING ACRU UTILITIES FOR MODELLING FUTURE WATER AVAILABILITY: A CASE STUDY OF THE OLDMAN RESERVOIR WATERSHED, ALBERTA. Masters of Science Thesis, Department of Geography, University of Lethbridge
- Cartwright, Kelsey. 2019. SPATIO-TEMPORAL VARIATIONS IN SNOW DEPTH AND ASSOCIATED DRIVING MECHANISMS IN A TEMPERATE MESOSCALE MOUNTAINOUS WATERSHED. Masters of Science Thesis, Department of Geography, University of Lethbridge.
- Deering, Trevor 2019, MAPPING FUNDAMENTAL SOIL PROPERTIES IN THE WEST CASTLE WATERSHED, ALBERTA. Masters of Science Thesis, Department of Geography, University of Lethbridge.
- McCaffrey, David. 2018. ASSESSING HISTORIC CHANGE IN SUBALPINE FOREST: A

CASE STUDY IN THE WEST CASTLE WATERSHED. Masters of Science Thesis, Department of Geography, University of Lethbridge.

Okhrimenko, Maxim. 2018. APPLICATIONS OF MULTI-SPECTRAL LIDAR: RIVER CHANNEL BATHYMETRY AND CANOPY VEGETATION INDICES. Masters of Science Thesis, Department of Geography, University of Lethbridge.

Incomplete Theses (4 x students that started with WIP support but did not finish):

Herridge-Berry, Sean. M.Sc. candidate. LEDDAR WATER LEVEL MONITORING PROOF OF CONCEPT. (Dropped out of MSc to take another BSc).

Parsons, Reed. M.Sc. candidate. DEVELOPING A LEDDAR-BASED SNOWPACK SURFACE PROFILE MONITORING SYSTEM (Transferred out of Geography to Computer Science)

Porter, Thomas. MSc candidate. UAV SFM VEGETATION REGENERATION MONITORING IN A MOUNTAINOUS WATERSHED. (Withdrew after being offered a commercial UAV pilot job).

Quick, Dennis. M.Sc. candidate. RADAR-BASED CIRQUE SNOW- AND BURIED ICE-MELT CONTRIBUTIONS TO STREAM FLOWS IN THE WESTCASTLE WATERSHED. (Withdrew after being offered a consultancy project manager job).

Supporting HQP (13 x research assistants & independent study students):

Josh Montgomery, MSc (UL) - field support

Ike Allred, BSc (UL) - field support

Gaelle Gilson, PhD (UL) - field support

Dr. Craig Mahoney, PDF (UL) - field support

Emily Jones, MSc candidate (UL) - field support

Kailyn Nelson, PhD candidate (UL) - field support

Dr. Rihanna Peiman, PDF (UL) - SAR water surface mapping analysis

Iuliia Andreichuk, MSc (UR) - RCM modeling for input to ACRU

Samantha Kerr, BSc (UR) - Tree ring analysis to reconstruct stream flows for ACRU modeling

Colin Langhorn, BSc (UL) - RCM data analysis

Kaily Scott, BSc (UL) - ACRU set up support

Mark Derksen, BSc (UL) - lidar flight path modeling

David Tavernini, BSc (UL) - lidar snowpack modeling

Journal Papers Published / In Review (13):

Barnes, C., Hopkinson, C., Porter, T., Xi Z. 2020. IN-SITU LED-BASED OBSERVATION OF SNOW SURFACE AND DEPTH TRANSECTS. *Sensors* 20 (8), 2292

Barrow, EB and DJ Sauchyn. 2019. UNCERTAINTY IN CLIMATE PROJECTIONS AND TIME OF EMERGENCE OF CLIMATE SIGNALS IN WESTERN CANADA. *The International Journal of Climatology*, <https://doi.org/10.1002/joc.6079>

Cartwright, K., Hopkinson, C. Kienzle, S., Rood, S. EVALUATING LIDAR-ESTIMATED

SNOW DEPTH DRIVERS AND THEIR CONSISTENCY DURING MID WINTER AND MELT ONSET FOR A ROCKY MOUNTAIN WATERSHED. *Hydrological Processes*. (In Review)

Cartwright, K., Mahoney, C. Hopkinson, C. SPATIAL AND TEMPORAL CONSISTENCY OF MACHINE LEARNING-BASED IMPUTATION OF SNOWPACK DEPTH WITHIN AN OPERATIONAL LIDAR SAMPLING FRAMEWORK. *Remote Sensing*. (In Review)

Foster, S. G., Mahoney, J. M., Rood, S. B. 2018. FUNCTIONAL FLOWS: AN ENVIRONMENTAL FLOW REGIME BENEFITS RIPARIAN COTTONWOODS ALONG THE WATERTON RIVER, ALBERTA. *Restoration Ecology*, 26(5), 921-932.

McCaffrey, D. and Hopkinson, C. 2017. ASSESSING FRACTIONAL COVER IN THE ALPINE TREELINE ECOTONE USING THE WSL MONOPLOTTING TOOL AND AIRBORNE LIDAR. *Canadian Journal of Remote Sensing*. 43(5): 504-512.

McCaffrey and Hopkinson. 2020. REPEAT OBLIQUE PHOTOGRAPHY SHOWS TERRAIN AND FIRE-EXPOSURE CONTROLS ON CENTURY-SCALE CANOPY COVER CHANGE IN THE ALPINE TREELINE ECOTONE. *Remote Sensing*. *Remote Sensing* 12(10): 1569

McCaffrey and Hopkinson. HISTORIC CHANGE IN THE ALPINE TREELINE ECOTONE USING A RANDOM FOREST MODEL. *Canadian Journal of Remote Sensing* (Accepted pending revision)

Okhrimenko, Coburn, Hopkinson. 2019. MULTI-SPECTRAL LIDAR: RADIOMETRIC CALIBRATION, CANOPY SPECTRAL REFLECTANCE, AND VEGETATION VERTICAL SVI PROFILES. *Remote Sensing* 11 (13), 1556

Peimann, R., Hopkinson, C., Brisco, B. AUTOMATED SAR INTENSITY AND TEXTURE-BASED MONITORING OF SURFACE WATER EXTENT. *Canadian Journal of Remote Sensing* (Accepted pending revision)

Sauchyn, D and Ilich, N. 2017. NINE HUNDRED YEARS OF WEEKLY STREAMFLOWS: STOCHASTIC DOWNSCALING OF ENSEMBLE TREE-RING RECONSTRUCTIONS. *Water Resources Research*, DOI: 10.1002/2017WR021585

Xi , Hopkinson, Rood, Barnes, Xu, Pearce, Jones. 2019. A LIGHTWEIGHT LEDDAR OPTICAL FUSION SCANNING SYSTEM (FSS) FOR CANOPY FOLIAGE MONITORING. *Sensors* 19, 3943; doi:10.3390/s19183943

Journal Papers In Preparation (6):

Barnes and Hopkinson. WATERSHED-WIDE SNOW DENSITY AND SWE ESTIMATION USING SPATIALLY DISTRIBUTED TIME SERIES METEOROLOGICAL OBSERVATIONS, AIRBORNE LIDAR, AND FIELD MEASUREMENTS. To be submitted to *Hydrological Processes*.

Barnes and Hopkinson. AIRBORNE LIDAR FLIGHT PATH SAMPLING OPTIMIZATION FOR WATERSHED-WIDE SNOWPACK DEPTH MONITORING. To be submitted to *Hydrological Processes*

Barnes, Bexte, Bonaventure, Hopkinson. SPATIO-TEMPORAL VARIATION IN ATMOSPHERIC INVERSIONS IN A CHINOOK-PRONE MOUNTAINOUS

ENVIRONMENT IN THE ROCKIES OF SW ALBERTA. To be submitted to Arctic, Antarctic and Alpine Research.

Kienzle SW and Bonifacio CMT: HOW MUCH IRRIGATION WATER WILL WE HAVE IN THE FUTURE FOR SOUTHERN ALBERTA? Manuscript to be submitted to the Canadian Water Resources Journal.

Kienzle SW, Deering T and Jensen T: MAPPING HYDROLOGICAL SOIL PROPERTIES IN THE CASTLE RIVER WATERSHED. Manuscript to be submitted to the journal Geoderma (Elsevier).

Kienzle SW: SOME ADVANCES IN SETTING UP HYDRO-CLIMATOLOGICAL VARIABLES FOR PHYSICALLY-BASED SIMULATIONS IN MOUNTAIN HYDROLOGY. To be submitted to a Special Issue on 'Climatological and Hydrological Processes in Mountain Regions' in the Open Access MDPI Journal 'Atmosphere'.

Peer-reviewed Conference Proceedings (4):

Okhrimenko, M., Coburn, C., Hopkinson, C., 2018. INVESTIGATING MULTISPECTRAL LIDAR RADIOMETRY: AN OVERVIEW OF THE EXPERIMENTAL FRAMEWORK. *Proceedings of the International Geosciences and Remote Sensing Symposium, Valencia, Spain, July 2018. Published online: <http://www.igarss2018.org/>*

Peiman, R., Ali, H., Brisco, B., Hopkinson, C. 2017. PERFORMANCE EVALUATION OF SAR TEXTURE ALGORITHMS FOR SURFACE WATER BODY EXTRACTION THROUGH AN OPEN SOURCE PYTHON-BASED ENGINE. *Proceedings of the International Geosciences and Remote Sensing Symposium, Fort Worth, Texas, USA, July 2017. Published online: <http://www.igarss2017.org/>*

Cartwright, K., Parsons, N.R. Hopkinson, C. 2016. TERRAIN CHARACTERIZATION INFLUENCE ON SNOW ACCUMULATION AND PERSISTENCE. *Proceedings of the 73rd Annual Eastern Snow Conference, Columbus, Ohio. June 14-16.*

Parsons, N.R. and Hopkinson, C. 2016. THE DEVELOPMENT OF AN IN SITU LED DETECTION AND RANGING TECHNOLOGY FOR MONITORING SNOW DEPTH AND WATER LEVEL. *Proceedings of the 73rd Annual Eastern Snow Conference, Columbus, Ohio. June 14-16.*

Invited Seminars (8):

Hopkinson and Xi. 2019. EVOLUTION OF TERRESTRIAL LASER SCANNING AS A VIABLE FOREST INVENTORY AND PLOT MONITORING TOOL. *Canadian Institute of Forestry Workshop on Indigenous Consultation and Forestry: Finding Common Ground and Learning from Indigenous Knowledge. Edmonton, Nov 14.*

Hopkinson. 2019. THE EYES OF ARTEMIS: USING LASERS TO WATCH OVER OUR ECOSYSTEMS. *Public Professor Talk, University of Lethbridge. Sep 13.*

Kienzle SW 2019: ANALYZING IMPACTS OF CLIMATE CHANGE: APPROACHES, CHALLENGES AND POTENTIAL SOLUTIONS. Invited public presentation,

- University of KwaZulu-Natal, Centre for Water Research, April 30, 2019.
- Hopkinson, C. and L. Chasmer 2018. "CAN FLOODS AND FOREST FIRES BE MORE ACCURATELY PREDICTED USING 3D TECHNOLOGY?" *Public lecture to Southern Alberta Council on Public Affairs*. Televised by Shaw Media. April 12, 2018.
- Hopkinson, C. 2018. "USING / PRODUCING IMAGES AS DATE: USING LIDAR TO STUDY ECOSYSTEM DYNAMICS." *University of Lethbridge 'Take Two' Public Speaker Series*, Lethbridge February 1st
- Hopkinson, C. 2017. "HYDROGEOMATICS & THE WEB OF THINGS: ADVANCING HYDROLOGICAL RESEARCH AND MONITORING BY INTEGRATING 3D REMOTE SENSING, TELEMETRY AND ONLINE HYDROMETEOROLOGY." Invited seminar at the *Cold Regions Research Centre, Wilfrid Laurier University*. Waterloo, Ontario, Canada. October 26th.
- Kienzle SW 2016. DID IT REALLY GET WARMER IN ALBERTA BETWEEN 1950 AND 2010? BETTER CLIMATE INFORMATION FOR A BETTER FUTURE. Public presentation at the Galt Museum. Nov. 16, 2016.
- Kienzle SW 2016. HOW DID THE CLIMATE CHANGE BETWEEN 1950 AND 2010, AND WHAT DOES IT MEAN? Lethbridge and District Horticultural Society, July 27, 2016

Conference Presentations (40):

- Barnes, Porter, Xi, Hopkinson. 2019. COMBINING A LOW COST IN SITU LED RANGING AND INTENSITY SYSTEM WITH METEOROLOGICAL VARIABLES TO DEVELOP A TEMPORAL SNOWPACK DENSITY MODEL. *Canadian Symposium on Remote Sensing*. Fredericton, NB. June 4-6.
- McCaffrey and Hopkinson. 2019. RANDOM FOREST MODEL OF ALPINE TREELINE ECOTONE: ESTIMATING UPSLOPE ADVANCE USING HISTORIC PHOTOGRAPHY. *Canadian Symposium on Remote Sensing*. Fredericton, NB. June 4-6.
- Okhrimenko and Hopkinson. 2019. INVESTIGATING MULTI-SPECTRAL LIDAR RADIOMETRY: RECONCILING ALTITUDE AND CANOPY INFLUENCES ON ACTIVE VEGETATION INDICES. *Canadian Symposium on Remote Sensing*. Fredericton, NB. June 4-6.
- Porter, Barnes, Collingwood, Chasmer, Okhrimenko, Hopkinson. 2019. UAV PHOTOGRAMMETRY (SFM), MULTI-SPECTRAL LIDAR, AND AERIAL PHOTOGRAPHY DATA FUSION FOR WILDFIRE RECOVERY MONITORING IN WATERTON LAKES NATIONAL PARK. *Canadian Symposium on Remote Sensing*. Fredericton, NB. June 4-6.
- Xi, Hopkinson, Chasmer, Rood. 2019. DECOMPOSING PLOT-LEVEL TERRESTRIAL LASER SCANNING POINT CLOUDS INTO WOOD COMPONENTS FOR BIOMASS ESTIMATION BASED ON DEEP 3-D FULLY CONVOLUTIONAL NETWORKS AND QUANTITATIVE STRUCTURAL MODELING. *Canadian Symposium on Remote Sensing*. Fredericton, NB. June 19-21.
- Xi, Hopkinson, Rood, Barnes, Xu, Pearce, Jones. 2019. GENERATING HEMISPHERICAL COLORED POINT CLOUDS BY FUSING MULTI-BEAM LEDDAR®, MONOCULAR CAMERA AND PAN-TILT ROBOTICS FOR STATIONARY ENVIRONMENTAL

- MONITORING. *Canadian Symposium on Remote Sensing*. Fredericton, NB. June 19-21.
- Kienzle SW 2018: MAPPING CHANGES OF 43 CLIMATE INDICES AT HIGH SPATIAL RESOLUTION FOR ALBERTA, CANADA, FOR THE PERIOD 1950 TO 2010. In: *Climate Data Homogenization and Analysis of Climate Variability, Trends and Extremes*. Poster Presentation, EGU General Assembly, 7-12 April 2019, Vienna, Austria.
- Porter, Barnes, Hopkinson. UNMANNED AERIAL VEHICLES AND STRUCTURE-FROM-MOTION (SFM) FOR CHANGE DETECTION IN ALPINE ENVIRONMENTS. *Western Division of the Canadian Association of Geographers Conference*, Victoria, March 8-9.
- Kienzle SW 2019: ON THE POTENTIAL OF MAPPING CLIMATE INDICES TO DELINEATE CURRENT AND FUTURE WEATHER EXTREMES AND ECOSYSTEM CHANGES. 17th Savanna Science Network Meeting. Skukuza March 3-7, 2019.
- Kienzle SW 2018: WATER SECURITY AND CLIMATE CHANGE IN THE OLDMAN RIVER BASIN. Lethbridge College Science Café, Nov. 07, 2018.
- Bexte, K. J., Bonnaventure, P. P. 2018. CHARACTERIZATION OF INVERSION EVENTS WITHIN THE WEST CASTLE WATERSHED, ALBERTA. *Association of Canadian Universities for Northern Studies Student Conference*, Edmonton, AB, Canada. November 1-3
- Okhrimenko, Coburn, Hopkinson, 2018. INVESTIGATING MULTISPECTRAL LIDAR RADIOMETRY: AN OVERVIEW OF THE EXPERIMENTAL FRAMEWORK. Invited talk at the *International Geoscience and Remote Sensing Symposium*, Valencia, Spain July 22-27
- Kienzle SW and Clark C 2018: Creation of an Interactive Website for Mapping Changes of 43 Climate Indices for Alberta, Canada, for the Period 1950 to 2010. Universities Council on Water Resources Conference, Pittsburgh, PA, USA, June 26-28, 2018.
- Barnes, Hopkinson, Porter, Xi. 2018. MOUNTAIN WATERSHED SNOWPACK DEPTH AND DENSITY SCALING BY INTEGRATING LIDAR, PHOTOGRAMMETRY, WEBCAM, METEOROLOGICAL AND IN SITU FIELD DATA. *Canadian Symposium on Remote Sensing*. Saskatoon, June 19-21.
- Bosak, Houtekamer, Porter, Van Der Laan, Hopkinson, 2018. EFFECTS OF THE KENOW FIRE ON RUNOFF RESPONSE AND POTENTIAL INFRASTRUCTURE RISK IN WATERTON NATIONAL PARK, AB; A REMOTE SENSING EDUCATIONAL CASE STUDY. *Canadian Symposium on Remote Sensing*. Saskatoon, June 19-21.
- Hopkinson, Barnes, Cartwright, Porter, Quick, Xi. 2018. INTEGRATING AIRBORNE LIDAR AND IN SITU LED RANGING FOR MOUNTAIN SNOWPACK MONITORING. *Canadian Symposium on Remote Sensing*. Saskatoon, June 19-21.
- McCaffrey and Hopkinson. 2018. REPLICATING A RANDOM FOREST MODEL OF ALPINE TREELINE ECOTONE DYNAMICS. *Canadian Symposium on Remote Sensing*. Saskatoon, June 19-21.
- McCaffrey, Hopkinson, Brisco. 2018. BOREAL WETLAND WATER FLUX MONITORING WITH ALS AND SAR CORRELATES TO REGIONAL PATTERNS OF PRECIPITATION. *Canadian Symposium on Remote Sensing*. Saskatoon, June 19-21.
- Quick, Hopkinson, Brisco. 2018. COMBINED UTILIZATION OF HIGH RESOLUTION

- LIDAR WITH POLARIMETRIC SAR FOR DETERMINATION OF SNOW WETTING FRONTS AND DERIVED MELT RATES IN COMPLEX TERRAIN. *Canadian Symposium on Remote Sensing*. Saskatoon, June 19-21.
- Xi and Hopkinson, 2018. A SMALL LOW-COST 3D SCANNING AND IMAGING SYSTEM FOR IN-SITU HYDRO-ECOLOGICAL MONITORING. *Canadian Symposium on Remote Sensing*. Saskatoon, June 19-21.
- Hopkinson, Barnes, Cartwright, Quick, Xi, Brisco. 2018. INTEGRATING AIRBORNE LIDAR, IN SITU LED RANGING AND SAR FOR SNOWPACK MONITORING *IAHS Remote Sensing Hydrology Symposium*, Cordoba, Spain, May 8-10
- Kienzle SW 2018: ALBERTA WATER UNDER CONDITIONS OF CLIMATE CHANGE. WHAT SHOULD WE DO? Public Lecture and Panel Discussion at Pat Carlson Speaker Series: Climate change and the future of energy. Medicine Hat, Jan. 25, 2018.
- Kienzle SW 2018: ALBERTA WATER UNDER CONDITIONS OF CLIMATE CHANGE. WHAT SHOULD WE DO? Public Lecture and Panel Discussion at Pat Carlson Speaker Series: Climate change and the future of energy. Lethbridge, Jan. 24, 2018.
- McCaffrey, Hopkinson. 2017. TOPOGRAPHIC PATTERNS OF MORTALITY AND SUCCESSION IN THE ALPINE TREELINE ECOTONE SUGGEST HYDROLOGIC CONTROLS ON POST-FIRE TREE ESTABLISHMENT. *American Geophysical Union, Fall Meeting*. New Orleans, USA, December 11-15
- Peiman, R. Ali, H., Brisco, B. Hopkinson, C. 2017. PERFORMANCE EVALUATION OF SAR TEXTURE ALGORITHMS FOR SURFACE WATER BODY EXTRACTION THROUGH AN OPEN SOURCE PYTHON-BASED ENGINE. *International Geosciences and Remote Sensing Symposium*. Fort Worth, Texas, USA. July 23-28.
- Hopkinson, C. Ali, Barnes, Brisco., McCaffrey, Montgomery, Patterson, Peiman. 2017. AN AUTOMATED ONLINE INTERACTIVE SAR & LIDAR SURFACE WATER ASSESSMENT & MONITORING PORTAL. *International Association of Hydrological Sciences Conference*. Port Elizabeth, South Africa. July 10-14.
- Hopkinson, C. Cartwright, McCaffrey, Kienzle, Rood. 2017. A MOUNTAIN WATER RESOURCES MONITORING AND MODELLING FRAMEWORK USING A TEMPORAL FUSION OF OPTICAL, LIDAR AND RADAR IMAGE DATA SOURCES AT CENTURY TO SEASON SCALES. *International Association of Hydrological Sciences Conference*. Port Elizabeth, South Africa. July 10-14.
- Cartwright, K. and Hopkinson, C. 2017. EVALUATION OF RANDOM FOREST SNOW DEPTH MODELING USING AIRBORNE LIDAR. *38th Canadian Remote Sensing Symposium*. Montreal, Quebec. June 20-22
- Hopkinson, C. Ali, Brisco., McCaffrey, Montgomery, Patterson, Peiman. 2017. AN ONLINE PORTAL FOR AUTOMATED SAR AND LIDAR INTEGRATION FOR SURFACE WATER MONITORING. *38th Canadian Remote Sensing Symposium*. Montreal, Quebec. June 20-22.
- Hopkinson, C. Cartwright, McCaffrey, Kienzle, Brisco. 2017. A MOUNTAIN WATER RESOURCES MONITORING FRAMEWORK USING A TEMPORAL FUSION OF OPTICAL, LIDAR AND RADAR IMAGE DATA SOURCES AT CENTURY TO SEASON SCALES. *38th Canadian Remote Sensing Symposium*. Montreal, Quebec. June 20-22.

- McCaffrey, D. and Hopkinson, C. 2017. OBLIQUE PHOTOGRAPHY AND CANOPY COVER CLASSIFICATION IN THE ALPINE TREELINE ECOTONE. 38th *Canadian Remote Sensing Symposium*. Montreal, Quebec. June 20-22.
- McCaffrey, D. and Hopkinson, C. 2017. TOPOGRAPHY AND CANOPY COVER CHANGE IN THE ALPINE TREELINE ECOTONE: CENTURY -SCALE OBSERVATIONS. 38th *Canadian Remote Sensing Symposium*. Montreal, Quebec. June 20-22.
- Neilsen, T. Smith, B., Houston, T., Hopkinson, C. 2017. FUSING LIDAR AND RUNOFF TIME SERIES DATA TO MODEL FLOOD SCENARIOS IN THE HEADWATERS OF THE OLDMAN RIVER BASIN, ALBERTA. 38th *Canadian Remote Sensing Symposium*. Montreal, Quebec. June 20-22.
- Peiman, R. Ali, H., Brisco, B. Hopkinson, C. 2017. AUTOMATED SURFACE WATER THRESHOLDING TECHNIQUES BASED ON SAR TEXTURE AND DB BACKSCATTER USING MULTI-TEMPORAL RADARSAT-2 IMAGERY. 38th *Canadian Remote Sensing Symposium*. Montreal, Quebec. June 20-22.
- Xi, Z., Hopkinson, C., Goulen, T. 2017. A LOW-COST LIDAR SCANNING SCHEME INTEGRATING MULTI-BEAM LEDDAR, SERVO, AND VIDEO. 38th *Canadian Remote Sensing Symposium*. Montreal, Quebec. June 20-22.
- Cartwright, K. and Hopkinson, C. 2017. INTER AND INTRA-ANNUAL CONSISTENCY OF SEASONAL SNOWPACK PROPERTIES USING CONVENTIONAL SNOW DATA AND AIRBORNE LIDAR. *Canadian Water Resources Association Conference*. Lethbridge, Alberta. June 5-7.
- Kienzle SW 2017: MAPPING TRENDS OF 43 CLIMATE INDICES AT HIGH SPATIAL RESOLUTION FOR ALBERTA, CANADA, FOR THE PERIOD 1950 – 2010. Canadian Association of Geographers, May 29-June 02, 2017, York University, Toronto.
- McCaffrey, D. and Hopkinson. 2016. ESTIMATION OF BIOMASS DYNAMICS IN ALPINE TREELINE ECOTONE USING AIRBORNE LIDAR AND REPEAT PHOTOGRAPHY. AGU Fall Meeting, San Francisco, December 12-16.
- Cartwright, K., Parsons, N.R. Hopkinson, C. 2016. TERRAIN CHARACTERISTIC INFLUENCE ON SNOW ACCUMULATION AND PERSISTENCE: CASE STUDY. 73rd *Annual Eastern Snow Conference*, Columbus, Ohio. June 14-16.
- Parsons, N.R. and Hopkinson, C. 2016. IN-SITU LIGHT EMITTING DIODE DETECTION AND RANGING FOR THE MAPPING OF SNOW SURFACE TOPOGRAPHY AND DEPTH. 73rd *Annual Eastern Snow Conference*, Columbus, Ohio. June 14-16.
- Kienzle SW 2016. HIGH-RESOLUTION MAPPING OF CLIMATE INDICES AND THEIR TRENDS ACROSS ALBERTA, CANADA, FOR THE PERIOD 1950 TO 2010. *The 13th International Meeting on Statistical Climatology*. June 6-10, Canmore, Alberta.

Media outputs:

Thomas Porter Productions Interview: UNDERSTANDING OUR HEADWATERS. online video interview / mini documentary, April 2017.

<https://www.youtube.com/watch?v=PoutMOD94IA>

Hopkinson, C. 2016. AIRBORNE LIDAR FIELD OPERATIONS LECTURE. Ad Hoc presentation given outside hangers of our lidar survey aircraft at end of summer missions. <https://www.youtube.com/watch?v=L1dfGL7158w>

Lethbridge Herald Newspaper article by Tijana Martin: CLASSROOM WITH A VIEW. May 27th, 2016: <http://lethbridgeherald.com/news/local-news/2016/05/27/classroom-with-a-view/>

Lethbridge Living online article: SHELL CANADA DONATION HELPS REVITALIZE UNIVERSITY OF LETHBRIDGE'S WESTCASTLE FIELD STATION. May 26th, 2016: <http://lethbridgeliving.com/news/item/shell-canada-donation-helps-revitalize-university-of-lethbridge-s-westcastle-field-station> (mention of West Castle research hydro-meteorological program)

3. Appendices

Appendix 3.1: Assessing Fractional Cover in the Alpine Treeline Ecotone Using the WSL Monoplotting Tool and Airborne Lidar

McCaffrey, D. and Hopkinson, C. 2017. ASSESSING FRACTIONAL COVER IN THE ALPINE TREELINE ECOTONE USING THE WSL MONOPLOTTING TOOL AND AIRBORNE LIDAR. *Canadian Journal of Remote Sensing*. 43(5): 504-512.

Assessing Fractional Cover in the Alpine Treeline Ecotone Using the WSL Monoplotting Tool and Airborne Lidar

Research Note

Author 1: David R. McCaffrey, Department of Geography, Alberta Water & Environmental Science Building, University of Lethbridge, 4401 University Drive, Lethbridge, Alberta, T1K 3M4, phone: 403-329-2225, fax: 403-332-4039, david.mccaffrey@uleth.ca

Author 2: Chris Hopkinson, Department of Geography, Alberta Water & Environmental Science Building, University of Lethbridge, 4401 University Drive, Lethbridge, Alberta, T1K 3M4, phone: 403-329-2225, fax: 403-332-4039, c.hopkinson@uleth.ca

Abstract

As forest cover in mountain areas impacts headwater properties like habitat extent and downstream water resources, it is important to assess and understand the changes that occur across forest boundaries like the alpine treeline ecotone (ATE). Such changes occur slowly and manifest at decadal to century time scales, which requires a high spatial and temporal resolution of analysis for accurate meaningful results. In this research note, fractional cover classification of the ATE is achieved through oblique photography analysis with the WSL Monoplotting Tool. Seven photographs of the West Castle Watershed (Alberta, Canada), collected by the Mountain Legacy Project, were gridded to a 20 m resolution and assigned canopy cover classes by manual interpretation. Four canopy cover classes (i.e. no cover, low vegetation, partial canopy, full canopy) were validated through comparison to lidar-derived fractional cover. The extraction of canopy cover information from oblique photography at a resolution of 20 m introduces the ability to assess and quantify changes in the ATE over century-scale oblique photograph records.

Introduction

Alpine Treeline Ecotone (ATE), the boundary between alpine tundra and closed canopy, is characterized by distinct patterns in forest structure; decreases in mature tree height, stand density, and fractional cover correlate with increases in elevation (Körner, 2012; Tranquillini, 1979). Environmental factors influencing these forest structure patterns are varied (Case & Duncan, 2014; Holtmeier & Broll, 2005; Weiss et al., 2015), and understanding the degree to which multi-scale interactions between anthropogenic, orographic, and climatic factors affect shifts in ATE has relied on observation methods such as repeated plot measurements, dendrochronology, remote sensing, and historic records (e.g. oblique photography).

ATE observation methods are limited in the temporal and spatial extents at which they can resolve change (Danby, 2011). For example, forest structure parameters that characterize ATE, such as fractional cover, can be observed at a high spatial resolution (<1 m²) using airborne lidar, but few extended lidar records presently exist for periods longer than 20 years. Conversely, oblique photograph records of ATE exist for periods exceeding 100 years (e.g. Trant et al., 2015), but quantitative analysis of century-scale land cover change using oblique

photographs has been restricted to low spatial resolutions ($\sim 100 \text{ m}^2$) (Stockdale et al., 2015). In the Rocky Mountains, where advance of ATE species such as subalpine fir (*Abies lasiocarpa*) has been observed at $0.28\text{-}0.62 \text{ m yr}^{-1}$ (Bekker, 2005; Luckman & Kavanagh, 2000) the ability to detect shifts in ATE using repeat oblique photograph will depend on increasing the spatial resolution of analysis above 100 m^2 ; e.g. if ATE has advanced on the order of $\sim 50 \text{ m}$ in a century, in order to detect change using a century scale photographic record, the spatial resolution of analysis must be higher than 100 m . This research note applies recent advances in oblique photography analysis to increase the spatial resolution of fractional cover classification from oblique images of ATE, while comparing the photogrammetrically interpreted observations of canopy cover with empirically-derived estimates from lidar.

Quantitative Analysis of Oblique Photographs

Prior limitations of quantitative oblique photography analysis have been overcome with the advent of software like the WSL monoplottting tool (Bozzini et al., 2012) (WSL is the German acronym for the Swiss Federal Institute for Forest, Snow and Landscape Research), which can produce georeferenced vector data from oblique photographs using tie points between oblique and aerial photographs, and high resolution topographic data. Previous analyses of oblique photographs to assess change in the ATE were mostly qualitative, relying on descriptions of land cover change and estimations of spatial extent (Butler & DeChano, 2001; Klasner & Fagre, 2002; Kullman & Öberg, 2009; Moiseev & Shiyatov, 2003; Roush et al., 2007). Stockdale et al. (2015) introduced a method of quantifying land cover change using the WSL monoplottting tool and repeat photography from the Mountain Legacy Project (MLP), a collection of over 120,000 historic survey images of the Canadian Rocky Mountains (1888 – 1958) (Trant et al., 2015). The study compared estimated UTM coordinates of tie points used in WSL monoplottting camera calibration with the coordinates of points that had been projected from the captured oblique to orthogonal perspective using the software; the mean distance error of all test points in an oblique image was observed to be 14.7 m , while mean displacement error (i.e. difference between the centroid of the set of all test points in a given image) was 2.9 m (Stockdale et al., 2015).

The high spatial resolution of airborne lidar makes it a promising candidate for validation of oblique photograph analysis, but at the time of writing, only a single study is known to have compared land cover data generated from the WSL monoplottting tool against vegetation measurements from airborne lidar; using an object based image analysis and lidar derived estimates of vegetation height and cover as validation, Kolecka et al. (2015) mapped areas of forest succession from oblique photographs with an overall accuracy of 95%.

The present research compares manual estimates of ATE canopy cover from oblique photography with lidar-derived estimates of fractional cover, at a spatial resolution which is sufficiently high to observe and measure shifts in the ATE over a century scale photographic record (i.e. $\sim 20 \text{ m}$). The rationale for comparing oblique photogrammetric and lidar canopy cover estimates is that temporal change analysis is assumed most effective when the image data being compared are analogous; i.e. it is more logical to assess change from contemporary and historical oblique photographs than, say, historical oblique imagery and contemporary lidar. However, lidar is a reliable and widely used technique of contemporary canopy cover assessment, so comparing canopy cover data derived from contemporary oblique photographs and lidar enables long term change assessments to be placed into a modern context with wide-

ranging applicability.

Data Collection and Analysis

Oblique Photographs

Research was conducted over the West Castle Watershed (WCW), Alberta, Canada (49.3° N, 114.4° W), in the headwaters of the Oldman River on the eastern slopes of the Rocky Mountains, with an area of ~103 km², and an elevation range of ~1400-2600 masl. Aside from a small ski resort and village near the downstream end of the WCW (2.9% of the watershed area), and trails in the valley bottoms, the forested slopes demonstrate limited anthropogenic disturbance, have not been subject to mountain pine beetle infestation, and have no record of major fires. Consequently, the WCW is an ideal study area for assessing recent natural shifts in the treeline ecotone in this part of the Canadian Rockies. Seven high-resolution oblique photographs of the WCW were selected from a set of 46 provided by the MLP, which maximized the spatial coverage of the valley, while also providing areas of overlap for classification error analysis. All photographs were collected in August 2006, and geolocations for photograph origins were provided by the MLP (Table 1).

Table 1 – Oblique Photograph Summary

| Photo | Elev (masl) | Latitude | Longitude | Mean Angle Error (°) | # of tie points | Mean Distance Error (m) | SE (m) | Mean Displacement Error (m) |
|-------|-------------|---------------|----------------|----------------------|-----------------|-------------------------|--------|-----------------------------|
| 1 | 2295 | 49°16'3.62"N | 114°22'58.97"W | 0.09 | 4 | 15.99 | 1.1 | 2.85 |
| 2 | 2216 | 49°20'14.00"N | 114°23'22.81"W | 0.11 | 10 | 22.06 | 2.0 | 12.12 |
| 3 | 2390 | 49°17'7.010"N | 114°20'44.21"W | 0.085 | 6 | 71.71 | 0.7 | 55.00 |
| 4 | 2283 | 49°16'4.94"N | 114°22'56.77"W | 0.06 | 5 | 5.21 | 2.5 | 4.17 |
| 5 | 2283 | 49°16'4.94"N | 114°22'56.77"W | 0.092 | 8 | 6.46 | 3.6 | 1.04 |
| 6 | 2390 | 49°17'7.010"N | 114°20'44.21"W | 0.059 | 4 | 49.54 | 0.4 | 42.70 |
| 7 | 2462 | 49°18'7.26"N | 114°22'18.63"W | 0.118 | 7 | 11.06 | 1.7 | 5.70 |
| Avg | | | | 0.088 | 6.29 | 26.00 | 1.7 | 17.65 |

Rasterized canopy cover data were generated from oblique photographs, using WSL monoplottting software (Bozzini et al., 2012), generally following the method of Stockdale et al. (2015): 1) co-registered aerial imagery (1.5 m, SPOT) and topographic data (1 m DEM, lidar collection described below) were used to calibrate camera parameters for each of the seven images; 2) tie points were selected on the aerial and oblique images, and removed until the mean angle error reached ~0.1°; 3) the camera parameters were exported to ArcGIS 10.3 (Esri, 2014), where the 1 m DEM was used to create a viewshed for the image (Fig. 1b); 4) a 20 m

fishnet (i.e. 400 m² grid cells) was overlaid and clipped to the viewshed, with any grid cells having less than 75% coverage (i.e. < 300 m²) being omitted (Fig. 1c), 5) the *World to Pixel* function of the WSL software was used to project fishnet grid cells from UTM coordinates to pixel coordinates, so that the 20 m x 20 m grid cells were draped over the oblique image (Fig. 1d). Once in an oblique projection, grid cells were manually assigned into one of four canopy cover classes based on observed canopy openness and texture (Fig. 1e). Visual estimates of canopy cover from photography are commonplace and can be compared against lidar derived fractional cover. The four canopy cover classes were: 1) *No Cover* - grid cells devoid of vegetation; 2) *Low Vegetation* - grid cells appear vegetated, but context and texture suggest shrubs or krummholz, not upright trees >2 m; 3) *Partial Canopy* - trees are present, but ground is visible in > 50% of the grid cell; 4) *Full Canopy* - trees cover > 50% of a grid cell. Finally, classified oblique images were converted back to orthogonal view for comparative analysis (Fig. 1f).

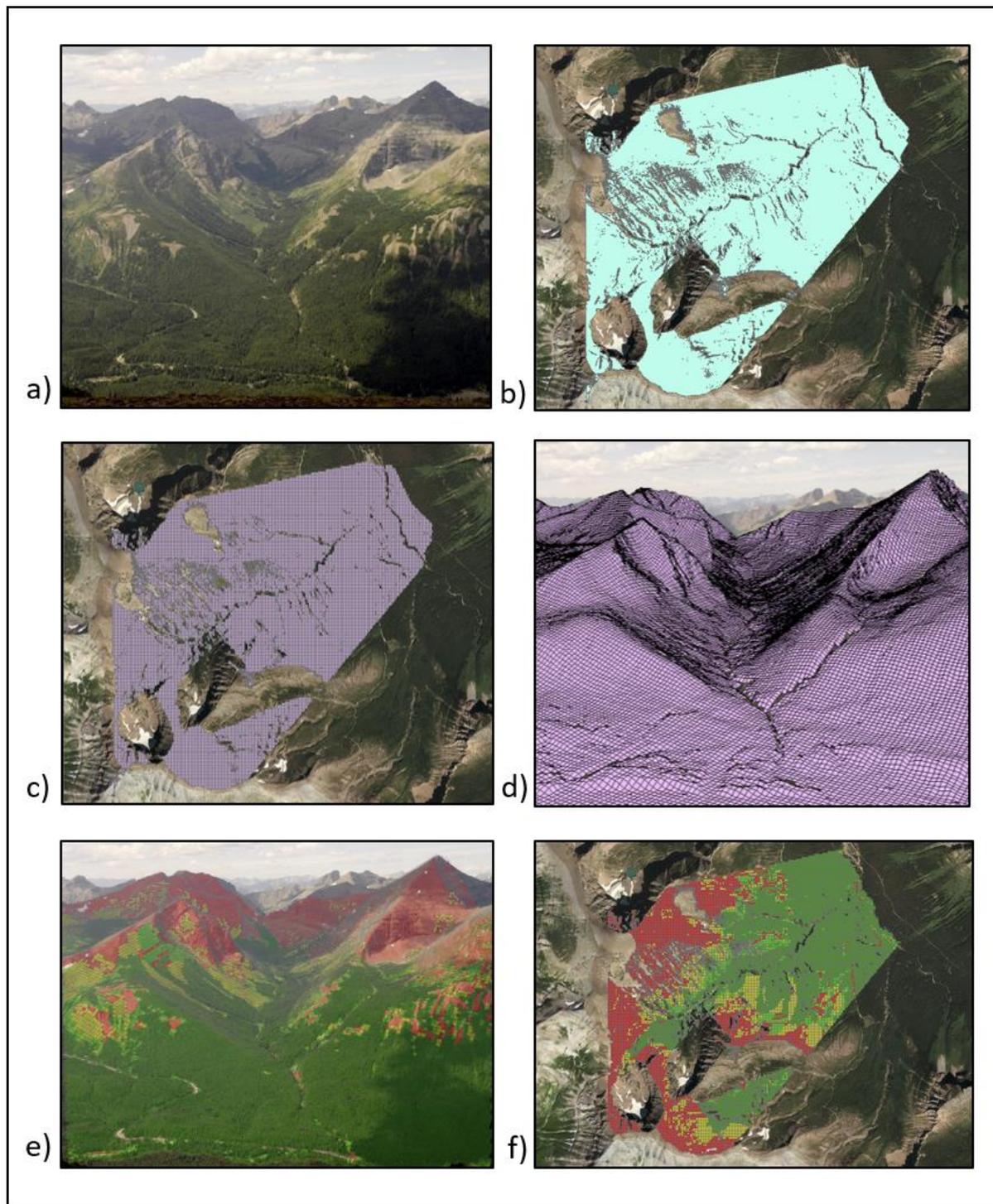


Figure 1 - The oblique photograph to raster workflow, as described by Stockdale et al. (2015); a) an oblique image of Mount Haig, in the West Castle valley, Copyright Mountain Legacy Project, 2006; b) view shed of Mount Haig image, using camera calibration from WSL Monoplotting tool; c) a 20 m fishnet applied to the Mount Haig view shed; d) fish net grid

project back to oblique view using WSL *World to Pixel* function; e) oblique image with canopy cover classification; f) orthogonal image with canopy cover classification.

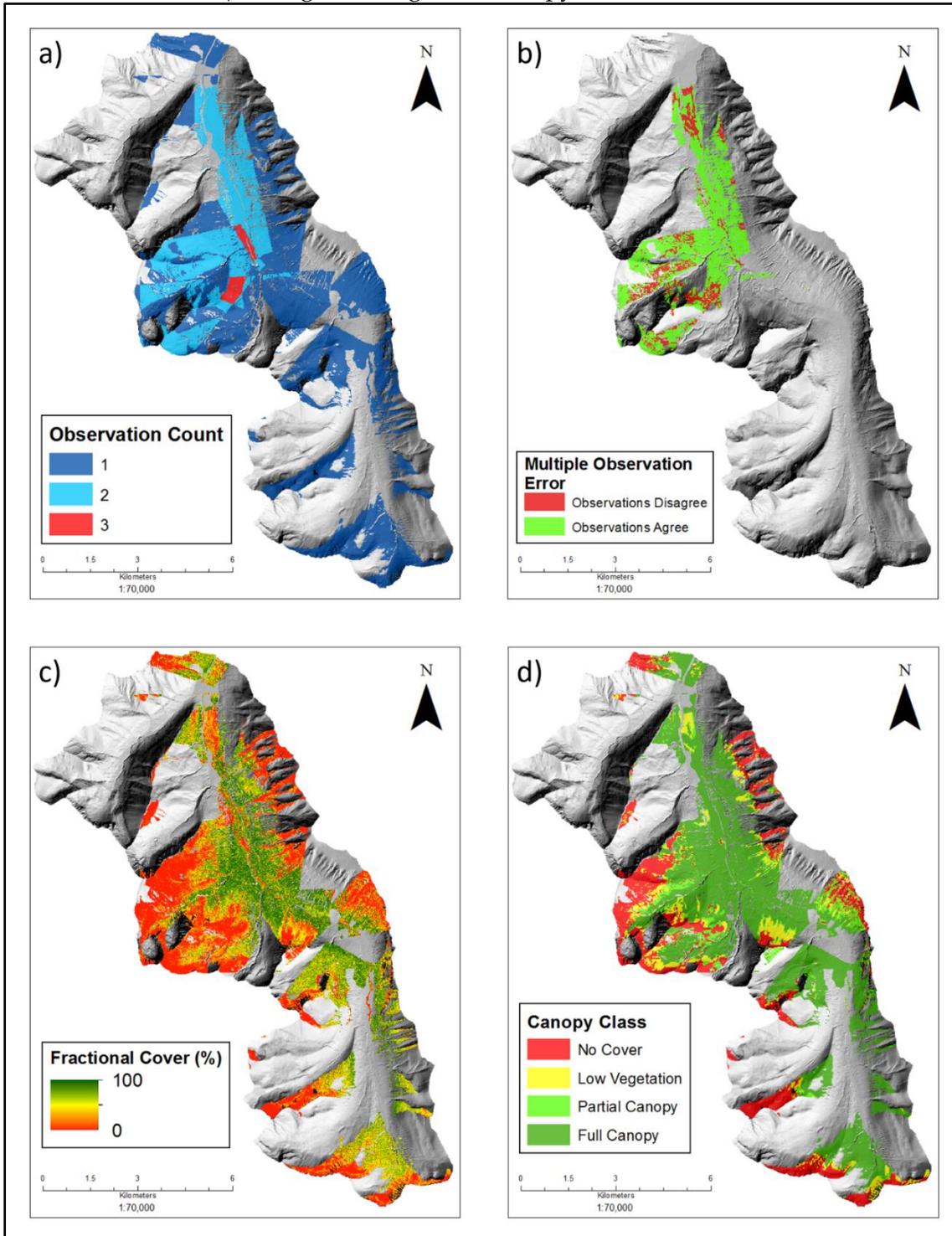


Figure 2 - Results of oblique photograph and fractional cover analyses in West Castle; a)

observation areas were viewed by either 1, 2, or 3 different oblique images; b) areas with multiple observation either agreed in all 2 or 3 observations, or disagreed; c) map of fractional cover (20), gridded from lidar using LAStools (Isenburg, 2013); d) canopy cover classifications, manually interpreted from oblique images and projected to orthogonal view.

Mean distance error and mean displacement error were calculated by importing the UTM coordinates (x,y,z) of registration points into the WSL monoplotting tool, and exporting the points to a shapefile using the camera calibration for each oblique image, as in Stockdale et al. (2015). Mean distance error was calculated as the 3d distance between each registration point and its projected counterpart, and mean displacement error was calculated as the 3d distance between the centroid of the set of registration points and the set of projected points.

Grid cells that extended above ridgelines or the horizon were omitted from analysis, leaving an area of 46.6 km², or 43% coverage of West Castle Watershed. Using the Alberta Vegetation Inventory (Alberta Sustainable Resources Development, 2005), vector data for known areas of disturbance (e.g. former cut blocks, oil well pads, roads, ski resort, and a small burn area) were aggregated into a single layer, and any grid cells intersecting with this disturbance layer were omitted, leaving a final observation area of 39.8 km², or 38.6% of WCW. Of the observed area, 10.7 km² (26.9%) had multiple observations allowing for an assessment of the manual classification consistency from different vantage points (Fig. 2a, Table 2). Accuracy of canopy cover classifications was assessed by determining whether the minimum and maximum classification values (for classes 1 to 4) for each grid cell agreed or disagreed. Terrain slope was investigated as a possible source of classification error; slope distribution of areas of agreement and disagreement were compared.

Table 2 – Observation

| Number of Observations | Area (km²) | Percent of Observed Area |
|-------------------------------|------------------------------|---------------------------------|
| 1 | 29.1 | 73.1% |
| 2 | 10.0 | 25.2% |
| 3 | 0.7 | 1.7% |
| Multiple Observations | | |
| Agree | 8.5 | 79.5% |
| Disagree | 2.2 | 20.5% |

Airborne Lidar

An airborne lidar survey of West Castle Valley was flown on October 18th, 2014 with a Leica ALS70 with a minimal point spacing of 3pts/m² at nominal altitude (1300 m above the mean ground surface elevation). Lidar point cloud processing and fractional cover analysis approximately followed the return-count ratio methods described in Hopkinson and Chasmer (2009). Data were classified into ground and non-ground using *Terrascan* (Terrasolid, Finland) and a 1m DEM was interpolated from the ground-classified points and aggregated using mean values to match the 20 m canopy cover classification grid cells. Fractional cover was calculated

using the LAStools suite (Isenburg, 2013); data were height normalized, and *lascanopy* was applied with a height threshold of 2.0 m, and a step of 20 m, again matching canopy cover classification grid (Fig. 2c). Differences in the distributions of fractional cover for each canopy cover classification type were analyzed using ANOVA and Tukey honest significant difference (HSD) pairwise comparison.

Results

Camera calibration of the seven oblique photographs using the WSL monoplotting tool resulted in a mean angle error of 0.088° (below the recommended target minimum of 0.100°), a mean distance error of 26.00 m per point, and mean displacement error of 17.65 m for the centroid of all points in a calibrated image (Table 1). Comparison of areas of multiple observation showed that 79.5% of the grid cells agreed between observations, while at least one classification value disagreed in 20.5% of cases (Fig. 2b, Table 2). Terrain slope was significantly higher in disagreement cases than in agreement cases (ANOVA, $F = 98.6$, $df = 1$, $p < 0.001$) (Fig. 3).

Results of ANOVA indicated a significant difference in the distribution of fractional cover for at least one of the four canopy cover classifications ($F = 1586$, $df = 3$, $p < 0.001$), and Tukey HSD pairwise comparison showed significant differences in fractional cover between all six pairwise cases (Fig. 4, Table 3). As this result may have been driven by expected areas of homogenous full canopy cover at lower elevations and no cover at higher elevations, the analysis was re-run, but restricted to elevations between 1800-2300 m, where ATE is expected to cause more variability in canopy cover. The results were repeated; ANOVA for ATE area only indicated a significant difference in the distribution of fractional cover for at least one of the four canopy classifications ($F = 576.4$, $df = 3$, $p < 0.001$), and Tukey HSD pairwise test showed significant differences in fractional cover between five of the six pairwise cases (Fig. 4, Table 3). With analysis restricted to ATE elevations, there was no significant difference in the distribution of fractional cover between low vegetation and partial canopy classes.

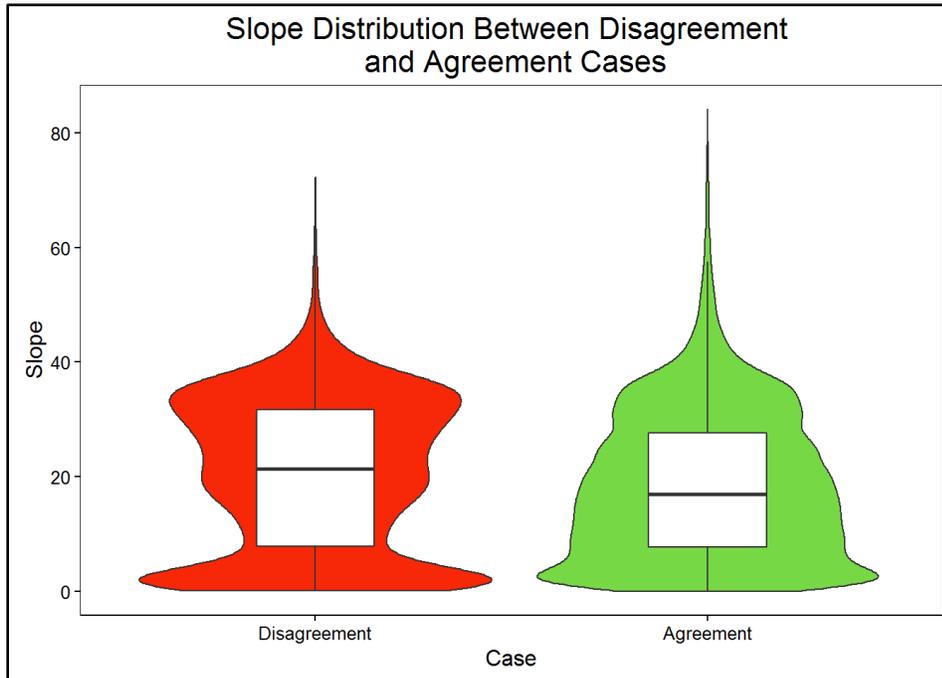


Figure 3 - Box and violin plot showing slope distribution for areas where multiple observations agreed and disagreed. Slope was higher in disagreement cases than in agreement cases (ANOVA, $F = 98.6$, $df = 1$, $p < 0.001^{***}$).

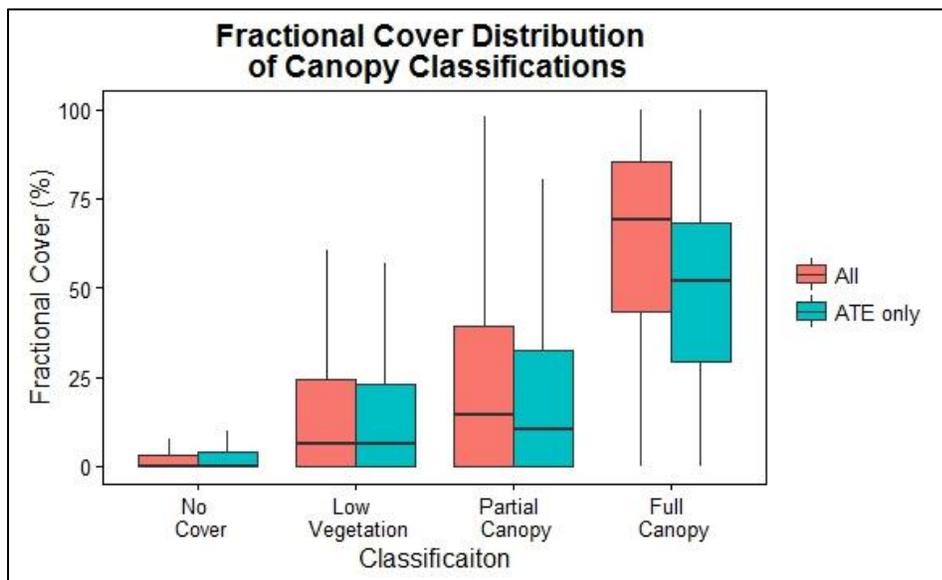


Figure 4 - Manual canopy cover classification from oblique images results in significantly different distributions of fractional cover; pink plots show analysis of all observation area, while blue plots show analysis restricted to approximate ATE elevations (1800-2300 masl). All within group differences are highly significant, with the exception of *low vegetation* and *partial canopy* in the ATE only analysis (see Table 3 for details).

Table 3 – Results of Tukey HSD Pairwise Comparison

| | All Observations | | | | ATE ONLY | | | |
|-------|------------------|-------|-------|------------|------------|-------|-------|------------|
| | difference | lower | upper | p value | difference | lower | upper | p value |
| 1 - 2 | 8.7 | 7.9 | 9.5 | < 0.001*** | 7.9 | 4.2 | 11.6 | < 0.001*** |
| 1 - 3 | 17.8 | 17.0 | 18.6 | < 0.001*** | 11.7 | 7.8 | 15.7 | < 0.001*** |
| 1 - 4 | 56.0 | 55.5 | 56.5 | < 0.001*** | 40.6 | 38.0 | 43.1 | < 0.001*** |
| 2 - 3 | 9.1 | 8.1 | 10.1 | < 0.001*** | 3.8 | -1.0 | 8.7 | 0.170 |
| 2 - 4 | 47.3 | 46.6 | 48.0 | < 0.001*** | 32.7 | 28.9 | 36.4 | < 0.001*** |
| 3 - 4 | 38.2 | 37.5 | 39.0 | < 0.001*** | 28.8 | 24.8 | 32.9 | < 0.001*** |

Discussion and Conclusion

The analyses indicated that fractional cover across the ATE could be discriminated from oblique imagery at a resolution of 20 m. Mean distance error and mean displacement error were both substantially higher than error values calculated in Stockdale et al. (2015), potentially as a result of the use of fewer test points. However, mean distance error was comparable to canopy cover grid cell resolution, and mean displacement error remained below this resolution.

Comparison of grid cells with multiple observations showed that canopy cover classification disagreed in 20.5% of cases. Projection error may arise from numerous factors, including distance from lens, and observation angle, which is a function of both the slope and aspect of the target. We tested the difference between slope in agreement and disagreement cases, and found that mean terrain slope was higher in cases of disagreement. Low observation angle is expected to increase vector length error in WSL calibration (Stockdale et al., 2015), so if terrain slope were the only control on observation angle, the expectation would be lower terrain slope in cases of disagreement. The observation that terrain slope is higher in cases of disagreement suggests that slope is playing a limited role in determining observation angle, and/or disagreement may result from difficulty in assessing fractional cover in high slope terrains. Cases of disagreement may result from projection error, manual classification error, or both.

Manual canopy cover classification produced significantly different fractional cover distributions, demonstrating that fractional cover assessment through oblique photography is a valid method of monitoring change in the ATE. Across the full watershed, this analysis was able to distinguish between four canopy cover classes (i.e. No Cover, Low Vegetation, Partial Canopy, and Full Canopy), but in the ATE the method could only confidently resolve three classes (i.e. No Cover, Low Vegetation/Partial Canopy, and Full Canopy). This result is unsurprising, given the well documented difficulty discerning low vegetation from trees in the ATE (Korner, 2012). With an abundance of repeat oblique photography in the Canadian Rocky Mountains (Trant et al., 2015), future research which combines quantitative analysis of oblique photography with lidar derived forest structure data can increase the spatial and temporal extent of ATE monitoring. This is important as we continue to explore the impacts of climatic changes on forested mountain ecosystems and downstream water resources.

References

- Alberta Sustainable Resources Development. (2005). *Alberta Vegetation Inventory, 2.1.1.*
- Bekker, M. F. (2005). Positive Feedback between Tree Establishment and Patterns of Subalpine Forest. *Arctic, Antarctic, and Alpine Research, 37*(1), 97-107.
- Bozzini, C., Conedera, M., & Krebs, P. (2012). A New Monoplotting Tool to Extract Georeferenced Vector Data and Orthorectified Raster Data from Oblique Non-Metric Photographs. *International Journal of Heritage in the Digital Era, 1*(3), 499-518.
- Butler, D. R., & DeChano, L. M. (2001). ENVIRONMENTAL CHANGE IN GLACIER NATIONAL PARK, MONTANA: AN ASSESSMENT THROUGH REPEAT PHOTOGRAPHY FROM FIRE LOOKOUTS. *Physical Geography, 22*(5), 291-304.
- Case, B. S., & Duncan, R. P. (2014). A novel framework for disentangling the scale-dependent influences of abiotic factors on alpine treeline position. *Ecography, 37*(9), 838-851.
- Danby, R. K. (2011). Monitoring Forest-Tundra Ecotones at Multiple Scales. *Geography Compass, 5*(9), 623-640.
- Esri. (2014). ArcGIS 10.3 for Desktop: Environmental Systems Research Institute Redlands, CA USA.
- Holtmeier, F. K., & Broll, G. (2005). Sensitivity and response of northern hemisphere altitudinal and polar treelines to environmental change at landscape and local scales. *Global Ecology and Biogeography, 14*(5), 395-410.
- Hopkinson, C., & Chasmer, L. (2009). Testing LiDAR models of fractional cover across multiple forest ecozones. *Remote Sensing of Environment, 113*(1), 275-288.
- Isenburg, M. (2013). LAStools-efficient tools for LiDAR processing. Version 161114.
- Klasner, F. L., & Fagre, D. B. (2002). A Half Century of Change in Alpine Treeline Patterns at Glacier National Park, Montana, U.S.A. *Arctic, Antarctic, and Alpine Research, 34*(1), 49-56.
- Kolecka, N., Kozak, J., Kaim, D., Dobosz, M., Ginzler, C., & Psomas, A. (2015). Mapping Secondary Forest Succession on Abandoned Agricultural Land with LiDAR Point Clouds and Terrestrial Photography. *Remote Sensing, 7*(7), 8300-8322.
- Korner, C. (2012). Treelines will be understood once the functional difference between a tree and a shrub is. *Ambio, 41 Suppl 3*, 197-206.
- Körner, C. (2012). *Alpine treelines: functional ecology of the global high elevation tree limits*: Springer Science & Business Media.
- Kullman, L., & Öberg, L. (2009). Post-Little Ice Age tree line rise and climate warming in the Swedish Scandes: a landscape ecological perspective. *Journal of Ecology, 97*(3), 415-429.
- Luckman, B., & Kavanagh, T. (2000). Impact of climate fluctuations on mountain environments in the Canadian Rockies. *Ambio: A journal of the human environment, 29*(7), 371-380.
- Moiseev, P. A., & Shiyatov, S. G. (2003). Vegetation Dynamics at the Tree-Line Ecotone in the Ural Highlands, Russia. *Ecological Studies, 167*.
- Roush, W., Munroe, J. S., & Fagre, D. B. (2007). Development of a spatial analysis method using ground-based repeat photography to detect changes in the alpine treeline ecotone, Glacier National Park, Montana, USA. *Arctic Antarctic and Alpine Research, 39*(2), 297-308.
- Stockdale, C. A., Bozzini, C., Macdonald, S. E., & Higgs, E. (2015). Extracting ecological information from oblique angle terrestrial landscape photographs: Performance evaluation of the WSL Monoplotting Tool. *Applied Geography, 63*, 315-325.

- Tranquillini, W. (1979). *Physiological ecology of the alpine timberline. Tree existence at high altitudes with special reference to the European Alps.*: Springer-Verlag.
- Trant, A. J., Starzomski, B. M., & Higgs, E. (2015). A publically available database for studying ecological change in mountain ecosystems. *Frontiers in Ecology and the Environment*, 13(4), 187-187.
- Weiss, D. J., Malanson, G. P., & Walsh, S. J. (2015). Multiscale Relationships Between Alpine Treeline Elevation and Hypothesized Environmental Controls in the Western United States. *Annals of the Association of American Geographers*, 105(3), 437-453.

Appendix 2: Repeat Oblique Photography shows Terrain and Fire-Exposure Controls on Century-Scale Canopy Cover Change in the Alpine Treeline Ecotone

McCaffrey D. and Hopkinson. C. 2020. REPEAT OBLIQUE PHOTOGRAPHY SHOWS TERRAIN AND FIRE-EXPOSURE CONTROLS ON CENTURY-SCALE CANOPY COVER CHANGE IN THE ALPINE TREELINE ECOTONE. *Remote Sensing* 12(10): 1569

Repeat Oblique Photography shows Terrain and Fire-Exposure Controls on Century-Scale Canopy Cover Change in the Alpine Treeline Ecotone

David McCaffrey^{1*} and Christopher Hopkinson²

¹ University of Lethbridge; david.mccaffrey@alumni.uleth.ca

² University of Lethbridge; c.hopkinson@uleth.ca

* Correspondence: david.mccaffrey@alumni.uleth.ca

Abstract: Alpine Treeline Ecotone (ATE), the typically gradual transition zone between closed canopy forest and alpine tundra vegetation in mountain regions, displays an elevational range that is generally constrained by thermal deficits. At landscape scales, precipitation and moisture regimes can suppress ATE elevation below thermal limits, causing variability in ATE position. Recent studies have investigated the relative effects of hydroclimatic variables on ATE position at multiple scales, but less attention has been given to interactions between hydroclimatic variables and disturbance agents, such as fire. Advances in monoplotted have enabled the extraction of canopy cover information from oblique photography. Using airborne lidar, and repeat photography from the Mountain Legacy Project, we observed canopy cover change in West Castle Watershed (Alberta, Canada; ~103 km²; 49.3° N, 114.4° W) over a 92-year period (1914-2006). Two wildfires, occurring 1934 and 1936, provided an opportunity to compare topographic patterns of mortality and succession in the ATE, while factoring by exposure to fire. Aspect was a strong predictor of mortality and succession. Fire-exposed areas accounted for 83.6% of all mortality, with 72.1% of mortality occurring on south- and east-facing slope aspects. Succession was balanced between fire-exposed and unburned areas, with 62.0% of all succession occurring on north- and east-facing slope aspects. The mean elevation increase of closed canopy forest (i.e. the lower boundary of ATE) on north- and east-facing undisturbed slopes was estimated to be 0.44 m yr⁻¹, or ~44 m per century. The observed retardation of treeline advance on south-facing slopes is likely due to moisture limitation.

Keywords: Alpine Treeline Ecotone, Repeat Photography, Monoplotted, Lidar, Fire

1. Introduction

Alpine treeline ecotone (ATE), the transition zone between closed canopy forest and alpine tundra, occurs where tree growth is limited by thermal thresholds at high elevations [1,2]. At continental scales, the approximate elevation of ATE is predicted by the *growth limitation hypothesis*, which suggests an inverse relationship exists between latitude and ATE elevation [1] as a function of temperature in the growing season [2].

However, numerous modulating factors have been described, which suppress treeline below this thermal limit (for detailed list, see [3]). The relative effects of these modulating factors on ATE elevation vary by spatial scale [4-6]. At landscape scales, ATE is suppressed below thermal limits by geomorphic conditions [7,8], physiological stressors (e.g. moisture limitation [9], wind exposure [10]), and disturbances such as rockslides [7], avalanches [11], and fires [12]).

A meta-analysis [13] identified 103 publications which describe long term records of treeline position (both ATE and latitudinal treeline) at 166 sites around the world, between 1900-2008. Approximately 52% of these sites have some form of treeline advance. Two studies report treeline recession and both cases are

associated with disturbances. Linear regression with climate records reveals that many of the advances in ATE position reported in [13] are associated with increases in atmospheric temperature.

Modulating factors, which suppress ATE below thermal limits regardless of rising atmospheric temperature, are often studied in isolation, and the potential for interaction between modulating factors is rarely considered. Our research investigated the interaction between two of these modulating factors, fire disturbance and moisture limitation, and their combined influence effect on ATE position. Using a novel monoplotted technique to extract quantitative spatial data from a historic photographic record, we investigated whether postfire moisture limitation caused topographic patterns of vegetative regeneration, thereby selectively suppressing ATE in dry, fire-exposed environments.

Fire and Topography in the ATE

Fires are less frequent in the ATE than at lower elevations, and fire return intervals (FRIs) are positively correlated with elevation [14,15]. While relatively infrequent, fire remains an important disturbance agent in the ATE [16,17], and the effects of fire on subalpine regions are potentially increasing with climate change [18].

Topographic factors like aspect and slope influence many of the processes that affect fire behavior in mountain landscapes. Topography can alter precipitation regimes, species composition, lightning strike frequency, and drying caused by both solar insolation and wind [14]. For example, in North America, south-facing mountain aspects receive more solar insolation and are generally drier than north-facing aspects. This impacts fire density (i.e. number of fires per unit area), with greater density observed on south-facing aspects [19-21]. However, the increase in fire density on warm aspects does not translate to an increase in area burned on south-facing aspects [22,23]. Thus, small fires are ignited more frequently on south-facing aspects, but a large fire will burn a range of topographic conditions, without aspect bias.

Patterns of moisture limitation also have the potential to impact of postfire regrowth, potentially suppressing ATE elevation below levels that would be expected if ATE was only limited by atmospheric temperature. At high elevations in the Rocky Mountains, postfire areas are predominantly colonized by spruce or fir. The initial postfire colonization in these areas can take 20-100+ years [24-27], leaving alpine meadows that persist for a century or more. Postfire rates of succession in subalpine forest can be twice as high on mesic slopes as on xeric slopes [28], and growth of subalpine fir has been demonstrated to be limited on south-facing aspects [29,30]. If delayed regrowth can inhibit colonization for almost a century, then the present position of fire-exposed treelines on south-facing aspects may be temporarily suppressed, appearing lower than would be expected if ATE were solely limited by atmospheric temperature. While stand-age reconstructions are useful at identifying the processes involved in postfire regrowth, few direct observations of an aspect effect on postfire regrowth have been made using imagery.

ATE Observation Techniques

Determining the interaction between aspect and regrowth in postfire environments using imagery requires high resolution observations that extend more than a century. These are uncommon. A trade-off exists between spatial resolution and temporal extent in ATE observation methods [31]. Multitemporal remote sensing can monitor vegetative land cover change at spatial resolutions required to correlate with topographic variables (10-100 m) [17], but observations of ATE using passive optical satellite imagery, like Landsat TM [11,32,33] only extend back to the 1970-80s. Aerial photographs are similarly useful for treeline change monitoring [17,34], but aerial photography records only extend to the late 1940s and early 1950s for much of North America.

Conversely, ATE observation records with very long temporal extents often have poor spatial resolution, prohibiting topographic analysis. For example, palynological records extend thousands of years before present, with some exceeding 10,000 years [35], but these observations can only provide a spatial resolution equivalent to the pollen distribution distance (<100 m). Dendrochronological records have offered the best combination of high spatial resolution (<10 m) and high temporal extent (~1000 yr). Tree-ring records are often used in ATE studies [9,36-38]. The spatial resolution and temporal extent of dendrochronological studies are both high, but the spatial extent is limited by onerous collection methods and the large sample size required to cover enough sites for topographic analysis [9,15].

The need for ATE observations that provide high spatial extent and resolution and long temporal extent can be addressed with repeat photography. This method is widely used to assess land cover change in the ATE. Patterns of ATE advance in eastern slope regions of the Canadian Rockies have been described using repeat photography [39-41]. High resolution photographs were used to identify shrub advance in the Swedish Scandes [42]. Repeat photography has also been used for a description of land cover change in the Ural Mountains [43]. While the temporal extent of repeat photographs, 100 years or more in cases, is useful for studies of change in the ATE, the method has been mostly restricted to qualitative analysis, given the previous inability to standardize spatial scale across oblique imagery. Quantitative spatial analysis of ATE has been attempted by draping a polygon *fishnet* over repeat photographs [44], in order to standardize the observation units in both historical and contemporary photographs, but spatial scale still varied between polygons, limiting the usefulness of this technique for spatial analysis.

Recent technological advances have enabled fully quantitative spatial analysis of repeat photographs, using the WSL Monoplotting Tool (WSL-MT, note: WSL is the German acronym for “Eidgenössische Forschungsanstalt für Wald, Schnee und Landschaft”, the Swiss Federal Institute for Forest, Snow and Landscape Research), which produces georeferenced vector data from oblique photographs using tie points between oblique and aerial photographs, and high resolution topographic data [45]. The WSL-MT tool has been used to monitor land cover change by extracting spatial data from oblique repeat photographs [46] and to observe century-scale vegetation transitions in the Rocky Mountains of southern Alberta, Canada [47]. These experiments were conducted on photo pairs from the Mountain Legacy Project, a collection of over 120,000 historic survey images of the Canadian Rockies (1888 – 1958), of which over 6,000 have 21st-century repeat photographs available under creative commons license [48]. Previous research, validated using airborne lidar, established that fractional cover classes could be assessed from oblique imagery at a spatial resolution of 20 m, using the WSL-MT [49].

Research Goals

We used a technique to assess fractional cover from oblique photographs, described in [49], to observe postfire vegetative change in repeat photographs of a watershed in the Canadian Rockies. We sought to determine: 1) if there were observable differences in canopy cover change between fire-exposed and non-fire-exposed regions of the watershed, 2) if changes in canopy cover correlate to topographic patterns that are consistent with expected edaphic controls on regrowth (e.g. less regrowth on south-facing aspects), and 3) if it was possible to detect the rate of ATE advance in non-fire-exposed areas, where ATE position is regulated by climate.

2. Materials and Methods

Study Area

The study was conducted over the West Castle Watershed (WCW), Alberta, Canada (49.3° N, 114.4° W). WCW is in the headwaters of the Oldman River, on the eastern slopes of the Canadian Rockies, with an area of ~103 km², and an elevation range of ~1400-2600 m a.s.l. Aside from a small ski resort and village near the downstream end of the WCW (2.9% of the watershed area), and trails in the valley bottoms, the forested slopes demonstrate limited anthropogenic disturbance. Consequently, the WCW is an ideal study area for assessing recent natural shifts in the ATE of this part of the Canadian Rockies. Dominant species at ATE elevation in WCW are subalpine fir and Engelmann spruce.

Repeat Photographs

Repeat photographs of WCW were provided by the Mountain Legacy Project [48] (Figure 1). The historical images were taken in the summer of 1914 by Morrison Parsons Bridgeland. The repeat photographs were taken by the Mountain Legacy Project in July of 2006 (Table 1). Photograph pairs were selected to maximize the spatial coverage of the observed area of WCW. Using the same set of photographs, it was determined that 38.5 km² (37.7%) of the watershed was observable in these images [49] (Figure 2).

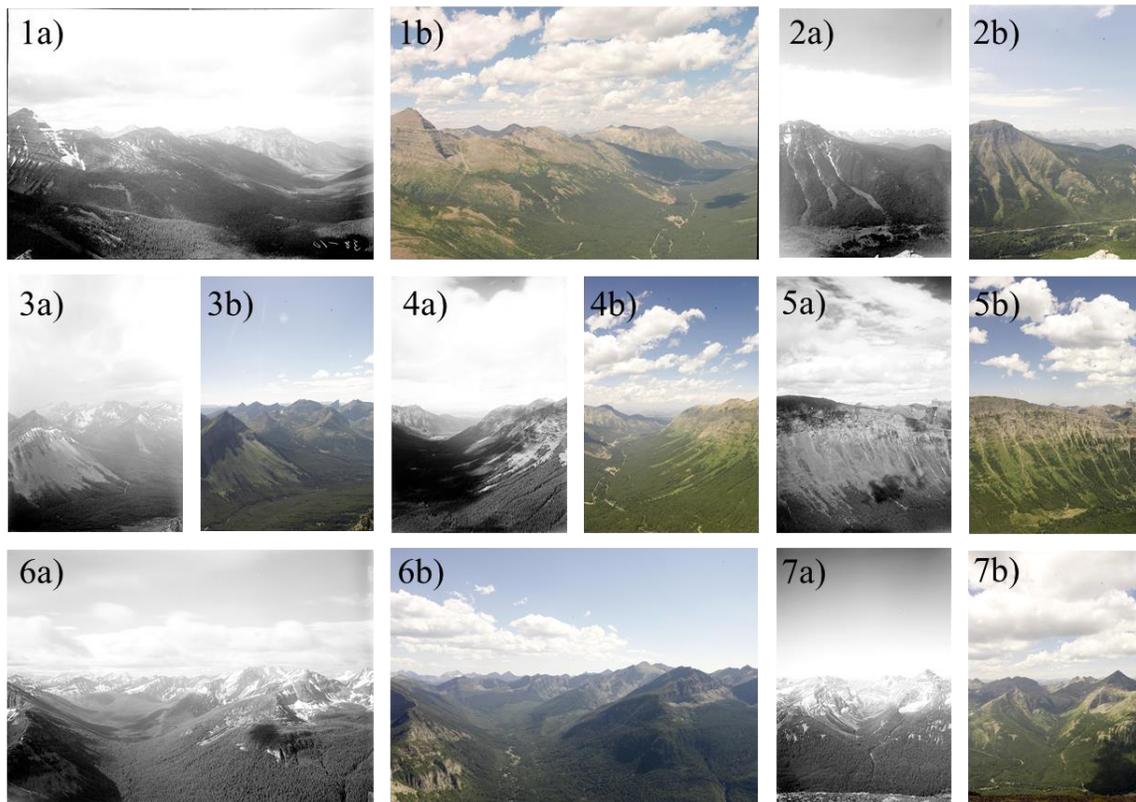


Figure 1. Seven repeat photograph pairs of WCW; a-series by M.P. Bridgeland, 1914, courtesy of Library and Archives Canada; b-series copyright Mountain Legacy Project, 2006.

Table 1. Collection data for Mountain Legacy Project 2006 photographs. Information for the corresponding 1914 data can be found at www.mountainlegacy.ca.

| Photograph Number | Elevation (m a.s.l.) | Latitude (N) | Longitude (W) | Photograph Date | MLP Photo ID (2006) |
|-------------------|----------------------|----------------|-----------------|-----------------|---------------------|
| 1 | 2295 | 49° 16' 3.62" | 114° 22' 58.97" | 28 July 2006 | B0032P0048 |
| 2 | 2216 | 49° 20' 14.00" | 114° 23' 22.81" | 29 July 2006 | B0033P0005 |
| 3 | 2390 | 49° 17' 7.01" | 114° 20' 44.21" | 28 July 2006 | B0032P0073 |
| 4 | 2283 | 49° 16' 4.94" | 114° 22' 56.77" | 28 July 2006 | B0032P0056 |
| 5 | 2283 | 49° 16' 4.94" | 114° 22' 56.77" | 28 July 2006 | B0032P0055 |
| 6 | 2390 | 49° 17' 7.01" | 114° 20' 44.21" | 28 July 2006 | B0032P0075 |
| 7 | 2462 | 49° 18' 7.26" | 114° 22' 18.63" | 30 July 2006 | B0034P0005 |

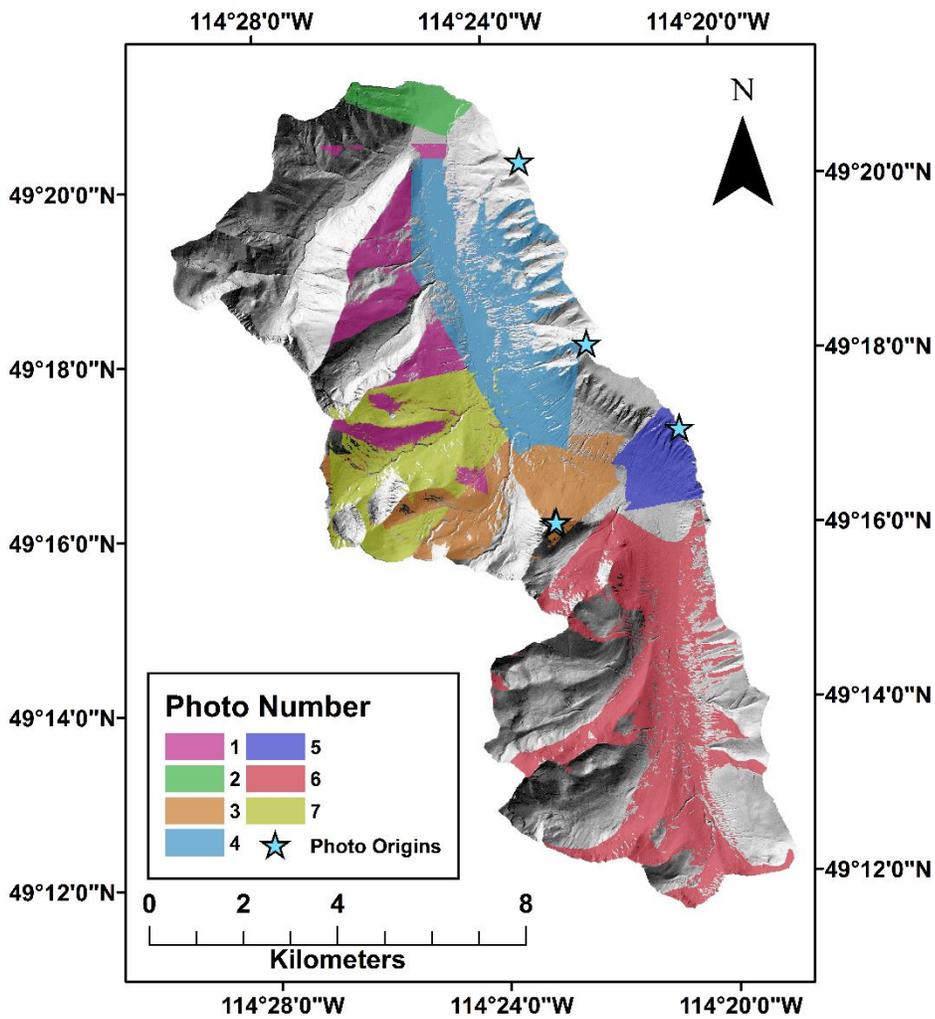


Figure 2. Visible region of WCW in each of the seven MLP photographs. In cases where the same region is visible in multiple photographs, the photograph used for analysis was determined by criteria described in [49].

Canopy Classification

Canopy cover was extracted from oblique photography with the WSL monoplottting tool [45], using a method described in [49]. The WSL-MT performs oblique to orthogonal vector transformation with a camera calibration, which generated by pairing control points between oblique photographs and a high resolution DEM, co-registered with aerial imagery. For the DEM requirement, a 1 m DEM was used, which was collected from an airborne lidar survey in October 2014. For the aerial imagery requirement, a SPOT 6 image of WCW was used (1.5 m, acquired July 31, 2014, © 2014 CNES). In the absence of high resolution topographic and aerial data from 1914, the analysis relied on the camera calibration from the 2014 lidar DEM and 2006 SPOT image being applied to the 1914 oblique photographs. This approach was deemed reasonable, as no extreme changes in topography (e.g. rock slides, slumping) are apparent between the 1914 and 2006 datasets.

The monoplottting procedure results in a *fishnet* vector layer, where each cell represents a 20 x 20 m area (viewed orthogonally). This vector layer is draped over the oblique raster image, such that cells are scaled according to the perspective in the image (i.e. cells in the foreground appear larger, and cells in the background appear smaller) (Figure 3). This method circumvents the problem of varying spatial scale in the oblique image, allowing quantitative spatial analysis to be performed on the resulting raster.

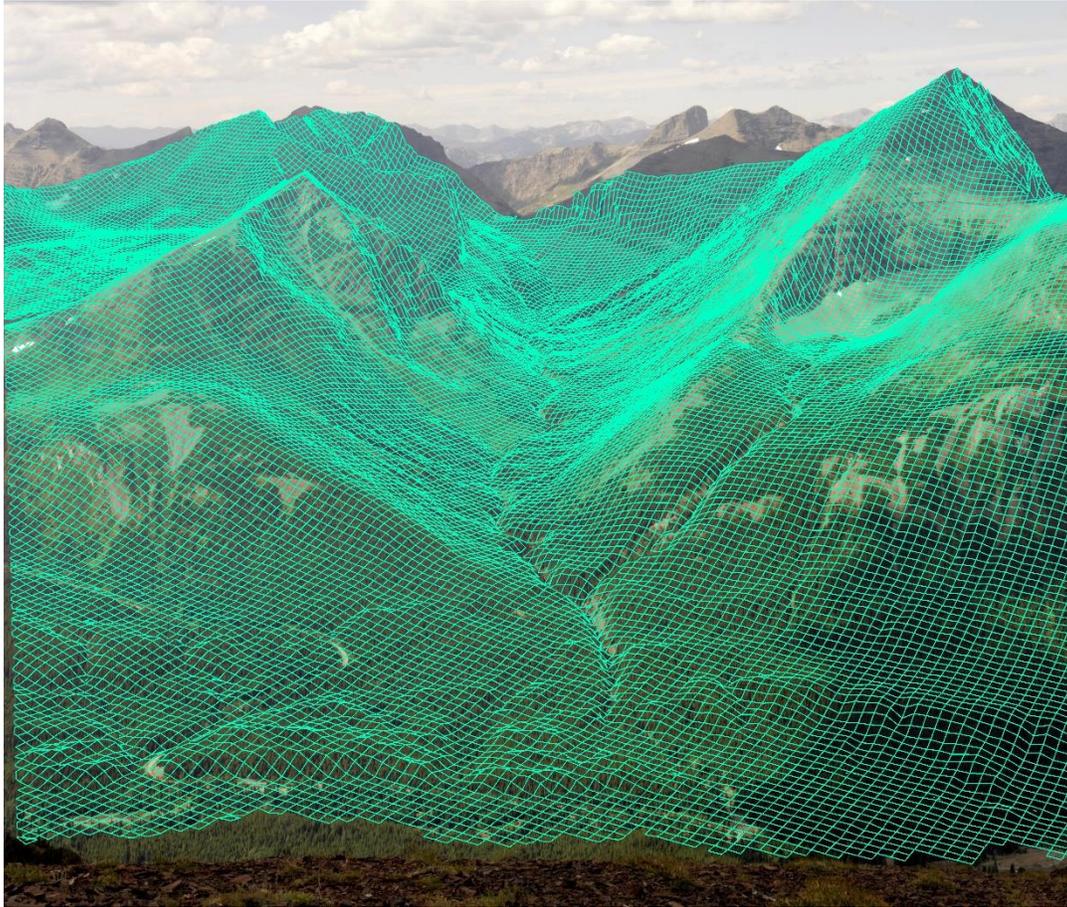
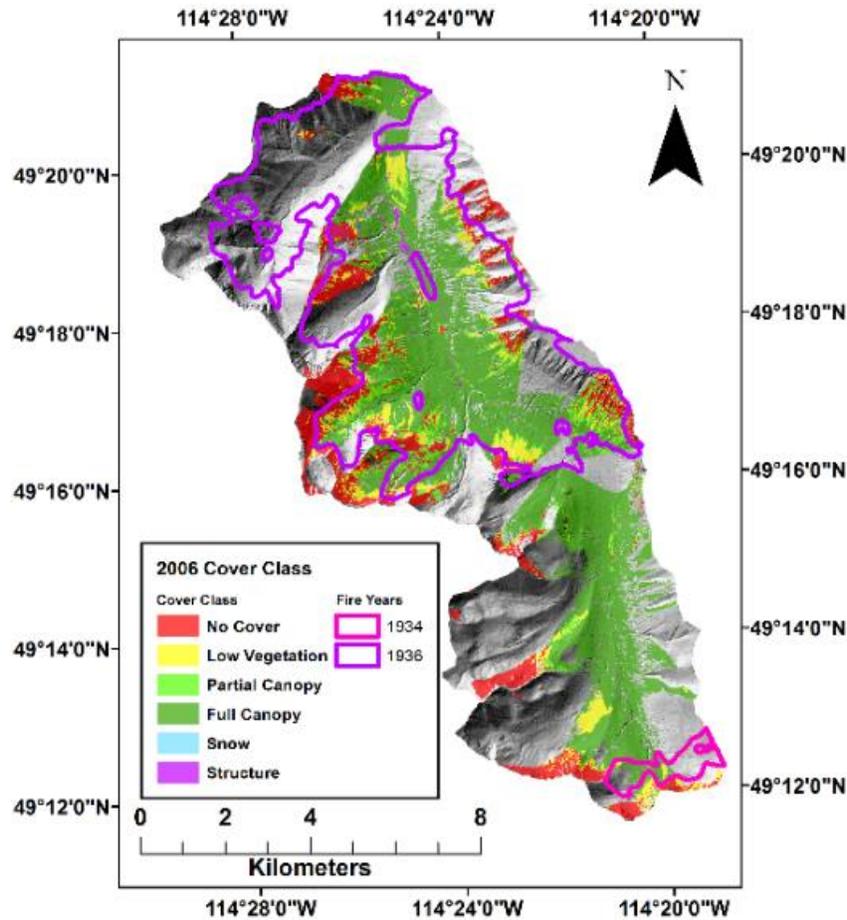


Figure 3. Fishnet vector layer, developed using the WSL monoplottting tool, draped on an oblique image. Each cell in the fishnet represents a 20 x 20 m area, viewed orthogonally.

Calibrated fishnets were applied to each of the seven oblique images. Any fishnet grid cell that intersected with less than 75% of the viewshed (i.e. < 300 m²) was omitted from the analysis (note: because these data were eventually rasterized, and each raster represented an area between 300-400 m², canopy cover and change is reported by grid cell count rather than surface area).

Once in an oblique projection, grid cells were manually assigned into one of six canopy cover classes, based on observed canopy openness and texture: 1) *No Cover* – grid cells devoid of vegetation; 2) *Low Vegetation* – grid cells appear vegetated, but context and texture suggest shrubs or krummholz, not upright trees > 2 m; 3) *Partial Canopy* – trees are present, but ground is visible in > 50% of the grid cell; 4) *Full Canopy* – trees cover > 50% of a grid cell; 5) *Snow* – snow covers > 50% of a grid cell; 6) *Structure* – any anthropogenic structure or non-vegetated land cover is present in the grid cell (e.g. buildings, roads, and trails) (Figure 4).

a)



b)

Figure 4. Canopy cover classification for WCW using oblique photography from a) 1914 and b) 2006. Each cell is a 20 x 20 m area, viewed orthogonally, where canopy cover has been assessed, ranging from *No Cover* to *Full Canopy*, and including cells with snow cover or modern structures. The extents of the 1934 and 1936 fires are provided for reference [50].

Given the focus on vegetative land cover change, further analysis focused on the first four, ordinal classes of vegetation (*no cover* through *full canopy*). Any grid cells having *snow* or *structure* in either 1914 or 2006 were omitted from analysis. The *snow* class was removed due to uncertainty of vegetation beneath the snow. The anthropogenic *structure* class was omitted because the processes causing vegetative land cover to change were distinct from the processes being studied. Of the total observed area, 7.2% was affected by the removal of snow cover, and 2.9% by the removal of anthropogenic structures.

Anthropogenic Disturbance

The Alberta Vegetation Inventory [51] was used to identify and omit areas of anthropogenic disturbance. It was produced from interpretation of aerial photographs, and identifies several types of

former and current disturbance, including former oil well pads, historical cut-blocks, settlements (including the Castle Mountain Ski Resort), and roads. These areas were aggregated into a single layer of anthropogenic disturbance so that topographic and other factors which affected land cover change could be considered independently of human activity.

Fire Disturbance

Two major historical fires occurred in the interval between the oblique photograph observations, one in 1934 and one in 1936. The extents of the historical fires were delineated in 2005, using a combination of visual evidence from a 1949 aerial photograph survey and descriptions from historical fire reports [50] (Figure 4). More detailed attributes of these fires, such as intensity and cause, are unknown, but previous research has demonstrated that fire intensity does not co-vary with topographic features such as aspect and slope in subalpine fir-spruce forests of the Rockies, as fires that reach these elevations are typically high intensity, crown replacing fires [22,23]. Thus, if we observe topographic differences in patterns of postfire regrowth, they are likely related environmental factors that affect regrowth, and not likely caused by topographic differences in fire intensity.

Fire suppression in early 20th-century was common practice, yet it is known that the Government of Alberta did not engage in fire suppression in the 1930s and 1940s, for reasons related to the Great Depression and the Second World War [52]. Fire suppression techniques did not affect the fire return interval of subalpine forests in the region [52]. Mean fire return intervals in the WCW range from 19.3 years near the mouth of the valley, to 70.8 years in the southern headwaters [53].

Topographic Analysis of Change

The four ordinal classes of canopy cover (1 - *no cover*, 2 - *low vegetation*, 3 - *partial canopy*, 4 - *full canopy*) were used to categorize vegetative change between 1914 and 2006. For each 20 x 20 m grid cell, the difference between the 2006 cover class and the 1914 cover class was used to generate a change class (Figure 5). For example, if canopy cover in 2006 was *no cover* (1), and canopy cover in 1914 was *full canopy* (4), then the change class for that area was $1 - 4 = -3$. The seven possible change classes were: -3 (high mortality), -2 (medium mortality), -1 (low mortality), 0 (no change), +1 (low succession), +2 (medium succession), +3 (high succession).

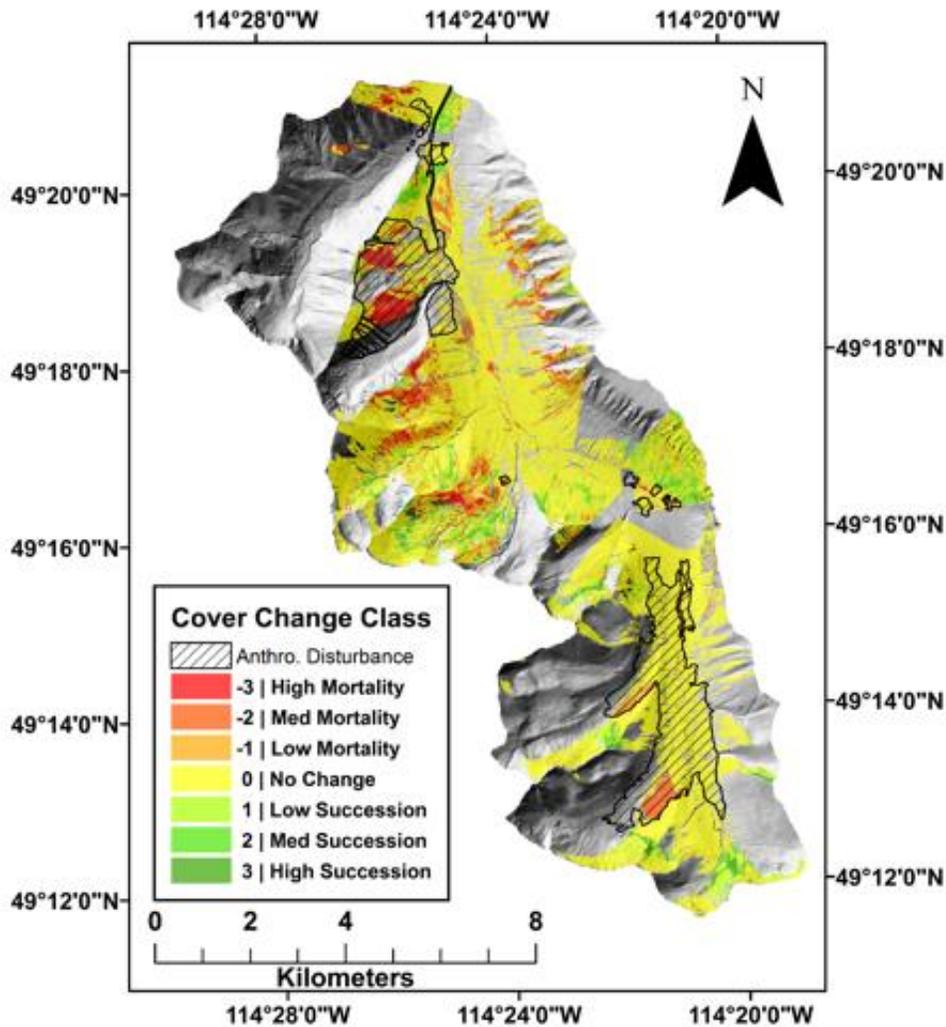


Figure 5. Cover change classification for period 1914-2006. Large areas of anthropogenic disturbance, such as a ski hill in the North of the watershed and a historic cut block in the South of the watershed, were omitted from analysis.

In preparation for the topographic analysis, a 1 m DEM was aggregated to 20 m and coregistered to the change classes: for elevation, the mean value of the 1 m cells was used for its corresponding 20 m raster; for slope, the degree angle of slope was calculated from the 1 m DEM, and the mean of degree values was used for the 20 m raster; for aspect, azimuth was calculated on the aggregated 20 m DEM, and discretized in eight intercardinal direction bins of 45° each: N = 337.5-22.5°, NE = 22.5-67.5°, E = 67.5-112.5°, SE = 112.5-157.5°, S = 157.5-202.5°, SW = 202.5-247.5°, W = 247.5-292.5°, NW = 292.5-373.5°.

Data were factored by fire exposure, with the 1934 and 1936 fire extents aggregated into a single layer. The distribution of topographic variables within each change cover class was contrasted between fire-exposed and non-fire-exposed areas in the time interval 1914-2006.

For aspect, we tested if the proportional distribution of aspects within each change class differed from expected aspect distribution in a way that suggested aspect influence on canopy cover. The 20 m DEM was sampled to establish the expected proportional distribution of aspect classes in the WCW. The observed proportional distribution of aspect within each change class was compared to the expected distribution with a chi-square goodness-of-fit test. To test the significance to each comparison, a Monte Carlo analysis using 100 repetitions was used to build a sampling distribution.

For elevation and slope, we tested if the distribution of these variables differed in between comparisons of change class. One-way ANOVA tested whether any of the seven classes' mean values differed from the others. For both elevation and slope comparisons, Welch's t-test was used on each of the 20 pairwise combinations of seven change classes to establish which classes differed from each other. Note that Welch's t-test corrects for both unequal variance between distributions and unequal sample size [54]. A Bonferroni correction was applied to the t-tests to account for multiple comparisons. All topographic analyses were completed solely outside of the anthropogenic disturbance area.

Finally, the rate of upslope ATE advance in undisturbed areas was estimated along five transects of canopy change cells. Along each transect, the boundary of closed canopy forest was identified in the 1914 and 2006 scenes. The closed canopy forest boundary is the lower extent of ATE (sometimes referred to as timberline). Given that the upper boundary (species limit) could vary within a 20 m grid cell, and may have been obstructed by snow, the closed canopy forest boundary was used as a proxy for total ATE change. Transects were selected where a distinct timberline (lower ATE boundary) could be observed across at least five grid cells (100 m), in areas without orographic boundaries to treeline advance (i.e. climatic treelines), and where there was no exposure to fire between observations. Five of these transects were identified. Elevation was sampled from each of the 20 m grid cells, and the mean elevations of the closed canopy fronts in the 1914 and 2006 scenes were compared.

3. Results

3.1 Aspect

Observed distribution of aspect significantly differed from expected distribution, with the majority of mortality seen on fire-exposed, warm aspects (Figure 6). In total, fire affected 63.2% of the observed area in WCW and 83.6% of all mortality occurred in fire-exposed areas. A strong interaction between aspect and fire was observed, with 72.0% of all mortality occurring on SW, S, SE, and E aspects. There was a high occurrence of succession classes in N and NE aspects, in both fire-exposed and non-fire-exposed areas. Succession was also notably high in fire-exposed SW aspects, and non-fire-exposed NE aspects. In total, north- and east-facing slope aspects accounted for 62.0% of all succession.

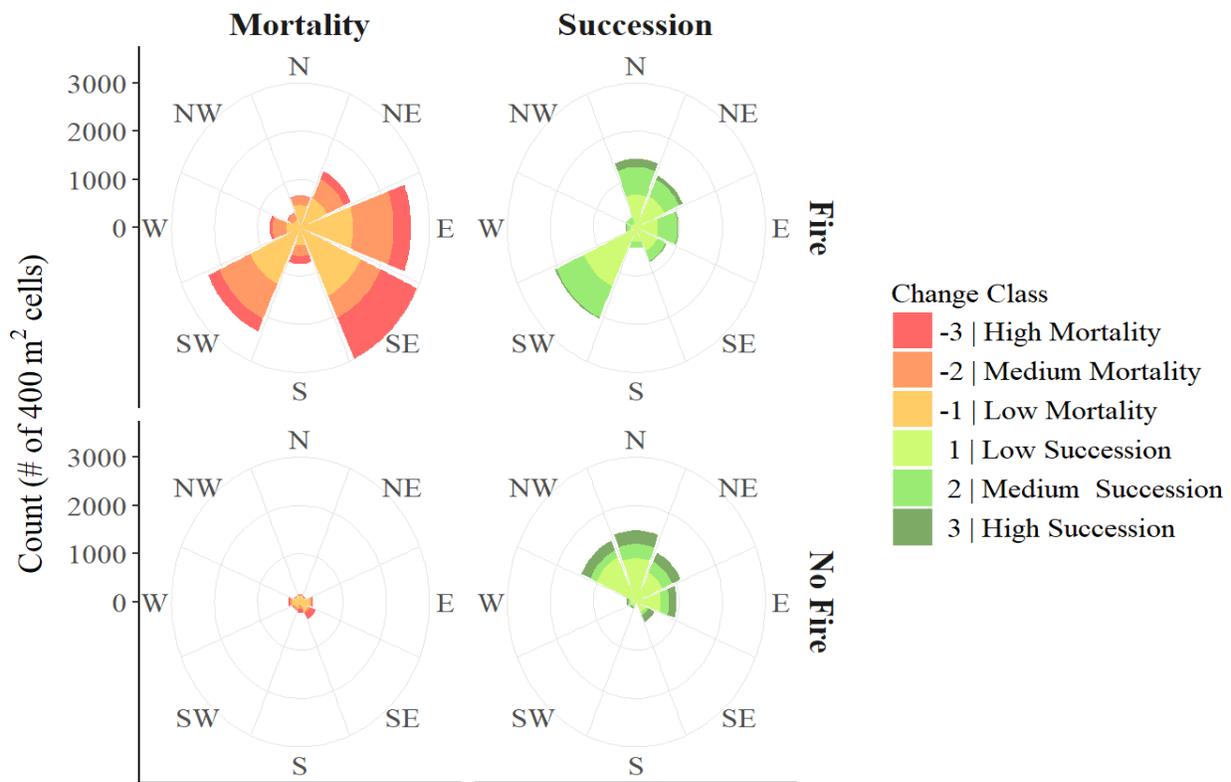


Figure 6. Mortality and succession in WCW over the period 1914-2006, factored by exposure to fire. Mortality is largely restricted to fire exposed areas, and is predominantly in the SW, SE, and E aspects.

The expected proportion distribution of aspect in WCW varied, as a function of the orientation of the valley; SW, E, and NE aspect had higher occurrence, while S, SE, and NW aspects had lower occurrence (Figure 7). We compared this expected proportion of aspects for each of the seven change classes (8 aspects \times 7 change classes = 56 comparisons). The results of the chi-squared test show that significant differences between expected and observed aspect proportion occurred frequently; 46 out of 56 (82.1%) of the change class/aspect cases showed significant differences (Table 2). The greatest departures from expected proportions were seen in high occurrences of mortality on S and SE aspects, and high occurrences of succession on N and NE aspects.

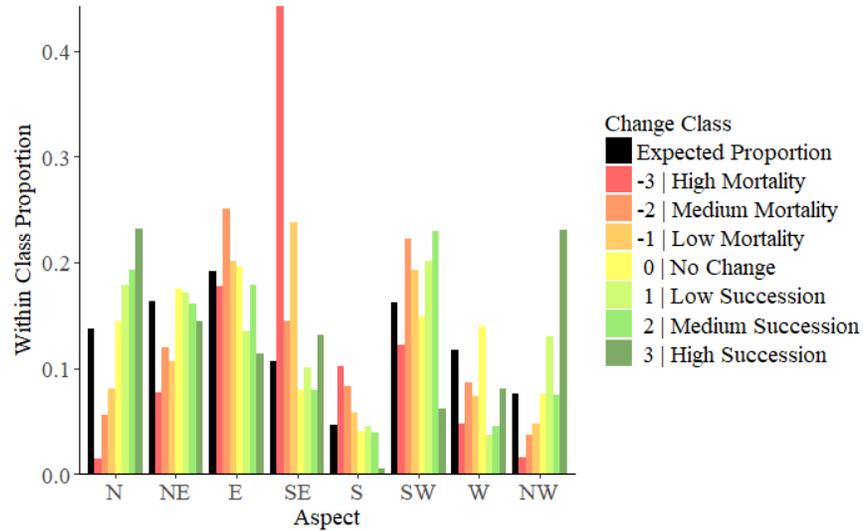


Figure 7. Expected proportion of aspects within the WCW (in black) compared to the observed proportion of aspect for each change class. For example, roughly 0.1 (or ~10%) of the land in WCW is on a SE aspect (black bar). If there were no aspect effect, we would expect the same proportion of high mortality cases to be on SE aspects. Instead, > 0.4 (or >40%) of high mortality occurs on SE aspects (red bar).

Table 2. Results of chi-squared test of aspect proportions. Significant values in bold. Significant differences in aspect proportion were seen in 43 out of 56 of the tests.

| Aspect | | Change Class | | | | | | |
|--------|----------|----------------|------------------|---------------|-----------|----------------|-------------------|-----------------|
| | | High Mortality | Medium Mortality | Low Mortality | No Change | Low Succession | Medium Succession | High Succession |
| N | χ^2 | 307.5 | 212.1 | 196.8 | 28.6 | 101.4 | 90.0 | 59.2 |
| | P | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |
| NE | χ^2 | 130.1 | 52.6 | 171.7 | 64.5 | 3.4 | 0.1 | 1.9 |
| | P | < 0.001 | < 0.001 | < 0.001 | < 0.001 | 0.066 | 0.728 | 0.163 |
| E | χ^2 | 3.3 | 85.9 | 4.4 | 6.1 | 147.0 | 4.2 | 30.7 |
| | P | 0.069 | < 0.001 | 0.037 | 0.013 | < 0.001 | 0.041 | < 0.001 |
| SE | χ^2 | 2863.3 | 55.9 | 1327.0 | 530.7 | 2.6 | 27.5 | 5.1 |
| | P | < 0.001 | < 0.001 | < 0.001 | < 0.001 | 0.109 | < 0.001 | 0.024 |
| S | χ^2 | 177.1 | 121.3 | 26.7 | 34.4 | 0.1 | 3.3 | 29.6 |
| | P | < 0.001 | < 0.001 | < 0.001 | < 0.001 | 0.702 | 0.070 | < 0.001 |
| SW | χ^2 | 27.8 | 101.1 | 52.0 | 71.2 | 81.5 | 114.1 | 58.0 |
| | P | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |
| W | χ^2 | 114.1 | 34.8 | 136.6 | 311.9 | 434.0 | 169.8 | 10.2 |
| | P | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |
| NW | χ^2 | 127.0 | 81.2 | 81.9 | 0.0 | 301.4 | 0.1 | 268.0 |
| | P | < 0.001 | < 0.001 | < 0.001 | 0.963 | < 0.001 | 0.806 | < 0.001 |

3.2 Elevation

The distribution of elevation within change classes shows that the greatest canopy changes were seen in high elevation. *High mortality* and *high succession* classes had the highest mean elevation, while the lowest mean elevation was in the *no change* class (Figure 8). This result suggests that the degree of change, for both mortality and succession, covaried with elevation. To test this hypothesis, change classes were pooled by absolute value (e.g. -3 combined with 3, -2 combined with 2), and a Spearman rank test was performed. The results showed that degree of canopy change was significantly correlated with elevation ($r = 0.174$, $p < 0.001$), though the correlation was weak.

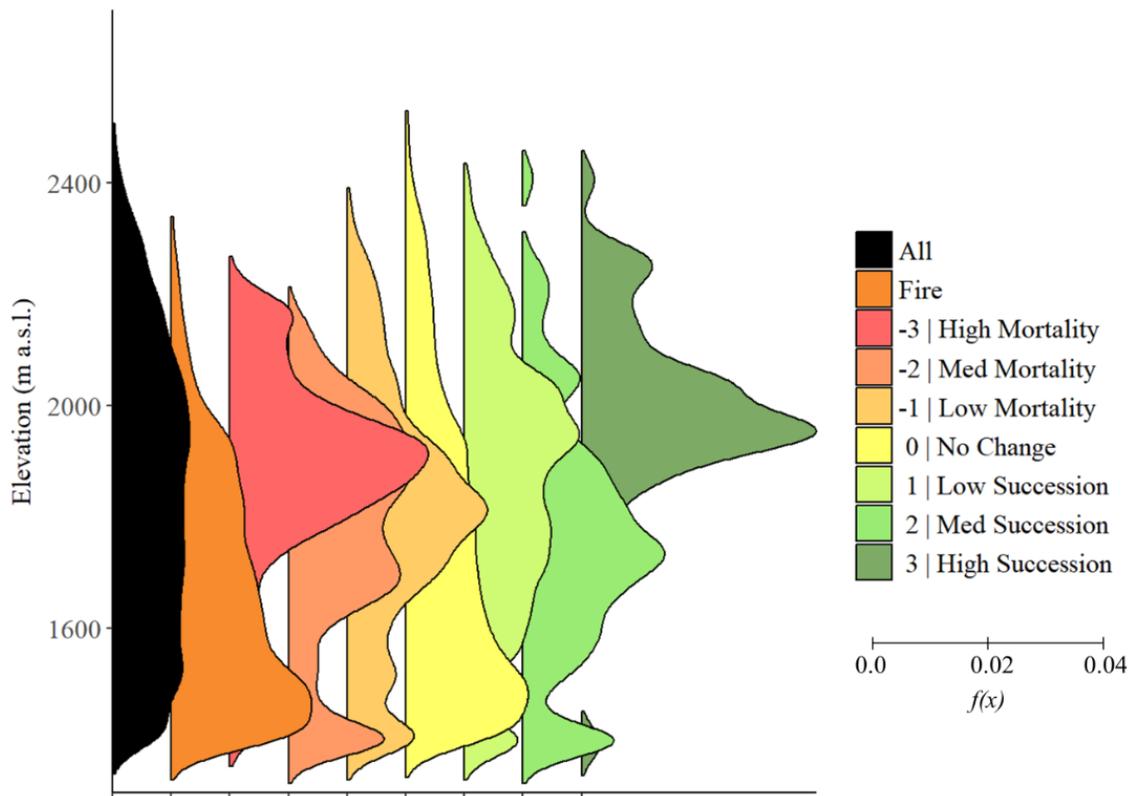


Figure 8. Elevation distributions in WCW, normalized with a probability density function. For each category, the hypsometric distribution of elevation within that category is displayed. The further right a curve extends from each baseline, the more likely it is to find values at that elevation within each category. Across all WCW, in black, elevation is evenly distributed, with elevation above 2200 m a.s.l. being less common. Fire, in orange, was more likely below 1600 m a.s.l., and high mortality and high succession were more likely at ATE elevations, between 1800-2100 m a.s.l. Corresponding t-tests can be seen in Table 3.

Between change class differences in elevation distribution were generally significant. One-way ANOVA showed that the mean elevation of at least one of the change classes was significantly different from the remaining groups ($F = 743.16$, $df = 6$, $p < 0.001$). The result of the pairwise t-tests showed that, of the 21 possible pairwise combinations, 20 showed significant differences (Table 3).

Mean elevation of fire-exposed areas was lower than mean elevation of non-fire-exposed areas in every change class (Figure 9). This is as expected, given the decreasing probability of fire with increasing elevation [15]. While factoring by fire showed between group differences in elevations, the significant correlation between degree of canopy change and elevation was preserved in both fire-exposed cases ($r = 0.340$, $p < 0.001$), and non-fire-exposed cases ($r = 0.110$, $p < 0.001$).

Table 3. Results of elevation t-test between change classes. Significant values in bold. Significant differences were seen in 20 out of 21 tests.

| | | Change Class | | | | | |
|------------------------------|---|-------------------|---------------------|-------------------|-------------------|-------------------|----------------------|
| | | High Mortality | Medium Mortality | Low Mortality | No Change | Low Succession | Medium Succession |
| Medium Mortality | t | -24.042 | - | - | - | - | - |
| | p | < 0.001 | - | - | - | - | - |
| Low Mortality | t | -20.153 | 6.643 | - | - | - | - |
| | p | < 0.001 | < 0.001 | - | - | - | - |
| No Change | t | -42.218 | -10.400 | -22.928 | - | - | - |
| | p | < 0.001 | < 0.001 | < 0.001 | - | - | - |
| Low Succession | t | -7.627 | 18.686 | 13.900 | 39.454 | - | - |
| | p | < 0.001 | < 0.001 | < 0.001 | < 0.001 | - | - |
| Medium Succession | t | -24.636 | -2.525 | -8.754 | 5.790 | -19.643 | - |
| | p | < 0.001 | 0.244 | < 0.001 | < 0.001 | < 0.001 | - |
| High Succession | t | 13.569 | 30.010 | 27.219 | 39.397 | 19.105 | 30.610 |
| | p | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |

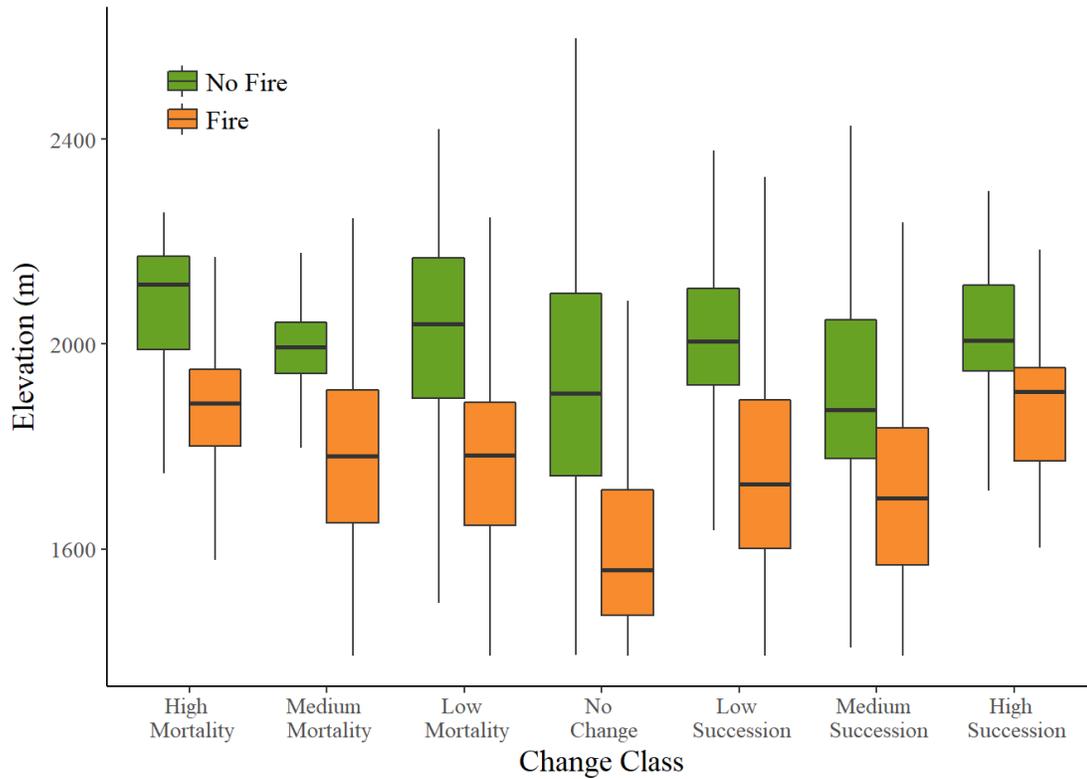


Figure 9. Distribution of elevation within change class, factored by fire exposure. In each canopy change class, median elevation is higher in areas that were not exposed to fire.

Slope

Significant differences in slope distribution were observed among change classes. The highest mean slope values were in the *high mortality* and *high succession* classes. The lowest mean slope value was in the *no change* category (Figure 10). As with elevation, the degree of change correlated positively with slope. Change classes were pooled by absolute value, and a Spearman rank test was performed. A weak correlation was found between slope and the degree of canopy change ($r = 0.126, p < 0.001$).

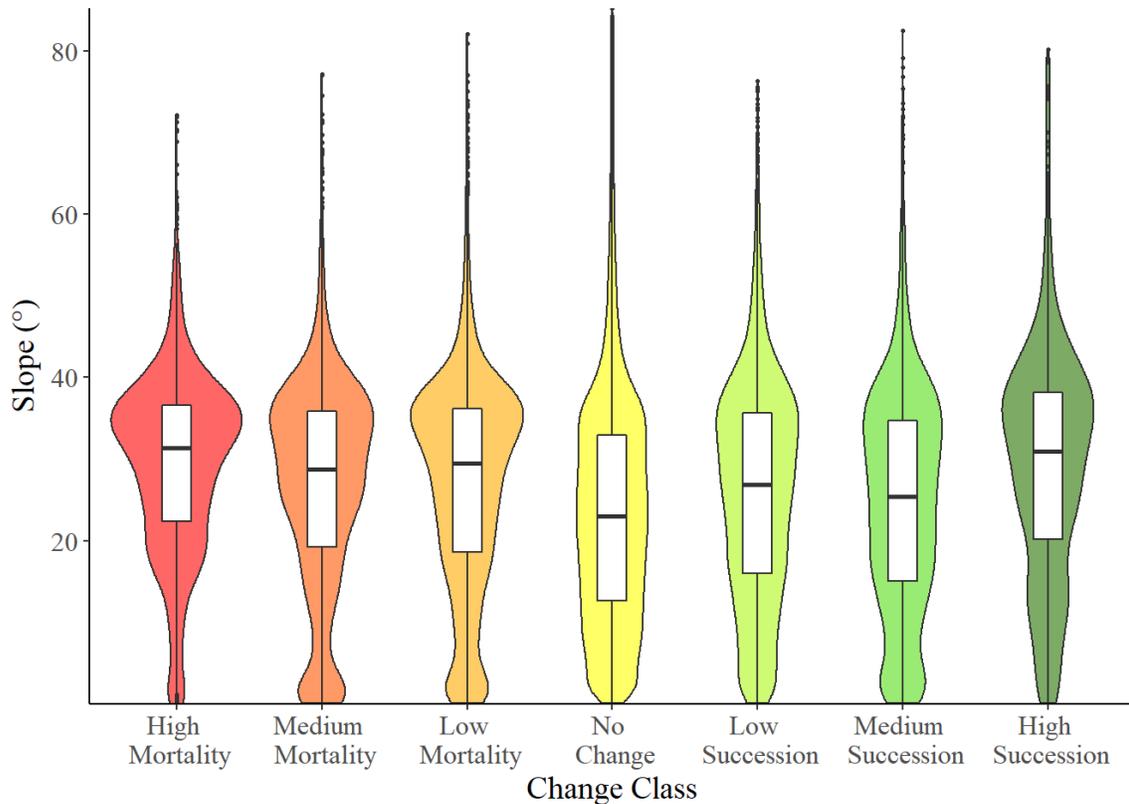


Figure 10. Distribution of slope within change class. The colored curves behind each box plot show a probability density function of the values in each distribution. Corresponding t-tests can be seen in Table 4.

The differences in slope distribution between change classes were generally significant. The slope distribution of at least one of the change classes was significantly different from the remaining groups (One-way ANOVA, $F = 213.69$, $df = 6$, $p < 0.001$). The result of the pairwise t-tests showed that, of the 20 possible pairwise combinations, 17 had significant differences (Table 4).

In a majority of change classes (-3 through 1), mean slope was lower in fire-exposed area than in non-fire-exposed areas (Figure 11). The hypsometry of this watershed shows a positive correlation between elevation and slope. Given that fires predominantly occurred at lower elevations, lower slope in fire-exposed areas was expected. In *medium* and *maximum succession* classes, fire-exposed areas had slightly higher mean slope than non-fire-exposed areas. Degree of canopy change and slope were significantly correlated in both fire-exposed cases ($r = 0.244$, $p < 0.001$), and non-fire-exposed cases ($r = 0.027$, $p < 0.001$).

Table 4. Results of slope t-test between change classes. Significant values in bold. Significant differences were seen in 18 out of 21 tests.

| Change Class | | | | | |
|----------------|------------------|---------------|-----------|----------------|-------------------|
| High Mortality | Medium Mortality | Low Mortality | No Change | Low Succession | Medium Succession |

| | | | | | | | |
|-------------------|---|---------|---------|---------|---------|---------|---------|
| Medium Mortality | t | -8.609 | - | - | - | - | - |
| | p | < 0.001 | - | - | - | - | - |
| Low Mortality | t | -8.727 | 1.097 | - | - | - | - |
| | p | < 0.001 | 1.000 | - | - | - | - |
| No Change | t | -24.947 | -14.348 | -21.092 | - | - | - |
| | p | < 0.001 | < 0.001 | < 0.001 | - | - | - |
| Low Succession | t | -12.396 | -2.926 | -4.793 | 13.790 | - | - |
| | p | < 0.001 | 0.072 | < 0.001 | < 0.001 | - | - |
| Medium Succession | t | -14.204 | -6.250 | -8.061 | 4.742 | -4.204 | - |
| | p | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 | - |
| High Succession | t | -0.064 | 4.710 | 4.350 | 11.106 | 6.297 | 8.120 |
| | p | 1.000 | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |

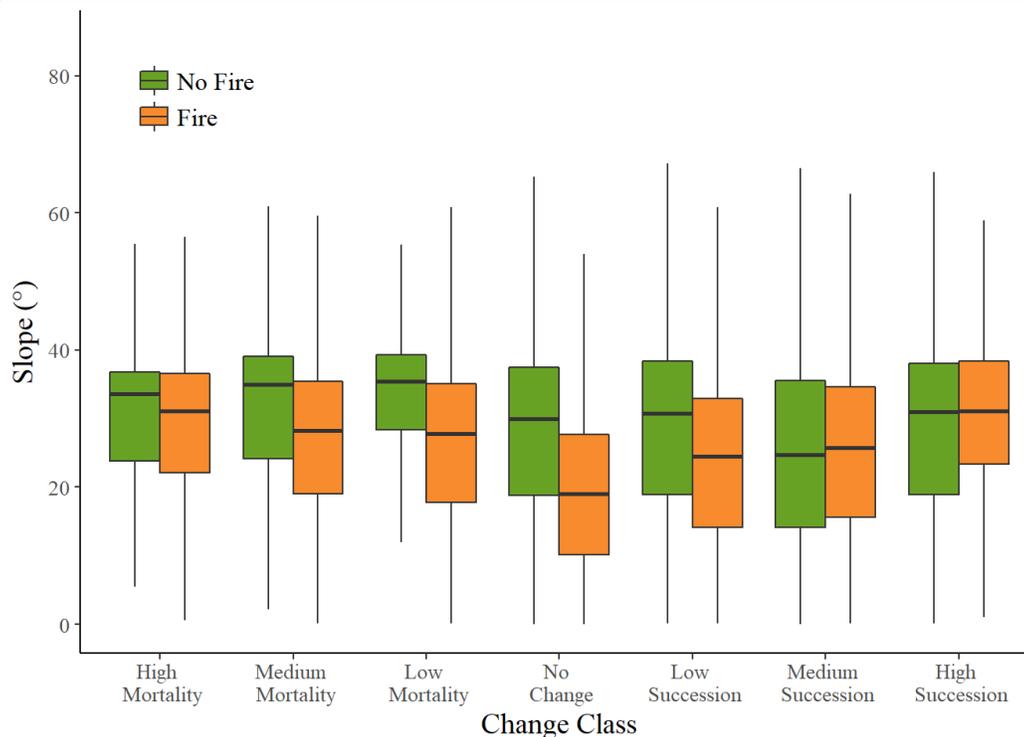


Figure 11. Distribution of slope within change class, factored by fire exposure. Median slope was generally higher in regions that were not exposed to fire.

Upslope Advance

Of the five closed canopy transects used to determine upslope elevation advance, all occurred in areas where Engelmann spruce was the dominant species in 2006. In each of these transects, mean elevation was higher in 2006 than in 1914. The mean elevation of closed canopy forest in these unburned areas increased by 40.5 m over 92 years, which translates to an annual increase of 0.44 m.

Table 5. Differences in mean elevation, between 1914 and 2006, across transects of cells in unburned areas, which represent the advanced of timberline, the lower boundary of ATE.

| Transect | Number of Cells 1914 | Mean Elevation 1914 (m a.s.l.) | Number of Cells 2006 | Mean Elevation 2006 (m a.s.l.) | Elevation Difference (m a.s.l.) |
|-----------------|-----------------------------|---------------------------------------|-----------------------------|---------------------------------------|--|
| 1 | 6 | 2046.0 | 8 | 2053.5 | 7.5 |
| 2 | 7 | 1894.6 | 9 | 1897.0 | 2.4 |
| 3 | 35 | 2002.9 | 28 | 2028.5 | 25.6 |
| 4 | 24 | 1945.1 | 24 | 2033.5 | 88.4 |
| 5 | 29 | 1876.1 | 31 | 1954.7 | 78.6 |
| Mean | 20.2 | 1952.9 | 20.0 | 1993.4 | 40.5 |

4. Discussion

Our study demonstrates an example of topographic patterns of postfire colonization that conforms to the edaphic process model of postfire colonization, using direct quantitative observation with monoplotted on repeat photography. In fir-spruce forests, postfire regrowth in high elevation regions, like the ATE, likely has edaphic controls. As a result, colonization in xeric areas, like south-facing slopes, can lag several decades or even centuries behind regrowth in mesic areas, like north-facing slopes. These edaphic controls on colonization were generally suggested based on research using stand level age-reconstructions, but given the long gap in colonization between mesic and xeric colonization, few direct observations of this effect have been made in imagery. By increasing the temporal range of observation using monoplotted on historic repeat photography, the topographic differences in postfire colonization are now visible in imagery.

Aspect

Greater than expected proportions of succession occurred on N and NW aspects, while higher than expected proportion of mortality occurred on S and SE aspects. Correlation between aspect and canopy cover change was associated with fire-exposure, suggesting soil moisture stress levels may have been a significant barrier to postfire recruitment.

It is known that seedling recruitment of subalpine fir and Engelmann spruce is inhibited on warm-aspects in the ATE, primarily as a function of soil moisture stress caused by high temperature and solar radiation [55]. This process has been demonstrated experimentally: heating experiments in trees at ATE elevations show that moisture limitation both inhibits conifer seedling recruitment [56] and reduces carbon gain [57]. In a postfire environment, insolation would increase due to reduced interception and increased gap fraction in the burned canopy. This increased insolation could have caused aspect specific mortality on warm aspects in WCW, particularly the SW aspect. Postfire erosion may also contribute to mortality. A study conducted in a watershed adjacent to WCW [58] found that total suspended sediment levels were eight times higher in fire-exposed areas, an effect that was exaggerated by steep terrain at high elevations. High levels of erosion in fire-exposed, subalpine areas may suppress the moisture retention capacity of the soil and exacerbate other obstacles to recruitment.

On the north- and east-facing slopes in the WCW, environmental factors may have mitigated soil moisture stress. Reduced energy inputs and wind exposure diminish the moisture losses to snowpack sublimation, soil and vegetation evapotranspiration. For example, it is known that end of winter mean snowpack depth on north-facing slopes can be ~40% greater than south-facing slopes in the southern Canadian Rockies [59]. North-facing snow packs also persist longer into the growing season, thus ensuring melt water is available to support processes of germination and plant maintenance. Aspect specific patterns of snow accumulation can affect the functional length of the growing season, in turn altering ATE position [9]. Additionally, aspect can predict postfire recruitment of subalpine fir [17], with cool, north-facing aspects having rates of establishment twice as high as on south-facing aspects. It has been demonstrated that interactions among solar radiation (aspect), fire history, and elevation (in addition to anthropogenic activity) predicted broad patterns of succession in a large region of the Rocky Mountains of southern Alberta [47]; this included the encroachment on closed canopy forest on alpine meadow, or ATE advance. Together, these findings suggest postfire colonization is highly dependent on soil moisture, and that this limitation is exacerbated at high elevations.

Aspect effects that are unrelated to fire disturbance have been reported in a number of studies of ATE dynamics [9,60-62]. However, the proposed mechanisms of these aspect effects are often site-specific, and vary by study region. These include differences in lapse rates among aspects [61] and diurnal cycles of cloud cover causing preferential aspects in equatorial rainforests [60].

The suggestion that soil moisture limitation, caused by warm and dry conditions, prevented postfire recruitment is consistent with known climatic history of the region. The early 20th-century is noted as being a historically warm period in the eastern slopes of the Canadian Rockies, with dendroclimatic records showing the periods from 1917-1936 and 1939-1958 as having the highest temperature anomalies since the 1400s [63]. In fact, one of these reconstructions listed 1936, the year of the Pass Creek fire in WCW, as the second warmest year in the ~1000 year record [63]. This observation suggests the possibility that abnormally high atmospheric temperatures may have a two-fold effect on subalpine forests, first causing the desiccated conditions required for a burn at high elevation, then limiting expected regrowth.

Elevation

High degree change was almost entirely restricted to elevations greater than 1800 m, which coincide with the lower extent of ATE. This finding is consistent with the expectation that, as a transitional boundary that occurs as a function of elevation, canopy cover at ATE would be subject to a high level of variability [13,64,65]. Fire-exposed areas had lower mean elevations in all change classes, but the majority of canopy cover change was seen in these areas. Elevations greater than 1800 m accounted for a proportionally small region of fire-exposed areas, and yet the highest degree of change was seen at those elevations. The positive correlation between change degree and elevation is maintained in fire exposed areas suggesting that postfire mortality and re-establishment are potentially affected by factors that covary with elevation.

In the Canadian Rockies, there is evidence of both upslope advance and downslope recession of ATE elevation over the past 300 years. Our estimate of ATE advance in areas undisturbed by fire is consistent with other observations in the region; reported a subalpine fir advance rate of 0.28-0.62 m yr⁻¹ along an upslope transect [36].

Slope

Treelines in high slope areas are known to be subject to mechanical disturbances by avalanches [11,66], or enhanced soil erosion following disturbance [17,67], which introduces the prospect that treelines in

WCW may be orographically regulated [6]. Cyclical patterns of disturbance and regrowth may explain the high degree of both mortality and succession seen in high slope areas. However, in a nearby site in Glacier National Park (GNP), avalanche-regulated treelines were stable over a >70 year repeat photograph record, showing no cycle of disturbance and advance [40]. If the observed slope pattern in WCW is a result of cyclical processes, then this discrepancy between WCW and GNP warrants further investigation.

The slope of fire-exposed areas was generally lower than non-fire-exposed areas. High terrain slopes can accelerate the spread of fire [68], but this effect was outweighed by the extent of the WCW fire that occurred on the low slopes of the valley bottom. In two cases, *medium* and *high succession* classes, slopes were greater in fire-exposed areas than in non-fire-exposed areas, which agrees with research showing that slope was one of the factors that predicted re-establishment of subalpine fir following disturbance by fire [17]. In that study, 75% of establishment occurred on slopes with angles between 40-60°. The authors reasoned that this intermediate slope range was steep enough to avoid competition by shrubs during re-establishment, but shallow enough to avoid frequent disturbance by rockslide and avalanche. That explanation is inconsistent with the present finding, as areas with slopes above 30° saw a high proportion of both mortality and succession.

5. Conclusions

The WSL-MT monoplotted technique provided the ability to perform quantitative spatial analysis of mountain landcover change on oblique, historic repeat photography from the MLP. Canopy cover estimates from these images agree with canopy cover measurements using airborne lidar [49]. The method provides the ability to observe ATE at a spatial resolution and temporal extent which is sufficient to identify topographic correlation to change, which is an important step in understanding interactions between modulating factors. The MLP dataset is ideal for multi-scale investigation of change in the ATE; the spatial extent of MLP photographs ranges from individual meadows (~100 m), to valleys (~10 km), at sites which can be several hundred kilometers apart [47].

In ATE in the WCW, patterns of postfire colonization related to edaphic controls have hindered regrowth and locally suppressed ATE on south-facing aspects, while considerable postfire colonization was seen on north-facing aspects. This observation appears to directly support hypothesis of edaphic controls on colonization postfire. While fire is not a direct control on ATE position, it is in this case a modulating factor, and acts to locally suppress ATE below the thermal limits. While much has been learned about these edaphic controls using stand reconstructions, these can be limited in their understanding of microsite-processes and larger regional trends. Repeat photography provides the opportunity to observe the patterns at a high spatial resolution, over broad regions and diverse landscapes. Future research would benefit from investigating the interactions between fire, moisture limitation, and other modulating factors, using the multi-scale approach to understanding ATE dynamics [4,5].

In unburned areas on north-facing slopes, where moisture is not limiting, the ATE has systematically advanced. These observations of successful regeneration and ATE advance are consistent with increasing alpine isotherms and energy availability, as are expected under many climatic change scenarios, provided sufficient moisture or nutrient availability. More research is needed to further refine these hypotheses, particularly with respect to snowpack quantities and persistence at and above ATE.

Supplementary Materials: Data for the analysis presented in this research are maintained in the following repository: https://osf.io/ha579/?view_only=9cdda5a7fd47413386851fc91ba1282b. The WSL Monoplotted software is available free of charge (<https://www.wsl.ch/en/services-and-products/software-websites-and-apps/monoplotted-tool.html>). Historic and repeat photography from the mountain legacy project are open (<http://mountainlegacy.ca>). Spot 6 imagery

is commercially available (<https://www.intelligence-airbusds.com/en/8693-spot-67>). The Alberta Vegetation Inventory are Fire Disturbance layers are open, by request (<https://www.alberta.ca/forest-and-vegetation-inventories-data.aspx>).

Author Contributions: Conceptualization, D.M. and C.H.; methodology, D.M. and C.H.; software, D.M.; validation, D.M.; formal analysis, D.M.; data curation, D.M.; writing—original draft preparation, D.M.; writing—review and editing, C.H.; visualization, D.M.; supervision, C.H.; project administration, C.H.; funding acquisition, D.M. and C.H. All authors have read and agreed to the published version of the manuscript.

Funding: McCaffrey acknowledges funding from the NSERC AMYTHEST program and the University of Lethbridge Graduate Students' Association. Hopkinson acknowledges funding from the Alberta Innovates, Water Innovation Program, the NSERC Discovery Grant Program, and Campus Alberta Innovates Program

Acknowledgments: Repeat photography was graciously provided by the Mountain Legacy Project. Dr. Claudio Bozzini is thanked for developing the WSL monoplotting tool and instructing our team on its use. Dr. Marie-Pierre Rogeau was instrumental in accessing the fire history data of the West Castle area.

Conflicts of Interest: The authors are unaware of any potential conflicts of interest.

References

1. Körner, C. A re-assessment of high elevation treeline positions and their explanation. *Oecologia* **1998**, *115*, 445-459.
2. Körner, C.; Paulsen, J. A world-wide study of high altitude treeline temperatures. *Journal of Biogeography* **2004**, *31*, 713-732.
3. Holtmeier, F.-K. *Mountain timberlines: ecology, patchiness, and dynamics*; Springer Science & Business Media: 2009; Vol. 36.
4. Case, B.S.; Duncan, R.P. A novel framework for disentangling the scale-dependent influences of abiotic factors on alpine treeline position. *Ecography* **2014**, *37*, 838-851.
5. Weiss, D.J.; Malanson, G.P.; Walsh, S.J. Multiscale Relationships Between Alpine Treeline Elevation and Hypothesized Environmental Controls in the Western United States. *Annals of the Association of American Geographers* **2015**, *105*, 437-453.
6. Holtmeier, F.K.; Broll, G. Sensitivity and response of northern hemisphere altitudinal and polar treelines to environmental change at landscape and local scales. *Global Ecology and Biogeography* **2005**, *14*, 395-410.
7. Butler, D.R.; Malanson, G.P.; Walsh, S.J.; Fagre, D.B. Influences of geomorphology and geology on alpine treeline in the American West—More important than climatic influences? *Physical Geography* **2007**, *28*, 434-450.
8. Resler, L.M. Geomorphic controls of spatial pattern and process at alpine treeline. *The Professional Geographer* **2006**, *58*, 124-138.
9. Elliott, G.P.; Cowell, C.M. Slope Aspect Mediates Fine-Scale Tree Establishment Patterns at Upper Treeline during Wet and Dry Periods of the 20th Century. *Arctic, Antarctic, and Alpine Research* **2015**, *47*, 681-692.
10. Broll, G.; Holtmeier, F.-K. Wind as an Ecological Agent at Treelines in North America, the Alps, and the European Subarctic. *Physical Geography* **2010**, *31*, 203-233.
11. Walsh, S.J.; Butler, D.R.; Allen, T.R.; Malanson, G.P. Influence of Snow Patterns and Snow Avalanches on the Alpine Treeline Ecotone. *Journal of Vegetation Science* **1994**, *5*, 657-672.

12. Stine, M.B.; Butler, D.R. Effects of fire on geomorphic factors and seedling site conditions within the alpine treeline ecotone, Glacier National Park, MT. *Catena* **2015**, *132*, 37-44.
13. Harsch, M.A.; Hulme, P.E.; McGlone, M.S.; Duncan, R.P. Are treelines advancing? A global meta-analysis of treeline response to climate warming. *Ecology Letters* **2009**, *12*, 1040-1049.
14. Baker, W.L. *Fire ecology in Rocky Mountain landscapes*; Island Press: 2009.
15. Rogeau, M.-P.; Armstrong, G.W. Quantifying the effect of elevation and aspect on fire return intervals in the Canadian Rocky Mountains. *Forest Ecology and Management* **2017**, *384*, 248-261.
16. Colombaroli, D.; Henne, P.D.; Kaltenrieder, P.; Gobet, E.; Tinner, W. Species responses to fire, climate and human impact at tree line in the Alps as evidenced by palaeo-environmental records and a dynamic simulation model. *Journal of Ecology* **2010**, *98*, 1346-1357.
17. Stueve, K.M.; Cerney, D.L.; Rochefort, R.M.; Kurth, L.L. Post-fire tree establishment patterns at the alpine treeline ecotone: Mount Rainier National Park, Washington, USA. *Journal of Vegetation Science* **2009**, *20*, 107-120.
18. Cansler, C.A.; McKenzie, D.; Halpern, C.B. Area burned in alpine treeline ecotones reflects region-wide trends. *International Journal of Wildland Fire* **2016**, *25*, 1209.
19. Barrows, J. Fire behavior in northern Rocky Mountain forests. USDA Forest Service Station Paper 29, Northern Rocky Mountain Forest and Range Experiment Station, Missoula. *Montana* **1951**.
20. Fowler, P.M.; Asleson, D.O. The location of lightning-caused wildland fires, northern Idaho. *Physical Geography* **1984**, *5*, 240-252.
21. Howe, E.; Baker, W.L. Landscape heterogeneity and disturbance interactions in a subalpine watershed in northern Colorado, USA. *Annals of the Association of American Geographers* **2003**, *93*, 797-813.
22. Baker, W.L.; Kipfmueller, K.F. Spatial Ecology of Pre-Euro-American Fires in a Southern Rocky Mountain Subalpine Forest Landscape. *The Professional Geographer* **2001**, *53*, 248-262.
23. Buechling, A.; Baker, W.L. A fire history from tree rings in a high-elevation forest of Rocky Mountain National Park. *Canadian Journal of Forest Research* **2004**, *34*, 1259-1273.
24. Aplet, G.H.; Laven, R.D.; Smith, F.W. Patterns of Community Dynamics in Colorado Engelmann Spruce-Subalpine Fir Forests. *Ecology* **1988**, *69*, 312-319.
25. Peet, R.K. Forest vegetation of the Colorado front range. *Vegetatio* **1981**, *45*, 3-75.
26. Romme, W.H.; Knight, D.H. Fire frequency and subalpine forest succession along a topographic gradient in Wyoming. *Ecology* **1981**, *62*, 319-326.
27. Veblen, T.T. Age and size structure of subalpine forests in the Colorado Front Range. *Bulletin of the Torrey Botanical Club* **1986**, 225-240.
28. Donnegan, J.A.; Rebertus, A.J. Rates and mechanisms of subalpine forest succession along an environmental gradient. *Ecology* **1999**, *80*, 1370-1384.
29. Kelsey, K.C.; Redmond, M.D.; Barger, N.N.; Neff, J.C. Species, climate and landscape physiography drive variable growth trends in subalpine forests. *Ecosystems* **2018**, *21*, 125-140.

30. Redmond, M.D.; Kelsey, K.C. Topography and overstory mortality interact to control tree regeneration in spruce-fir forests of the southern Rocky Mountains. *Forest ecology and management* **2018**, *427*, 106-113.
31. Danby, R.K. Monitoring Forest-Tundra Ecotones at Multiple Scales. *Geography Compass* **2011**, *5*, 623-640.
32. Allen, T.R.; Walsh, S.J. Spatial and compositional pattern of alpine treeline, Glacier National Park, Montana. *Photogrammetric Engineering and Remote Sensing* **1996**, *62*, 1261-1268.
33. Bolton, D.K.; Coops, N.C.; Hermosilla, T.; Wulder, M.A.; White, J.C. Evidence of vegetation greening at alpine treeline ecotones: three decades of Landsat spectral trends informed by lidar-derived vertical structure. *Environmental Research Letters* **2018**, *13*, 084022.
34. Danby, R.K.; Hik, D.S. Evidence of Recent Treeline Dynamics in Southwest Yukon from Aerial Photographs. *Arctic* **2007**, *60*.
35. Tinner, W.; Theurillat, J.-P. Uppermost limit, extent, and fluctuations of the timberline and treeline ecocline in the Swiss Central Alps during the past 11,500 years. *Arctic, Antarctic, and Alpine Research* **2003**, *35*, 158-169.
36. Bekker, M.F. Positive Feedback between Tree Establishment and Patterns of Subalpine Forest. *Arctic, Antarctic, and Alpine Research* **2005**, *37*, 97-107.
37. Mamet, S.D.; Kershaw, G.P. Subarctic and alpine tree line dynamics during the last 400years in north-western and central Canada. *Journal of Biogeography* **2012**, *39*, 855-868.
38. Sakulich, J. Reconstruction and spatial analysis of alpine treeline in the Elk Mountains, Colorado, USA. *Physical Geography* **2016**, *36*, 471-488.
39. Butler, D.R.; DeChano, L.M. Environmental Change in Glacier National Park, Montana: An Assessment Through Repeat Photography from Fire Outlooks. *Physical Geography* **2001**, *22*, 291-304.
40. Butler, D.R.; Malanson, G.P.; Cairns, D.M. Stability of Alpine Treeline in Northern Montana, USA. *Phytocoenologia* **1994**, *22*, 485-500.
41. Klasner, F.L.; Fagre, D.B. A Half Century of Change in Alpine Treeline Patterns at Glacier National Park, Montana, U.S.A. *Arctic, Antarctic, and Alpine Research* **2002**, *34*, 49-56.
42. Kullman, L.; Öberg, L. Post-Little Ice Age tree line rise and climate warming in the Swedish Scandes: a landscape ecological perspective. *Journal of Ecology* **2009**, *97*, 415-429.
43. Moiseev, P.A.; Shiyatov, S.G. Vegetation Dynamics at the Tree-Line Ecotone in the Ural Highlands, Russia. In *Alpine Biodiversity in Europe*, Ecological Studies ed.; Springer Science and Business Media: Berlin, Germany, 2003; Vol. 167, pp. 423-435.
44. Roush, W.; Munroe, J.S.; Fagre, D.B. Development of a spatial analysis metho using ground-based repeat photography to detect changes in the alpine treeline ecotone, Glacier National Park, Montana, USA. *Arctic, Antarctic, and Alpine Research* **2007**, *39*, 297-308.
45. Bozzini, C.; Conedera, M.; Krebs, P. A New Monoplotting Tool to Extract Georeferenced Vector Data and Orthorectified Raster Data from Oblique Non-Metric Photographs. *International Journal of Heritage in the Digital Era* **2012**, *1*, 499-518.

46. Stockdale, C.A.; Bozzini, C.; Macdonald, S.E.; Higgs, E. Extracting ecological information from oblique angle terrestrial landscape photographs: Performance evaluation of the WSL Monoplotting Tool. *Applied Geography* **2015**, *63*, 315-325.
47. Stockdale, C.A.; Macdonald, S.E.; Higgs, E. Forest closure and encroachment at the grassland interface: a century-scale analysis using oblique repeat photography. *Ecosphere* **2019**, *10*, e02774.
48. Trant, A.J.; Starzomski, B.M.; Higgs, E. A publically available database for studying ecological change in mountain ecosystems. *Frontiers in Ecology and the Environment* **2015**, *13*, 187-187.
49. McCaffrey, D.R.; Hopkinson, C. Assessing Fractional Cover in the Alpine Treeline Ecotone Using the WSL Monoplotting Tool and Airborne Lidar. *Canadian Journal of Remote Sensing* **2017**, *43*, 504-512.
50. Wildfire Management Branch - Alberta Agriculture and Forestry. Fire History Polygons C5 FMU. 2017.
51. Alberta Environment and Parks. Alberta Vegetation Inventory. Branch, R.I.M., Ed. Edmonton, Alberta, Canada, 2012.
52. Rogeau, M.-P.; Flannigan, M.D.; Hawkes, B.C.; Parisien, M.-A.; Arthur, R. Spatial and temporal variations of fire regimes in the Canadian Rocky Mountains and Foothills of southern Alberta. *International Journal of Wildland Fire* **2016**, *25*, 1117-1130.
53. Rogeau, M. *Fire history study Castle River watershed*; Alberta. Technical report prepared for Alberta Environment and Sustainable ...: 2012.
54. Lumley, T.; Diehr, P.; Emerson, S.; Chen, L. The importance of the normality assumption in large public health data sets. *Annual review of public health* **2002**, *23*, 151-169.
55. Germino, M.J.; Smith, W.K.; Resor, A.C. Conifer seedling distribution and survival in an alpine-treeline ecotone. *Plant Ecology* **2002**, *162*, 157-168.
56. Kueppers, L.M.; Conlisk, E.; Castanha, C.; Moyes, A.B.; Germino, M.J.; de Valpine, P.; Torn, M.S.; Mitton, J.B. Warming and provenance limit tree recruitment across and beyond the elevation range of subalpine forest. *Glob Chang Biol* **2017**, *23*, 2383-2395.
57. Moyes, A.B.; Germino, M.J.; Kueppers, L.M. Moisture rivals temperature in limiting photosynthesis by trees establishing beyond their cold-edge range limit under ambient and warmed conditions. *New Phytologist* **2015**, *207*, 1005-1014.
58. Silins, U.; Stone, M.; Emelko, M.B.; Bladon, K.D. Sediment production following severe wildfire and post-fire salvage logging in the Rocky Mountain headwaters of the Oldman River Basin, Alberta. *Catena* **2009**, *79*, 189-197.
59. Hopkinson, C.; Collins, T.; Anderson, A.; Pomeroy, J.; Spooner, I. Spatial snow depth assessment using LiDAR transect samples and public GIS data layers in the Elbow River watershed, Alberta. *Canadian Water Resources Journal* **2012**, *37*, 69-87.
60. Bader, M.Y.; Ruijten, J.J.A. A topography-based model of forest cover at the alpine tree line in the tropical Andes. *Journal of Biogeography* **2008**, *35*, 711-723.
61. Dang, H.; Zhang, Y.; Zhang, Y.; Zhang, K.; Zhang, Q. Variability and rapid response of subalpine fir (*Abies fargesii*) to climate warming at upper altitudinal limits in north-central China. *Trees* **2015**, *29*, 785-795.

62. Greenwood, S.; Chen, J.C.; Chen, C.T.; Jump, A.S. Strong topographic sheltering effects lead to spatially complex treeline advance and increased forest density in a subtropical mountain region. *Global Change Biol* **2014**, *20*, 3756-3766.
63. Luckman, B.H.; Wilson, R.J.S. Summer temperatures in the Canadian Rockies during the last millennium: a revised record. *Climate Dynamics* **2005**, *24*, 131-144.
64. Körner, C. *Alpine treelines: functional ecology of the global high elevation tree limits*; Springer Science & Business Media: 2012.
65. Tranquillini, W. *Physiological ecology of the alpine timberline. Tree existence at high altitudes with special reference to the European Alps.*; Springer-Verlag: Berlin, Germany, 1979.
66. Butler, D.R.; Walsh, S.J. Site characteristics of debris flows and their relationship to alpine treeline. *Physical Geography* **1994**, *15*, 181-199.
67. Shakesby, R.; Doerr, S. Wildfire as a hydrological and geomorphological agent. *Earth-Science Reviews* **2006**, *74*, 269-307.
68. Werth, P.A.; Potter, B.E.; Clements, C.B.; Finney, M.; Goodrick, S.L.; Alexander, M.E.; Cruz, M.G.; Forthofer, J.A.; McAllister, S.S. Synthesis of knowledge of extreme fire behavior: volume I for fire managers. **2011**.

Appendix 3: Simulating Historic Change in the Alpine Treeline Ecotone using a Random Forest Model

McCaffrey and Hopkinson. HISTORIC CHANGE IN THE ALPINE TREELINE ECOTONE USING A RANDOM FOREST MODEL. *Canadian Journal of Remote Sensing* (Accepted pending revision)

Simulating Historic Change in the Alpine Treeline Ecotone using a Random Forest Model

Author 1: David R. McCaffrey, Department of Geography, Alberta Water & Environmental Science Building, University of Lethbridge, 4401 University Drive, Lethbridge, Alberta, T1K 3M4, phone: 403-329-2225, fax: 403-332-4039, david.mccaffrey@alumni.uleth.ca

Author 2: Chris Hopkinson, Department of Geography, Alberta Water & Environmental Science Building, University of Lethbridge, 4401 University Drive, Lethbridge, Alberta, T1K 3M4, phone: 403-329-2225, fax: 403-332-4039, c.hopkinson@uleth.ca

Abstract

Historic changes in Alpine Treeline Ecotone were modeled using 17 topographic, climatic, and disturbance variables in a random forest model. Airborne lidar and oblique historic repeat photography were used to identify changes in canopy cover in the West Castle Watershed (WCW), Alberta Canada (49.3° N, 114.4° W). A random forest model was trained on ~36% of the watershed which was observable in oblique imagery, then used for a spatial extension to predict change classes in the unobserved regions of the watershed. Overall accuracy compared to the training data was significant (0.84, $n = 89,440$, $p < 0.001$), and kappa showed moderate agreement at 0.58. Ranked variable importance demonstrated that elevation and solar radiation were strong predictors of ATE change.

Introduction

In the alpine treeline ecotone (ATE) tree growth is limited at high elevations by low air temperature (Körner & Paulsen 2004). Heterogenous tree cover in the ATE is generally attributed to patterns of climate, topography, and response to disturbance (Butler et al. 2007; Holtmeier & Broll 2005; Weiss et al. 2015), but it is understood that the absolute limit to tree growth at the upper boundary of ATE is thermal (Körner 1998). Increases in atmospheric temperature are expected over the 21st-century, and the effects of warming on mountain ecosystems may be amplified by elevation-dependent processes (Pepin et al. 2015). The potential for atmospheric warming to increase the elevation of treelines globally has been a matter of debate (Grace et al. 2002; Harsch et al. 2009). Some models predict a linear increase in the elevation of ATE, correlating to increases in atmospheric temperature (Schwörer et al. 2014). Other models suggest that ATE elevation will remain static (Paulsen & Körner 2014), or that advance may be limited by geomorphic barriers (Macias-Fauria & Johnson 2013). Ultimately, the ability to predict what effect, if any, a warming climate will have on ATE position will rely on the accuracy of models of ATE dynamics.

Most models of change in the ATE are generated by using recent climate history and local topography to explain the present location of ATE, then extrapolating future ATE limits under anticipated climate warming scenarios (Bader et al. 2008; Macias-Fauria & Johnson 2013; Paulsen & Körner 2014). Alternatively, ATE dynamics can be understood by observing historical change and associating advance or recession with the climatic and topographic conditions under which the change occurred (Dang et al. 2015; Elliott 2012; Greenwood et al. 2014; Leonelli et al. 2016; Liang et al. 2016; Lloyd & Graumlich 1997; Yadava et al. 2017)

This study falls into the latter category, adding to the body of research that attempts to

understand ATE dynamics through modeling historic change. It adapts a previously published model of ATE change in the Canadian Rockies (Macias-Fauria & Johnson 2013), and trains that model on historic changes in ATE canopy observed using airborne lidar and repeat photography McCaffrey (2018).

Literature Review

There are several methods of observing historic change in the ATE, each of which has limits on the spatial and temporal resolution (Danby 2011). The oldest records of ATE dynamics are generally from temperate regions, and show correlations between climate and ATE position throughout the Holocene (Heiri et al. 2006; Lloyd & Graumlich 1997; Schwörer et al. 2017; Schwörer et al. 2014). Schwörer et al. (2017) used paleobotanical records to show that in the Canadian Cordillera, the position of subalpine fir (*Abies lasiocarpa*) and Engelmann spruce (*Picea engelmannii*) was primarily correlated to summer temperature throughout the Holocene. In that research, study sites did not have concurrent ATE advance in response to increases in atmospheric temperature, suggesting local variations in topography or geomorphology may have modified ATE position.

Dendrochronological records have been used to estimate the relative effects of temperature and moisture availability on ATE position, with evidence from the Sierra Nevadas indicating treeline elevation decreasing during periods of extended drought, ~750-850 BP (Lloyd & Graumlich 1997). These inferences about ATE position in previous millennia often lack the spatial resolution to determine site-specific impacts of climate and topography. One exception is found in Schwörer et al. (2014), who developed a dynamic vegetation model of climatic and topographic conditions surrounding a lake in the Swiss Alps, at a resolution of 25 x 25 m, adjusting model parameters until conditions at 9800 BP were simulated; results suggested dry conditions with increased seasonality. But while paleobotanical and dendrochronological records provide excellent temporal range, other methods must be used to provide the spatial resolution needed to model the specific effects of topography on ATE position.

Records of ATE dynamics in the 20th-century are abundant and typically collected using remote sensing, dendrochronology, or long-term plots (Harsch et al. 2009). In a meta-analysis, Harsch et al. (2009) identified 103 publications which describe long term records of treeline position (both ATE and latitudinal treeline) at 166 sites around the world, between 1900-2008. Approximately 52% of these sites saw some form of treeline advance, and the only two studies that reported treeline recession were associated with disturbances. Linear regression with climate records revealed that many of these increases were associated with increases in atmospheric temperature. This finding is in line with multi-scale models that have emphasized the primacy of temperature in determining ATE position at a coarse scale (latitudinal or continental), while demonstrating local variation that is often topographically mediated (Case & Buckley 2015; Paulsen & Körner 2014; Weiss et al. 2015).

When topographic factors are considered, variables like slope, curvature, or compound topographic index have all been cited as factors regulating landscape-scale ATE position (Bader & Ruijten 2008; Case & Buckley 2015; Macias-Fauria & Johnson 2013). Topographic variables may influence mechanical disturbance (rockslide and avalanche), moisture accumulation, microsite formation, pedogenesis, erosion, and geomorphic conditions, all of which have the potential to affect the position of ATE.

Among topographic variables, aspect is strongly associated with heterogeneity in treeline landscapes (Bader & Ruijten 2008; Dang et al. 2015; Elliott & Cowell 2015; Greenwood et al. 2014). ATE advance can occur preferentially on one aspect in a landscape, with proposed mechanisms of this effect varying by location. Dang et al. (2015) found that in the Chinese Qinling Mountains, ATE advance occurred predominantly on north-facing slopes. This is surprising, considering that northern slopes of the region are cooler. Dang et al. (2015) note that a warming trend in lapse rates has been shown to occur at a faster rate on northern slopes, suggesting the aspect effect is a result of air mass movement. In a model of ATE in the Ecuadorian Andes, Bader and Ruijten (2008) found that treeline was higher on western slopes. They hypothesized this was a result of diurnal cycles of cloud cover in the rainforest; clear mornings led to excess insolation on eastern slopes, resulting in drying or photoinhibition. In dendrochronological records of treeline advance along a 1,000 km latitudinal transect of the American Rockies, Elliott and Cowell (2015) showed that pulses of treeline advance were more pronounced on northern aspects than on southern aspects during periods of drought. While this may seem paradoxical, it must be considered that drought in this context does not mean a reduction in summer precipitation, but rather a reduction in snow accumulation and a subsequent lengthening of the growing season. While not all research demonstrates an aspect effect on ATE dynamics (Leonelli et al. 2016), it is evident that ecologically relevant variables that are affected by aspect can cause asymmetric treeline change, and that the nature of this aspect effect is site specific.

Heterogeneity in the ATE may also be impacted as a result of fire disturbance. Fire return intervals in the Rockies have varied by aspect throughout the Holocene, with south-facing aspects seeing more frequent fires (Mustaphi et al. 2013). Rogeau and Armstrong (2017) show that south-facing aspects in the front ranges of the Canadian Rockies are more than twice as likely to burn as north-facing aspects. Aspect may not only impact fire frequency in the ATE, but also post-fire patterns of re-establishment and mortality in the ATE. In a study site on Mount Rainier, Washington, Stueve et al. (2009) used aerial imagery from the period 1970-2003 to analyze topographic effects on ATE advance following a fire that occurred in 1930. They found that tree establishment rates were twice as high on north-facing aspects than on south-facing aspects. These findings introduce the possibility that the underlying climatic and topographic variables which lead to low fire return interval on southern aspects may be correlated, and potentially causally related, to environmental variables which influence ATE position.

The aspect effect described in Stueve et al. (2009) supports similar observations made in the West Castle Watershed (WCW), in the Canadian Rockies (McCaffrey 2018), using an oblique repeat photography analysis to assess ATE change between 1914 and 2006. As in Stueve et al. (2009), the region experienced wildfires in 1930s. Distinct topographic patterns of ATE change were observed, and differences emerged when these patterns were factored by exposure to fire. High rates of mortality (i.e. downslope ATE recession) were seen on south-facing slopes that had been exposed to fire between observation periods. North-facing aspects experienced high rates of succession (i.e. upslope ATE advance). These observations led to questions about the interaction between fire exposure, aspect, and ATE position. The research highlighted the need to better understand the effect of topographic and climatic variables on post-fire tree establishment in the ATE (McCaffrey, 2018; McCaffrey and Hopkinson, submitted).

Objectives

The primary objective of this study is to develop a model of historic canopy cover change in the WCW that builds on and gap fills remote sensing-based observations. To do so, this study adapts a modeling technique from Macias-Fauria and Johnson (2013) which uses proxy values for climatic, topographic, and disturbance variables to predict changes in canopy cover. The accuracy of these predictions is tested using the method of Macias-Fauria and Johnson (2013).

The secondary object of this study is to test the findings Macias-Fauria and Johnson (2013) with respect to the ranked importance of the climatic, topographic, and disturbance variables in their ability to explain ATE dynamics in the Canadian Rockies. Using a model trained on contemporary ATE position, Macias-Fauria and Johnson (2013) suggested that ATE advance would primarily be obstructed by geologic barriers under future warming scenarios. It is important to test these assumptions using historical observations of ATE dynamics, such as those from WCW.

The final objective of this study is to explore the potential for ranked variable importance to disentangle the climatic and topographic factors leading to the apparent interaction between fire-exposure and aspect, described in Stueve et al. (2009) and McCaffrey (2018). To that end, the modeling procedure is repeated in separate analysis which are restricted to areas below ATE, and areas at and above ATE. Ranked variable importance is compared between these areas, to determine if the processes driving change in the ATE are distinct from those driving change at lower elevations.

Methods

Study Area and Data

This research was conducted using a land cover change dataset from the WCW, described in McCaffrey (2018). Located in the Canadian Rockies (Figure 1), the WCW is ~103 km² in area, with elevation ranging from 1400-2600 m. ATE occurs between ~1800-2300 m a.s.l. and is dominated by subalpine fir and Engelmann spruce.

The land cover change dataset was generated using fractional cover estimates from repeat oblique photography (McCaffrey & Hopkinson 2017). The repeat photographs, provided by the Mountain Legacy Project (Trant et al. 2015), contain pairs of photographs from a 1914 survey of the Canadian Rockies and repeat photographs from 2006. With the aid of a high-resolution DEM and monoploting software, spatial information was extracted from the oblique photographs and projected to an orthogonal view. Differences between the 1914 and 2006 land cover were cataloged (for details, see (McCaffrey 2018; McCaffrey & Hopkinson 2017; submitted). The result was a dataset which classified change into one of seven categories: high, medium, and low mortality, high, medium, and low succession, and no change (Figure 2). A total of 89,440 grid cells were classified, 20 x 20 m each, representing ~36% of the WCW area.

Random Forest

Random forest is a classification and modeling algorithm that uses bootstrapped decision trees based on random samples of multivariate data (Breiman 2001). The technique provides strong predictive capability, while also demonstrating the ability to rank the relative importance of classification variables in ecologically relevant ways (Cutler et al. 2007). Random forest analysis was used to investigate variability of *Nothofagus* treeline in New Zealand (Case &

Buckley 2015). The analysis used 17 topographic and climatic variables, showing that in addition to the expected effects of precipitation and moisture gradients, regional differences in topo-climate were able to explain up to half of the variance in treeline position.

Random forest has also been used to predict the response of ATE to climate change in the Canadian Rockies (Macias-Fauria & Johnson 2013). Similar to Case and Buckley (2015), this work used variable importance ranking to understand the relative influence of climate, topography, and geology/geomorphology on contemporary ATE position. When historic climate values were substituted with warming projections, geologic and geomorphic barriers to upslope tree establishment became apparent, with ATE advance largely restricted to gentle, colluvium-mantled slopes.

The random forest model presented by Macias-Fauria and Johnson (2013) predicts ATE change in the Kananaskis Valley, a site that shares many characteristics with the WCW, including similar elevation range, climate, area, species mix, and topography. As such, the present work attempted to closely replicate the climate and topography proxy variables used by Macias-Fauria and Johnson (2013), and also adopted their modeling procedure. If upslope ATE advance is limited primarily by geology, any advances observed over the 20th-century in WCW may have already hit geologic barriers. In this case, geology would have high variable importance rank in the present study, as it did in Macias-Fauria and Johnson (2013). Similarly, if historic upslope advance were restricted to colluvium, as Macias-Fauria and Johnson (2013) predict future upslope advance will be, then the importance of surficial geology will rank high in the model.

To date, no known studies have used random forest for classification of historical change in the ATE.

Topographic Variables

Five topographic variables were used to investigate potential correlations in ATE change: elevation, slope, aspect, curvature, and compound topographic index (Table 1). All topographic variables were calculated using a 1 m digital elevation model (DEM), generated from airborne lidar flown on October 18, 2014. Topographic conditions in 2014 are an accurate representation of conditions in 2006, as there were no notable geomorphic disturbances (landslides, etc.) in the intervening period.

Elevation was gridded at 1 m resolution, then aggregated to match the 20 m land cover change dataset, using mean values. Slope values were calculated on the 1 m DEM using the *slope* function in ArcGIS 10.3 (Esri 2014), then aggregated to 20 m resolution, again using mean values. The 1 m DEM is expected to produce greater estimations of slope, compared to slopes calculated on the 20 m, aggregated DEM (Kienzle 2004). Aspect was calculated on the 20 m grid, as averaging azimuth values in an aggregation process may have led to errors. Aspect was classified into nine categorical values, based on 45° bins, eight for the intercardinal directions (N = 337.5-22.5°, NE = 22.5-67.5°, E = 67.5-112.5°, SE = 112.5-157.5°, S = 157.5-202.5°, SW = 202.5-247.5°, W = 247.5-292.5°, NW = 292.5-337.5°) and one for flat surfaces without slope.

Curvature refers to the convexity or concavity of a DEM grid cell, in relation to neighboring cells. It is calculated by fitting a fourth order polynomial to the 3 x 3 window around a DEM grid cell; positive values indicate convexity and negative values represent concavity (Zevenbergen & Thorne 1987). Curvature was calculated on the 20 m DEM, using the *curvature*

function in ArcGIS 10.3. This function calculates standard curvature, which combines both profile and plan curvature.

Compound topographic index (CTI), is an index of the ratio of upstream area to slope (Beven & Kirkby 1979). In ATE research, CTI has been used as a proxy for both cold air pooling (Case & Duncan 2014) and soil accumulation (Schwörer et al. 2014). CTI is measured using the formula:

$$CTI = \ln (A / \tan \beta) \quad (1)$$

Where A is the contributing upslope area, and β is the slope in degrees. These variables were calculated in ArcGIS 10.3, using the *flow direction* and *flow accumulation* functions for A , the *slope* function for β .

Geologic Variables

It has been demonstrated that soil properties, surficial geology, bedrock type, and mountain architecture all have a significant effect on the position of ATE (Fagre et al. 2007; Macias-Fauria & Johnson 2013). To account for potential geologic barriers to change in the ATE, two geologic variables were included in this model: surficial geology and bedrock type. For surficial geology, areas of bedrock, colluvial deposits, fluvial deposits, glacio-fluvial deposits, and moraine were categorized using data from the Alberta Geological Survey (Fenton et al. 2013). For bedrock types, nine geologic formations in the WCW were identified using the Lexicon of Canadian Stratigraphy (Glass 1990) and digitized from a historic geologic map of the region (Geological Survey of Canada & Price 1961).

Climatic Variables

Seven climatic variables were used to model change: precipitation (annual and JJA), temperature (annual and JJA), solar radiation (annual and JJA), and potential evapotranspiration. Given the geographic similarity in the research areas, this study followed the methods of Macias-Fauria and Johnson (2013) to develop precipitation and temperature model inputs. In alpine areas, steep gradients in temperature and precipitation are caused by lapse rates, which generally result in cooler temperatures and increased precipitation with increased elevation. To create proxy values that spatially distribute the anticipated effect of lapse rates, Macias-Fauria and Johnson (2013) adapted a method from Randin et al. (2009): 1) historic normal values (30-year) for temperature and precipitation were collected from three meteorological stations, within 32 km of their study site in the Canadian Rockies; 2) using locally observed temperature and precipitation enhancement rates (Cullen & Marshall 2011), monthly averages were normalized to sea-level (i.e. the residual if elevation is equal to 0), 3) residual temperature and precipitation values were interpolated, at 'sea-level', using inverse distance weighting (IDW), 4) using a DEM of the study area and the definition of the linear regression of lapse rate, temperature and precipitation values were calculated for each grid cell according to elevation, then adjusted by either adding or subtracting the value from the interpolated residual surface.

One issue with the Randin et al. (2009) method used by Macias-Fauria and Johnson (2013) is that the lapse rates described by Cullen and Marshall (2011) were from observations made between 2005-2009. Macias-Fauria and Johnson (2013) applied these to climate records for a period between 1971-2000. Given that the present study seeks to identify climatic variables associated with historic change in the ATE, a longer observation period was required (1914-

2006). Five Environment Canada meteorological stations were identified within 20 km of the WCW (West Castle, Spionkop Creek, Gardiner Creek, Castle, and Carbondale Lo). While these stations provided adequate elevation range to calculate lapse rates, the temporal coverage did not extend earlier than 1958. The search radius was extended to 50 km, and stations with the longest temporal coverage were selectively added. These historic records came from three nearby towns (Coleman, Beavers Mines, and Pincher Creek Town). The addition of the towns provided adequate temporal coverage, but they were located at low elevations compared to the study site. Using monthly mean temperature and monthly precipitation totals, annual and summer (JJA) averages for temperature and precipitation were calculated.

If a record did not contain all 12 months in a calendar year, that year was omitted from annual average calculation. Similarly, if a record did not contain all three summer months (June, July, and August), the year was omitted from summer average calculation. Sporadic temporal coverage caused the Carbondale Lo site to be omitted entirely, leaving seven meteorological stations, summarized in Figure 2, Figure 3, Table 2.

Following the Randin et al. (2009) method, data from the seven meteorological stations were used to regress a line between elevation and either temperature or precipitation. In all cases, elevation was significantly correlated to average annual and summer values for temperature and precipitation (Table 3). The lapse rates for both annual and JJA temperature were $-4.0^{\circ}\text{C} / 1 \text{ km}$, and were comparable to regional values reported in Cullen and Marshall (2011), (-4.2°C and -5.0°C respectively).

Solar radiation was calculated in ArcGIS 10.3 using the *area solar radiation* tool on the 20 m DEM. Solar radiation was modeled at 30-minute intervals, with all monthly totals summed for an estimate of annual insolation, and JJA totals summed for summer insolation.

Potential evapotranspiration (PET) values were modeled using the method of Oudin et al. (2005), which gives comparable results to Penman-Monteith calculations in rainfall-runoff models, but can be implemented when values for wind speed and humidity are unreliable. This method of PET calculation was based on the models of Jensen and Haise (1963) and McGuinness and Bordne (1972), which use the general form:

$$PE = \frac{R_e T_a + K_2}{\lambda \rho K_1} \quad \text{if } T_a + K_2 > 0 \quad (2)$$

$$PE = 0 \quad \text{otherwise}$$

where PE is the rate of potential evapotranspiration (mm day^{-1}), R_e is extraterrestrial solar radiation ($\text{MJ m}^{-2} \text{ day}^{-1}$), λ is latent heat flux (2.45 MJ kg^{-1}), ρ is the density of water (1000 kg m^{-3}), T_a is the mean daily air temperature ($^{\circ}\text{C}$). The fixed parameters K_1 ($^{\circ}\text{C}$) and K_2 ($^{\circ}\text{C}$) scale the importance of T_a , and set a threshold at which $PE = 0$. Oudin et al. (2005) empirically determined 100 and 5 to be the optimal values for K_1 and K_2 , which are the values used in this study. Average summer potential evapotranspiration values were modeled by calculating average daily JJA solar radiation (total JJA solar radiation in Wh/m^2 , divided by number of days in summer, converted to $\text{MJ m}^{-2} \text{ day}^{-1}$) for R_e , and using the average JJA temperatures (calculated from the previously described lapse rate analysis) for T_a .

Disturbance Variables

Two disturbance types were accounted for in the model: wildfire and anthropogenic

disturbance.

Two wildfires affected the WCW in the period between the 1914 and 2006 oblique photographs which were used to generate the change dataset. The Pass Creek Fire of 1936 mainly affected the northern extent of the WCW, covering 45.2% of the area of the watershed. A smaller fire occurred in 1934 and was restricted to the southern-most slope, covering only 1.5% of the watershed. Fire extents were delineated using historic descriptions, and aerial imagery from 1949 where the remnant fire scars were still visible (Wildfire Management Branch - Alberta Agriculture and Forestry 2017).

Finally, areas of WCW featuring anthropogenic disturbance between observations were identified and omitted. The Alberta Vegetation Inventory (Alberta Environment and Parks 2012) was used to aggregate disturbances including former oil well pads, historical forestry cut-blocks, a small ski resort (Castle Mountain), and roads. These features were excluded from model training, and were not included in the extent of the model prediction.

Model Execution

The topographic, geomorphic, climatic, and disturbance variables listed in Table 4 were used as model inputs for classification of land cover change in R v3.2.3 (R Core Team 2017) using the *randomForest* package (Liaw & Wiener 2002). A correlation matrix for the continuous variables was generated using the *corrplot* package (Wei & Simko 2016), and groups of variables that were highly correlated ($|r| > 0.60$) were reduced to a single, representative variable.

Canopy cover change in unobserved areas was predicted using the method of Macias-Fauria and Johnson (2013). The model was run 11 times, each time with a random sample of 50% of the training data used as input variables. After each model run, the results using that specific sample of training data were used to generate a prediction of land cover change for the whole extent of WCW, including the area of the watershed that was not observed in the initial oblique photography (note: anthropogenic disturbance areas were omitted from both training and prediction datasets, as they are a result of processes independent of historical ATE change). An odd number of model runs was used, so that once 11 predictive surfaces had been generated, the median value for each grid cell could be joined to make a final prediction raster (Macias-Fauria & Johnson 2013). Model accuracy was assessed by clipping the extent of the training data from the prediction raster, and comparing against the observed land cover change in a confusion matrix, using the *caret* package (Williams et al. 2017). Additionally, the out-of-bag assessment of error (overall accuracy when the 50% random sample is trained to predict the remaining 50% of the data) was recorded for each of the 11 runs.

Variable importance was compared by cataloging the mean decrease in Gini coefficient (MDG) for each variable, in each model run, and taking the mean MDG value among the 11 runs as the final MDG for each variable. Ranked variable importance is compared to the results of Macias-Fauria and Johnson (2013).

Finally, in an attempt to disentangle the climatic and topographic factors leading to the apparent interaction between fire-exposure and aspect in the ATE, a second analysis compared variable importance between the ATE and lower elevations. The training dataset was split in two; areas where elevation > 1800 m, which include ATE and alpine areas, are referred to as *high elevation*, and areas < 1800 m are referred to as *low elevation*. The value of 1800 m was selected as a break point, as it fell approximately between the median elevation of the training data (1742 m) and the median elevation for the extent of WCW (1838 m). The previous analyses

were repeated with 11 model runs, confusion matrix accuracy assessment, and ranking of variable importance using MDG.

Results

When the random forest model input variables were arranged in a correlation matrix, three groups of highly correlated continuous variables were identified (Figure 7, Table 4). In the first group, annual precipitation, annual temperature, summer precipitation, summer temperature, and elevation were found to be highly correlated ($|r| = 0.85 - 1.00$). These correlations reflected the results of Macias-Fauria and Johnson (2013), which also failed to prevent auto-correlation in the proxy variables developed using the method of Randin et al. (2009). In the second group, annual solar radiation, summer potential evapotranspiration, and summer solar radiation were found to be highly correlated ($|r| = 0.85 - 0.96$). In the third group, summer potential evapotranspiration and slope were found to be highly correlated ($|r| = 0.65$). Following the method of Macias-Fauria and Johnson (2013), one variable was selected from each of the highly correlated groups (elevation, summer solar radiation, and slope), with the other variables omitted.

Using the remaining nine variables (Figure 6), the median prediction raster was compared to the training data. Overall accuracy was significant (0.840, $n = 89,440$, $p < 0.001$), and kappa showed moderate agreement at 0.577. The training data had high representation of the *no change* cover class, which is evident from the high sensitivity value of *no change* compared to other variables. Conversely, specificity was much lower for *no change* than for the other change variables. To understand the effect of the large *no change* sample on measures of accuracy, sensitivity, and specificity, a second confusion matrix was constructed with *no change* values omitted. Sensitivity increased, and specificity remained high. Out-of-bag estimates of error showed 75.8% overall accuracy when training on a 50% random sample of the data, 11 times.

Prediction of canopy change class in unobserved areas was generated using the full training sample area (Figure 8). While this spatial extension was not validated, the out-of-bag estimate of error from the model training (i.e. 75.8%) provides the best assessment of model predictive accuracy on naïve data.

Variable ranking showed that elevation had the highest MDG, followed by the other continuous variables (summer solar radiation, slope, CTI, and curvature) (Figure 9, Table 6). The lowest MDG was seen in the categorical variables (bedrock geology, aspect, surficial geology, and fire exposure).

The second analysis, based on elevation partition, showed that *high elevation* areas resulted in model accuracy of 0.761, $n = 38,014$, $p < 0.001$, and kappa = 0.490. *Low elevation* areas had model accuracy of 0.826, $n = 51,426$, $p < 0.001$, and kappa = 0.300. Order of variable importance was generally consistent across both conditions, with summer solar radiation and slope having high MDG (Figure 9, Table 7). The categorical variables remained below the continuous variables in order importance. The largest difference in MDG between the *high elevation* and *low elevation* training samples was observed in the fire exposure variable.

Discussion

Accuracy

The random forest model demonstrated the ability to predict historic change in the WCW,

based on proxies of climatic, topographic, geologic, and disturbance variables. However, the large sample of areas without land cover change over the 92-year observation period introduced uncertainty in accuracy assessment. The high level of sensitivity for the *no change* class (0.998) compared to the other change classes (0.394-0.568) indicated that the accuracy predictions were potentially a function of the large sample improving the probability of a correct classification. Sensitivity of change class identification increases when the potential for confusion with *no change* class is removed (0.778-0.943), suggesting that future attempts to model ATE change with this method might benefit from using a training dataset with a large sample of ATE variability.

Accuracy of the *high elevation* model was lower than accuracy of the *low elevation* model, which was potentially a result in the discrepancy in the size of the training dataset (high elevation $n = 38,014$, low elevation $n = 51,426$). Canopy cover change in *low elevation* data was generally more homogenous, and experienced less change over the 92-year observation period. Thus, with less variability in change classes, the model provided more accurate classification at lower elevations. The low kappa value of the *low elevation* area (0.30) compared to the *high elevation* area (0.49) supports this reasoning.

The reported accuracies are only valid for the area of overlap between the training data and the prediction area, and do not reflect the accuracy of the spatial extension of the model outside of the training area. The best estimate of spatial extension accuracy come from the out-of-bag estimate of error, which showed overall accuracy to be 75.8%.

Elevation, Air Temperature, and Precipitation

Given that ATE is recognized as being a function of lower temperatures at high elevation, it is unsurprising that elevation ranked as the most important variable by MDG. This finding supports the work in McCaffrey (2018), which showed that the degree of both mortality and succession in the ATE was positively correlated to elevation; stated differently, changes in mortality and succession were more exaggerated at higher elevations. The finding is also consistent with a century-scale observation based model demonstrating a high degree of meadow to forest succession at high elevation in the region (Stockdale et al. 2019).

One of the objectives of this research was to disentangle the relative importance of temperature and precipitation on tree establishment following fire disturbance, in an effort to explain observed patterns of mortality on north- and south-facing aspects. The method that was used to generate proxy values for the effect of lapse rates on temperature and precipitation (Randin et al. 2009) did not adequately distinguish between the effect of elevation on temperature and precipitation. The method used IDW to interpolate the residual values at each meteorological station across the study area, but did not provide sufficient resolution to outweigh the impact of elevation in each calculation. The impact of the interpolated residual surfaces was small enough that summer temperature and precipitation values were correlated to elevation at $r = 1.0$, as they were both primarily functions of elevation.

The results call into question whether the method of Macias-Fauria and Johnson (2013) accurately distinguished between climatic and geologic factors in determining upslope treeline advance. As in the present study, Macias-Fauria and Johnson (2013) were not able to separate elevation, air temperature, and precipitation using the method of Randin et al. (2009). This is not insignificant, as the general finding of their research was that upslope treeline advance will be

limited by a lack of adequate geologic and geomorphic conditions at high elevations.

The inability to distinguish between the effects of lapse rate on temperature and precipitation underscores the need for improved understanding of local, meso-scale variations climate, and their relationship to elevation. Until modeled variables have the resolution required for this analysis, an understanding of the relative importance of air temperature versus precipitation on ATE dynamics can be gained through experimentation. Kueppers et al. (2017) used experimental heating and watering across an elevation transect to study the effect of changing air temperature and precipitation regimes on tree establishment in the ATE. Heating limited the establishment of Engelmann spruce at elevations above treeline, but this effect was mitigated with watering. This finding suggests that precipitation or moisture provided through accumulated snow could be essential for upslope treeline advance to results from increases in atmospheric temperature.

Solar Radiation, PET, and Aspect

Solar radiation also had high variable importance, as ranked by MDG. Previous models have demonstrated that solar radiation can explain variability of ATE at multiple spatial scales (Case & Duncan 2014), and that this effect was independent of the influence of insolation on soil moisture availability. Unfortunately, insolation and moisture availability could not be directly compared in the present model, as the Oudin et al. (2005) equation to calculate PET used solar radiation as a parameter, causing a high degree of correlation. Given the a priori expectation that these variables would be highly correlated, it is surprising they were used in Macias-Fauria and Johnson (2013).

In previous research (McCaffrey 2018), a strong aspect effect was observed, where the majority of mortality seen over the observation period occurred on south-facing, fire-exposed slopes. One of the expectations of this model would be that aspect would have a high MDG ranking, but paradoxically, aspect ranked low. It is worth considering that this result may be a function of the way in which random forest handles continuous versus categorical variables. Strobl et al. (2006) showed that the classification and regression tree algorithm (Breiman et al. 1984) used in random forest is biased such that continuous variables rank higher in importance than categorical variables. Additionally, categorical variables with a greater number of classes are biased towards higher variable importance rank. Indeed, when looking at variable importance rank in the present study, they are perfectly sorted, first by continuous variable, then in order of the number of classes in the categorical variables. Macias-Fauria and Johnson (2013) treated aspect as a categorical variable, also finding that it ranked low in variable importance; aspect was treated as a categorical variable in the present study, in order to maintain comparison between the studies.

Slope, CTI, and Curvature

The topographic variables of slope, CTI, and curvature were midway in rankings of importance. Slope has been shown to affect ATE through mechanical disturbance, with many treelines in Montana regulated by rockslide and avalanche (Butler & Walsh 1994; Walsh & Butler 1997; Walsh et al. 1994). WCW has numerous chutes where ATE appears to be suppressed, potentially by avalanche, but the episodic nature of these events may not have been captured in the two observations, separated by 92 years. Previous research indicated that the high

mortality class had significantly higher slope than other change classes, suggesting that mortality may have been the result of avalanche. It is conceivable that in specific circumstances, the effect of slope might outweigh that of elevation and aspect, but if this were broadly the case then slope would have had higher predictive value in the model.

CTI has been used in ATE research as a proxy for cool air-pooling (CAP) (Case & Duncan 2014), where radiative transfer and advection reduce air temperatures at ground surface. CAP is known to occur in the Rocky Mountains (Holden et al. 2011), which implies several possibilities: 1) CAP is occurring, and affecting ATE position, but CTI is not acting as a proxy for CAP, 2) CAP is occurring, and is correlated to CTI, but the effect is not dominant over other temperature mediators like elevation and insolation, 3) CAP is not occurring in the WCW. The results suggest that options 2 and 3 are the most plausible, but future ATE research in WCW should incorporate either high resolution, topographically distributed temperature measurements or modeling methods other than CTI to account for the potential effect of air-pooling.

The response of the curvature variable had middling rank, indicating that local convexity and concavity had a marginal effect on ATE dynamics. One issue with this finding is potentially the resolution. The curvature operates on a 3 x 3 moving window, which means that curvature was based on a 60 x 60 m observation space. This resolution is ideal for distinguishing gullies from ridgelines, but may be inadequate to identify smaller features associated with microsite formation and pedogenesis. Curvature was identified as the primary variable driving divergence from the climatic treeline maximum in the random forest model of Case and Buckley (2015).

Surficial and Bedrock Geology

Surficial and bedrock geology both rank low in variable importance, but this finding must be considered with the understanding that they are both categorical variables, and are subject to bias in the random forest model. Bedrock geology ranked higher in importance than surficial geology, but this is likely a function of bedrock geology containing more categories. These variables were included in the model in order to account for the finding of Macias-Fauria and Johnson (2013), that upslope ATE advance under climate warming projections would be limited by geologic constraints such as bedrock. If these geologic constraints did limit ATE advance in the WCW, then these variables would have had higher ranks. This finding does not preclude future geologic limitations to ATE advance in the WCW.

Fire

The ranked importance of fire exposure in the present research diverged the most from expectation. In the preceding research, an asymmetric response to fire exposure was observed, with fire causing higher mortality on south-facing slopes than on north-facing slopes. Given that finding, the fire-exposure variable was expected to rank high, however the results showed that fire-exposure was the least important variable. It must be taken into account that as a categorical variable with only two classes, fire-exposure may have been subject to the same model bias described for the other categorical variables.

An intriguing finding came from the elevation partition analysis. The largest difference in MDG between *high elevation* and *low elevation* datasets was seen in the fire-exposure variable. Fire-exposure provided a higher explanatory power for land cover change in *high elevation* areas than *low elevation areas*. One interpretation of this finding is that fire may interact with other

topo-climatic variables at ATE elevations in a way that it does not at lower elevations. For example, a stand-clearing fire in the 1930s which occurred at low elevation would have ample resources (soil moisture, insolation, adequate growing season temperature) to replace the stand before the 2006 observation, thus recording *no change*. At ATE elevation, where these resources are limited, soil moisture and growing season temperatures which may have been adequate for a mature tree stand pre-fire may have posed limits to tree establishment post-fire. This effect would result in the observed increase in the ability for fire to predict land cover change at *high elevations* versus *low elevations*.

Evidence of this threshold effect of fire can be seen in a recent studies of post-fire tree establishment in the ATE (Stine & Butler 2015). This study reported significant soil erosion by wind and water following the removal of vegetation and duff layers by fire. While effective soil depth did not significantly vary between burned and unburned sites, seedling establishment was significantly higher in burned areas with deeper soil and more rock shelter. This finding was curious, as previous studies had shown these variables did not affect seedling establishment in unburned areas (Malanson et al. 2002). Thus, variables related to topography and climate that do not appear to affect tree establishment in normal ATE conditions can become critically important to tree establishment following disturbance by fire. Future research should investigate patterns of post-fire tree establishment using a method that controls for the effect of aspect on climatic variables.

Conclusion

A random forest model of WCW was trained to classify historic changes in forest cover, based on a set of climatic, topographic, geologic, and disturbance variables. Elevation had the highest explanatory power in determining forest cover change. Air temperature and precipitation may also be considered to have high explanatory power in the model but there is high correlation among variables. Insolation also ranked highly. While several multi-scale models suggest that at a landscape scale, topography can outweigh the importance of temperature in ATE position (Case & Duncan 2014; Weiss et al. 2015), the topographic factors of slope, CTI, and curvature rank lower in importance than climatic variables. Lowest variable importance was seen in the categorical variables surficial geology, bedrock geology, and fire-exposure; these findings were unexpected, as previous studies had suggested strong effects from geologic variables (Macias-Fauria & Johnson 2013) and fire-exposure. Interpretation of variable importance ranking was made difficult by an apparent bias towards continuous variables. Future implementations of this method will better understand variable importance by applying random forests which use conditional interference trees for variable selection, avoiding bias (Strobl et al. 2006).

Partitioning the data into areas at and above ATE and below had little effect on the rank of variable importance. However, a discrepancy in the MDG of fire-exposure between elevation ranges provided an indication of the potential for ‘threshold effects’, where climatic and topographic stressors which are not limiting to tree establishment in the pre-fire ATE become limiting post-fire. These threshold effects may explain the correlation between fire-exposure, aspect, and mortality/succession at ATE described in previous research (McCaffrey 2018). Future research should focus on an understanding of how climatic variables which correlate to aspect (i.e. insolation, wind exposure, PET, soil moisture, soil temperature) alter patterns of post-fire tree establishment. While the modeling techniques used were generally insufficient to

confidently distinguish between factors at the resolution required to observe these micro-site formation effects, *in situ* measurements of hydroclimate at a range of elevations and aspects in WCW will provide this ability. Additionally, recent wildfires in southern Alberta, such as the Lost Creek Fire of 2003 and the Kenow Fire of 2017, could be useful for observing ‘threshold effects’ of post-fire tree establishment in the ATE in a variety of aspects and growth stages.

Given the increased frequency of ATE fire (Cansler et al. 2016), there is a need to add to the growing body of research investigating fire at high elevation (Rogeanu & Armstrong 2017; Stine & Butler 2015). The potentially limiting effect of fire on upslope treeline advance should remain an area of research focus.

Acknowledgments

McCaffrey acknowledges funding from the NSERC AMYTHEST program and the University of Lethbridge Graduate Students’ Association. Hopkinson acknowledges funding from the Alberta Innovates, Water Innovation Program, the NSERC Discovery Grant Program, and Campus Alberta Innovates Program. The authors are unaware of any conflict of interest in this research. Repeat photography was graciously provided by the Mountain Legacy Project. Dr. Claudio Bozzini is thanked for developing the WSL monoploting tool and instructing our team on its use. Dr. Marie-Pierre Rogeanu was instrumental in accessing the fire history data of the West Castle area.

Declaration of Interest Statement

The authors know of no conflicts of interest in the present research.

References

- Alberta Environment and Parks. (2012). *Alberta Vegetation Inventory*. Edmonton, Alberta, Canada.
- Bader, M. Y., Rietkerk, M., & Bregt, A. K. (2008). A simple spatial model exploring positive feedbacks at tropical alpine treelines. *Arctic, Antarctic, and Alpine Research*, 40(2), 269-278.
- Bader, M. Y., & Ruijten, J. J. A. (2008). A topography-based model of forest cover at the alpine tree line in the tropical Andes. *Journal of Biogeography*, 35(4), 711-723.
- Beven, K. J., & Kirkby, M. J. (1979). A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrological Sciences Bulletin*, 24(1), 43-69.
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.
- Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*: CRC press.
- Butler, D. R., Malanson, G. P., Walsh, S. J., & Fagre, D. B. (2007). Influences of geomorphology and geology on alpine treeline in the American West—More important than climatic influences? *Physical Geography*, 28(5), 434-450.
- Butler, D. R., & Walsh, S. J. (1994). Site characteristics of debris flows and their relationship to alpine treeline. *Physical Geography*, 15(2), 181-199.
- Cansler, C. A., McKenzie, D., & Halpern, C. B. (2016). Area burned in alpine treeline ecotones reflects region-wide trends. *International Journal of Wildland Fire*, 25(12), 1209.

- Case, B. S., & Buckley, H. L. (2015). Local-scale topoclimate effects on treeline elevations: a country-wide investigation of New Zealand's southern beech treelines. *PeerJ*, 3, e1334.
- Case, B. S., & Duncan, R. P. (2014). A novel framework for disentangling the scale-dependent influences of abiotic factors on alpine treeline position. *Ecography*, 37(9), 838-851.
- Cullen, R. M., & Marshall, S. J. (2011). Mesoscale temperature patterns in the rocky mountains and foothills region of Southern Alberta. *Atmosphere-Ocean*, 49(3), 189-205.
- Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology*, 88(11), 2783-2792.
- Danby, R. K. (2011). Monitoring Forest-Tundra Ecotones at Multiple Scales. *Geography Compass*, 5(9), 623-640.
- Dang, H., Zhang, Y., Zhang, Y., Zhang, K., & Zhang, Q. (2015). Variability and rapid response of subalpine fir (*Abies fargesii*) to climate warming at upper altitudinal limits in north-central China. *Trees*, 29(3), 785-795.
- Elliott, G. P. (2012). The Role of Thresholds and Fine-Scale Processes in Driving Upper Treeline Dynamics in the Bighorn Mountains, Wyoming. *Physical Geography*, 33(2), 129-145.
- Elliott, G. P., & Cowell, C. M. (2015). Slope Aspect Mediates Fine-Scale Tree Establishment Patterns at Upper Treeline during Wet and Dry Periods of the 20th Century. *Arctic, Antarctic, and Alpine Research*, 47(4), 681-692.
- Esri. (2014). ArcGIS 10.3 for Desktop: Environmental Systems Research Institute Redlands, CA USA.
- Fagre, D. B., Walsh, S. J., Malanson, G. P., & Butler, D. R. (2007). Influences of Geomorphology and Geology on Alpine Treeline in the American West—More Important than Climatic Influences? *Physical Geography*, 28(5), 434-450.
- Fenton, M., Waters, E., Pawley, S., Atkinson, N., Utting, D., & McKay, K. (2013). Surficial geology of Alberta. *Alberta Geological Survey, AER/AGS Map*, 601.
- Geological Survey of Canada, & Price, R. A. (1961). *Fernie map-area, east half, Alberta and British Columbia 82 GE 1/2 (Report and Map 35-1961)*.
- Glass, D. (1990). Lexicon of Canadian Stratigraphy, Volume 4, Western Canada, including Eastern British Columbia, Alberta, Saskatchewan and Southern Manitoba. *Canadian Society of Petroleum Geologists*, 8, p109-118.
- Grace, J., Berninger, F., & Nagy, L. (2002). Impacts of climate change on the tree line. *Annals of Botany*, 90(4), 537-544.
- Greenwood, S., Chen, J. C., Chen, C. T., & Jump, A. S. (2014). Strong topographic sheltering effects lead to spatially complex treeline advance and increased forest density in a subtropical mountain region. *Global Change Biology*, 20(12), 3756-3766.
- Harsch, M. A., Hulme, P. E., McGlone, M. S., & Duncan, R. P. (2009). Are treelines advancing? A global meta-analysis of treeline response to climate warming. *Ecology Letters*, 12(10), 1040-1049.
- Heiri, C., Bugmann, H., Tinner, W., Heiri, O., & Lischke, H. (2006). A model-based reconstruction of Holocene treeline dynamics in the Central Swiss Alps. *Journal of Ecology*, 94(1), 206-216.
- Holden, Z. A., Abatzoglou, J. T., Luce, C. H., & Baggett, L. S. (2011). Empirical downscaling of daily minimum air temperature at very fine resolutions in complex terrain. *Agricultural and Forest Meteorology*, 151(8), 1066-1073.

- Holtmeier, F. K., & Broll, G. (2005). Sensitivity and response of northern hemisphere altitudinal and polar treelines to environmental change at landscape and local scales. *Global Ecology and Biogeography*, *14*(5), 395-410.
- Jensen, M. E., & Haise, H. R. (1963). Estimating evapotranspiration from solar radiation. *Proceedings of the American Society of Civil Engineers, Journal of the Irrigation and Drainage Division*, *89*, 15-41.
- Kienzle, S. (2004). The effect of DEM raster resolution on first order, second order and compound terrain derivatives. *Transactions in GIS*, *8*(1), 83-111.
- Körner, C. (1998). A re-assessment of high elevation treeline positions and their explanation. *Oecologia*, *115*(4), 445-459.
- Körner, C., & Paulsen, J. (2004). A world-wide study of high altitude treeline temperatures. *Journal of Biogeography*, *31*(5), 713-732.
- Kueppers, L. M., Conlisk, E., Castanha, C., Moyes, A. B., Germino, M. J., de Valpine, P., Torn, M. S., & Mitton, J. B. (2017). Warming and provenance limit tree recruitment across and beyond the elevation range of subalpine forest. *Glob Chang Biol*, *23*(6), 2383-2395.
- Leonelli, G., Masseroli, A., & Pelfini, M. (2016). The influence of topographic variables on treeline trees under different environmental conditions. *Physical Geography*, *37*(1), 56-72.
- Liang, E., Wang, Y., Piao, S., Lu, X., Camarero, J. J., Zhu, H., Zhu, L., Ellison, A. M., Ciais, P., & Peñuelas, J. (2016). Species interactions slow warming-induced upward shifts of treelines on the Tibetan Plateau. *Proceedings of the National Academy of Sciences*, *113*(16), 4380-4385.
- Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R news*, *2*(3), 18-22.
- Lloyd, A. H., & Graumlich, L. J. (1997). Holocene dynamics of treeline forests in the Sierra Nevada. *Ecology*, *78*(4), 1199-1210.
- Macias-Fauria, M., & Johnson, E. A. (2013). Warming-induced upslope advance of subalpine forest is severely limited by geomorphic processes. *Proceedings of the National Academy of Sciences*, *110*(20), 8117-8122.
- Malanson, G. P., Butler, D. R., Cairns, D. M., Welsh, T. E., & Resler, L. M. (2002). Variability in an edaphic indicator in alpine tundra. *Catena*, *49*(3), 203-215.
- McCaffrey, D. R. (2018). *Assessing historic change in subalpine forest: a case study in the West Castle Watershed*. Lethbridge, Alta.: University of Lethbridge, Department of Geography.
- McCaffrey, D. R., & Hopkinson, C. (2017). Assessing Fractional Cover in the Alpine Treeline Ecotone Using the WSL Monoplotting Tool and Airborne Lidar. *Canadian Journal of Remote Sensing*, *43*(5), 504-512.
- McGuinness, J. L., & Bordne, E. F. (1972). *A comparison of lysimeter-derived potential evapotranspiration with computed values*: US Dept. of Agriculture.
- Mustaphi, C. J. C., Pisaric, M. F. J., & Williams, J. (2013). Varying influence of climate and aspect as controls of montane forest fire regimes during the late Holocene, south-eastern British Columbia, Canada. *Journal of Biogeography*, *40*(10), 1983-1996.

- Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., & Loumagne, C. (2005). Which potential evapotranspiration input for a lumped rainfall–runoff model? *Journal of Hydrology*, *303*(1-4), 290-306.
- Paulsen, J., & Körner, C. (2014). A climate-based model to predict potential treeline position around the globe. *Alpine Botany*, *124*(1), 1-12.
- Pepin, N., Bradley, R. S., Diaz, H. F., Baraer, M., Caceres, E. B., Forsythe, N., Fowler, H., Greenwood, G., Hashmi, M. Z., Liu, X. D., Miller, J. R., Ning, L., Ohmura, A., Palazzi, E., Rangwala, I., Schöner, W., Severskiy, I., Shahgedanova, M., Wang, M. B., Williamson, S. N., & Yang, D. Q. (2015). Elevation-dependent warming in mountain regions of the world. *Nature Climate Change*, *5*(5), 424-430.
- R Core Team. (2017). A language and environment for statistical computing R Foundation for Statistical Computing Department of Agronomy, Faculty of Agriculture of the University of the Free State. *Vienna, Austria*. www.R-project.org.
- Randin, C. F., Engler, R., Normand, S., Zappa, M., Zimmermann, N. E., Pearman, P. B., Vittoz, P., Thuiller, W., & Guisan, A. (2009). Climate change and plant distribution: local models predict high-elevation persistence. *Global Change Biology*, *15*(6), 1557-1569.
- Rogeanu, M.-P., & Armstrong, G. W. (2017). Quantifying the effect of elevation and aspect on fire return intervals in the Canadian Rocky Mountains. *Forest Ecology and Management*, *384*, 248-261.
- Schwörer, C., Gavin, D. G., Walker, I. R., & Hu, F. S. (2017). Holocene tree line changes in the Canadian Cordillera are controlled by climate and topography. *Journal of Biogeography*, *44*(5), 1148-1159.
- Schwörer, C., Henne, P. D., & Tinner, W. (2014). A model-data comparison of Holocene timberline changes in the Swiss Alps reveals past and future drivers of mountain forest dynamics. *Global Change Biology*, *20*(5), 1512-1526.
- Stine, M. B., & Butler, D. R. (2015). Effects of fire on geomorphic factors and seedling site conditions within the alpine treeline ecotone, Glacier National Park, MT. *Catena*, *132*, 37-44.
- Stockdale, C. A., Macdonald, S. E., & Higgs, E. (2019). Forest closure and encroachment at the grassland interface: a century-scale analysis using oblique repeat photography. *Ecosphere*, *10*(6), e02774.
- Strobl, C., Boulesteix, A.-L., Zeileis, A., & Hothorn, T. (2006). Bias in Random Forest Variable Importance Measures: Illustrations, Sources and a Solution.
- Stueve, K. M., Cerney, D. L., Rochefort, R. M., & Kurth, L. L. (2009). Post-fire tree establishment patterns at the alpine treeline ecotone: Mount Rainier National Park, Washington, USA. *Journal of Vegetation Science*, *20*(1), 107-120.
- Trant, A. J., Starzomski, B. M., & Higgs, E. (2015). A publically available database for studying ecological change in mountain ecosystems. *Frontiers in Ecology and the Environment*, *13*(4), 187-187.
- Walsh, S. J., & Butler, D. R. (1997). Morphometric and multispectral image analysis of debris flows for natural hazard assessment. *Geocarto International*, *12*(1), 59-70.
- Walsh, S. J., Butler, D. R., Allen, T. R., & Malanson, G. P. (1994). Influence of snow patterns and snow avalanches on the alpine treeline ecotone. *Journal of Vegetation Science*, *5*(5), 657-672.

- Wei, T., & Simko, V. (2016). corrplot: Visualization of a Correlation Matrix. R package version 0.77. CRAN, Vienna, Austria.
- Weiss, D. J., Malanson, G. P., & Walsh, S. J. (2015). Multiscale Relationships Between Alpine Treeline Elevation and Hypothesized Environmental Controls in the Western United States. *Annals of the Association of American Geographers*, 105(3), 437-453.
- Wildfire Management Branch - Alberta Agriculture and Forestry. (2017). *Fire History Polygons C5 FMU*.
- Williams, C. K., Engelhardt, A., Cooper, T., Mayer, Z., Ziem, A., Scrucca, L., Tang, Y., Candan, C., Hunt, T., & Kuhn, M. M. (2017). Package ‘caret’.
- Yadava, A. K., Sharma, Y. K., Dubey, B., Singh, J., Singh, V., Bhutiyani, M. R., Yadav, R. R., & Misra, K. G. (2017). Altitudinal treeline dynamics of Himalayan pine in western Himalaya, India. *Quaternary International*, 444, 44-52.
- Zevenbergen, L. W., & Thorne, C. R. (1987). Quantitative analysis of land surface topography. *Earth Surface Processes and Landforms*, 12(1), 47-56.

Tables

Table 1.1 - Summary of variables used in random forest model of ATE change.

| Variable | Abbreviation | Unit | Mean | Std dev. | Min | Max |
|----------------------------|--------------|---|------------|-----------|-----------|-------------|
| <i>Continuous</i> | | | | | | |
| <i>Topographic</i> | | | | | | |
| Elevation | ELEV | m | 1843.20 | 263.24 | 1391.52 | 2624.44 |
| Compound Topographic Index | CTI | unitless | 5.86 | 2.07 | 1.46 | 23.16 |
| Curvature | CURV | unitless | 0.00 | 2.24 | -62.62 | 69.99 |
| Slope | SLOPE | ° | 26.42 | 13.59 | 0.00 | 87.38 |
| <i>Climatic</i> | | | | | | |
| Annual Precipitation | ANNPREC | mm | 1477.41 | 279.38 | 901.04 | 2298.08 |
| Annual Solar Radiation | ANNSOLR | Wh/m ² | 1103737.10 | 233753.55 | 169984.64 | 15990846.00 |
| Annual Temperature | ANNTEMP | °C | 2.01 | 1.07 | -1.12 | 3.93 |
| Summer Potential | | | | | | |
| Evapotranspiration | JJAPET | mm/day | 1.27 | 0.22 | 0.16 | 1.58 |
| Summer Precipitation | JJAPREC | mm | 255.02 | 34.53 | 195.68 | 357.95 |
| Summer Solar Radiation | JJASOLR | Wh/m ² | 505483.39 | 80525.16 | 69832.80 | 652794.63 |
| Summer Temperature | JJATEMP | °C | 10.97 | 1.06 | 7.83 | 12.77 |
| <i>Categorical</i> | | | | | | |
| Aspect | ASPECT | 9 classes: N,NE,E,SE,S,SW,W,NW,Flat | | | | |
| Fire | FIRE | 2 classes: Fire Exposed, Non-Fire-Exposed | | | | |
| Geological Formation | SUBGEO | 9 classes: Phillips Formation, Gateway Formation, Sheppard Formation, Purcell Lava, Siyeh Formation, Grinnell Formation, Appekunny Formation, Altyn Formation, Waterton Formation | | | | |
| Surficial Geology | SURFGEO | 5 classes: Bedrock, Colluvial Deposit, Fluvial Deposit, Glaciofluvial Deposit, Moraine | | | | |

Table 1.2 - Summary of Meteorological Station Attributes

| Station Name* | Latitude | Longitude | Elevation (m) | Years on Record | Dist. to WCW (km) |
|----------------------|-----------------|------------------|----------------------|------------------------|--------------------------|
| Castle | 49.40 | -114.33 | 1360 | 1958-2007 | 7.8 |
| Westcastle | 49.28 | -114.37 | 1524 | 1999-2007 | 0.0 |
| Pincher Creek Town | 49.52 | -113.97 | 1154 | 1893-1963 | 36.6 |
| Spionkop | 49.22 | -114.08 | 1861 | 1998-2007 | 17.3 |
| Gardiner | 49.37 | -114.52 | 1920 | 1998-2007 | 5.1 |
| Coleman | 49.63 | -114.58 | 1341 | 1912-1997 | 32.5 |
| Beaver Mines Town | 49.47 | -114.18 | 1257 | 1912-2007 | 21 |

Table 1.3 - Results of Lapse Rate Regression

| | slope | y-int | r² | F | p |
|-------------------------|--------------|--------------|----------------------|----------|----------|
| Annual Precipitation | 1.050 | -650.29 | 0.77 | 20.73 | 0.006 |
| Summer Precipitation | 0.130 | 18.20 | 0.74 | 17.81 | 0.008 |
| Annual Temperature | -0.004 | 8.79 | 0.65 | 12.25 | 0.017 |
| Summer Temperature | -0.004 | 19.01 | 0.74 | 17.81 | 0.008 |

Table 1.4 - Correlated variable groups.

| | Group 1 | Group 2 | Group 3 |
|-----------------------------|--|------------------------------|-----------------|
| <u>Correlated Variables</u> | ANNPREC ANNTEMP ELEV JJAPREC JJATEMP | ANNSOLR JJASOLR JJAPET | JJAPET SLOPE |
| <u>Selected Variable</u> | ELEV | JJASOLR | SLOPE |

Table 1.5 - Within-training-data confusion matrix and accuracy assessment.

| Prediction | Reference | | | | | | | Including 0 | | Excluding 0 | |
|-----------------|-----------|------|------|-------|------|------|-----|-------------|-------------|-------------|-------------|
| | -3 | -2 | -1 | 0 | 1 | 2 | 3 | Sensitivity | Specificity | Sensitivity | Specificity |
| -3 | 1381 | 149 | 155 | 107 | 12 | 1 | 0 | 0.568 | 0.995 | 0.841 | 0.972 |
| -2 | 58 | 1500 | 82 | 83 | 25 | 11 | 0 | 0.394 | 0.997 | 0.778 | 0.984 |
| -1 | 198 | 244 | 3325 | 262 | 90 | 46 | 4 | 0.450 | 0.990 | 0.910 | 0.936 |
| 0 | 790 | 1880 | 3729 | 63802 | 3530 | 1758 | 316 | 0.988 | 0.516 | - | - |
| 1 | 3 | 32 | 82 | 271 | 3322 | 123 | 70 | 0.471 | 0.993 | 0.943 | 0.967 |
| 2 | 2 | 4 | 13 | 62 | 71 | 1411 | 18 | 0.420 | 0.998 | 0.883 | 0.990 |
| 3 | 0 | 0 | 1 | 5 | 4 | 6 | 373 | 0.478 | 0.999 | 0.803 | 0.999 |
| Accuracy | | | | | | | | 0.840 | | 0.883 | |
| Kappa | | | | | | | | 0.577 | | 0.850 | |

Table 1.6 - Confusion Matrix Accuracy Results.

| | Accuracy | 95% CI | n | p-value | Kappa |
|-------------------|----------|---------------|--------|---------|-------|
| All | 0.840 | 0.838 - 0.843 | 89 440 | < 0.001 | 0.577 |
| High Elev. | 0.761 | 0.756 - 0.764 | 38 014 | < 0.001 | 0.490 |
| Low Elev. | 0.826 | 0.823 - 0.829 | 51 426 | < 0.001 | 0.300 |

Table 1.7 - Variable Importance MDG Values.

| Variable | All | High Elev. | Low Elev. | Difference (High Elev. – Low Elev.) |
|----------|---------|------------|-----------|-------------------------------------|
| FIRE | 472.18 | 478.59 | 152.28 | 326.32 |
| SUBGEO | 991.52 | 643.29 | 416.94 | 226.35 |
| JJASOLR | 2413.21 | 1403.11 | 1204.39 | 198.72 |
| ASPECT | 988.08 | 590.13 | 416.54 | 173.58 |
| CURV | 1870.38 | 1108.34 | 977.3 | 131.04 |
| CTI | 1996.19 | 1167.64 | 1059.17 | 108.47 |
| SURFGEO | 521.44 | 352.2 | 287.43 | 64.76 |
| SLOPE | 2249.92 | 1272.1 | 1280.58 | -8.47 |
| ELEV | 3435.21 | - | - | - |

Figure captions

- Figure 1 Land cover change in WCW. Areas of anthropogenic disturbance, bounded in black lines, were omitted from training data in the present model.
- Figure 2 Meteorological Stations in the West Castle Region.
- Figure 3 Length of data record for headwater meteorological stations
- Figure 4 Annual and Summer (JJA) Temperature Lapse Rates.
- Figure 5 Annual and Summer (JJA) Precipitation Enhancement.
- Figure 6 Nine variables used in the random forest model: a) elevation, b) compound topographic index, c) standard curvature, d) slope, e) summer solar radiation, f) aspect, g) fire occurrence, h) bedrock geology, i) surficial geology.
- Figure 7 Correlation matrix of continuous variables.
- Figure 8 Result of canopy change spatial extension into unobserved areas of WCW. Grey areas are omitted anthropogenic disturbance.
- Figure 9 Variable Importance Ranked by MDG.

Figure 1 Land cover change in WCW. Areas of anthropogenic disturbance, bounded in black lines, were omitted from training data in the present model.

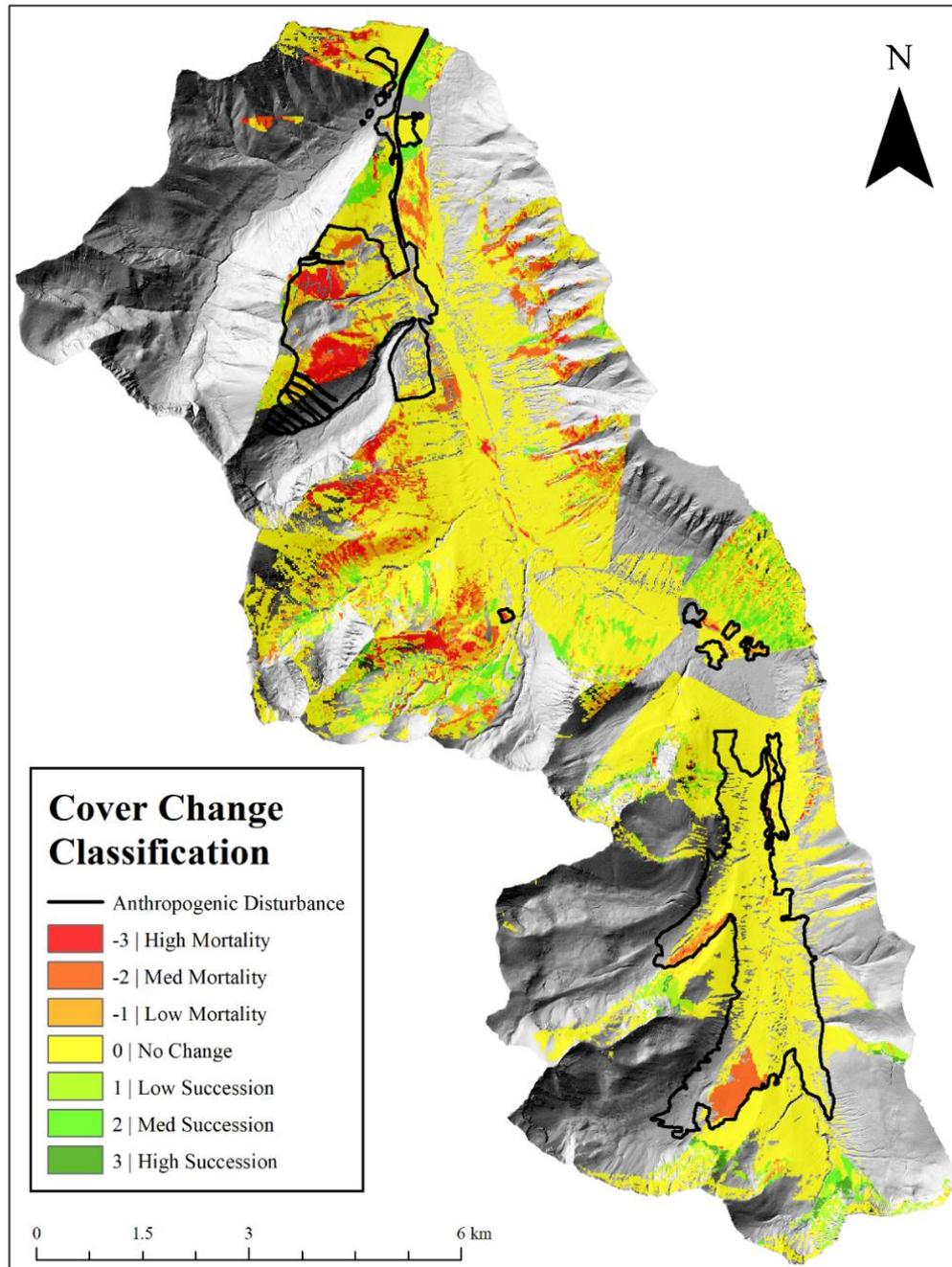


Figure 2 Meteorological Stations in the West Castle Region.

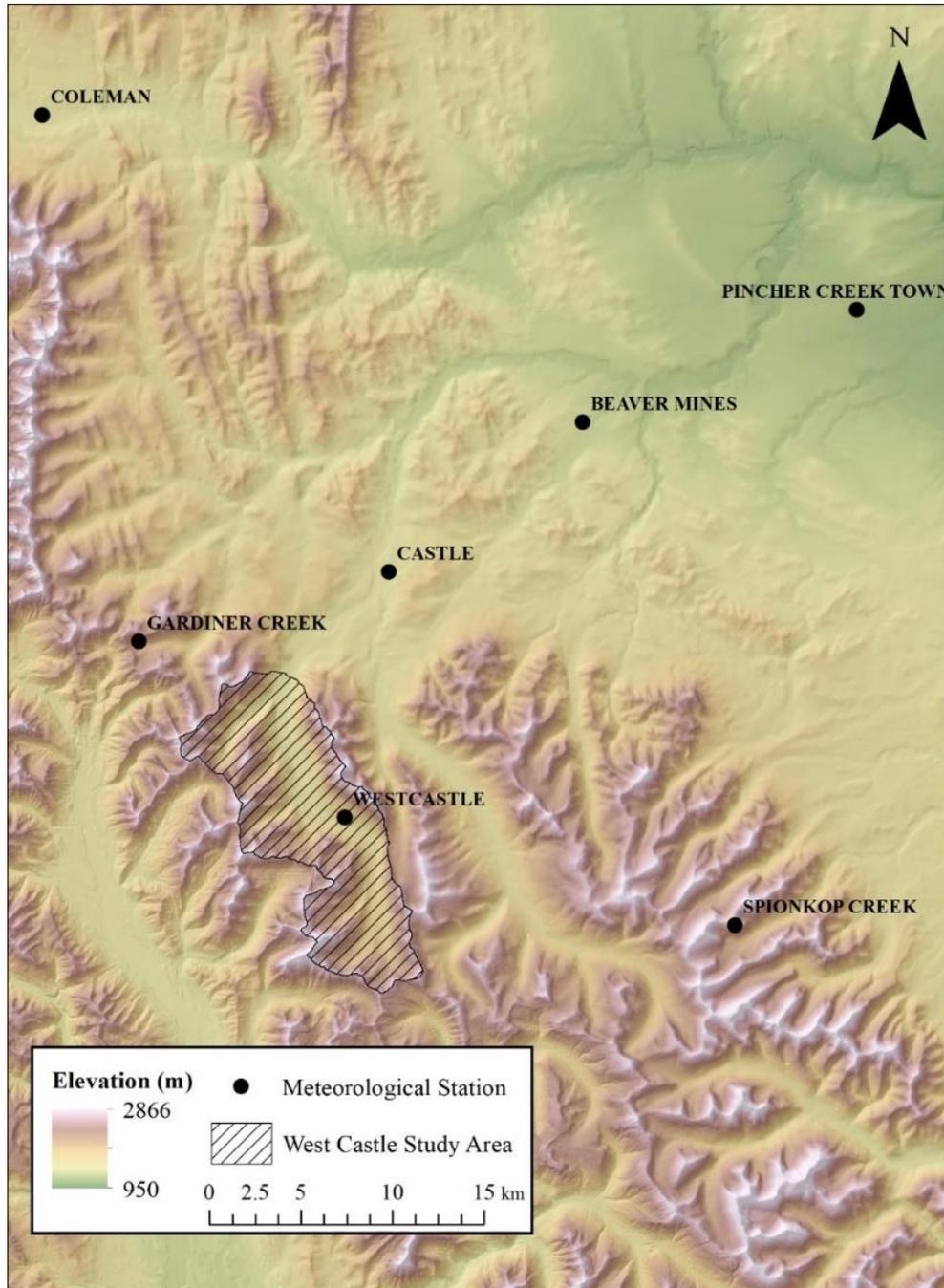


Figure 3 Length of data record for headwater meteorological stations

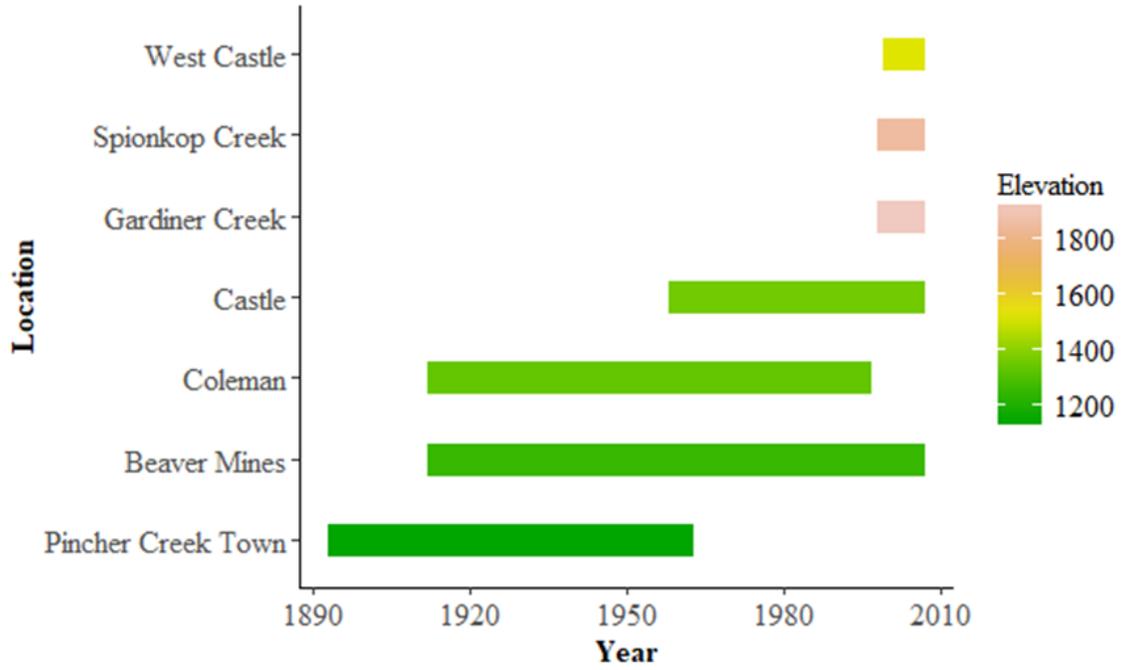


Figure 4 Annual and Summer (JJA) Temperature Lapse Rates.

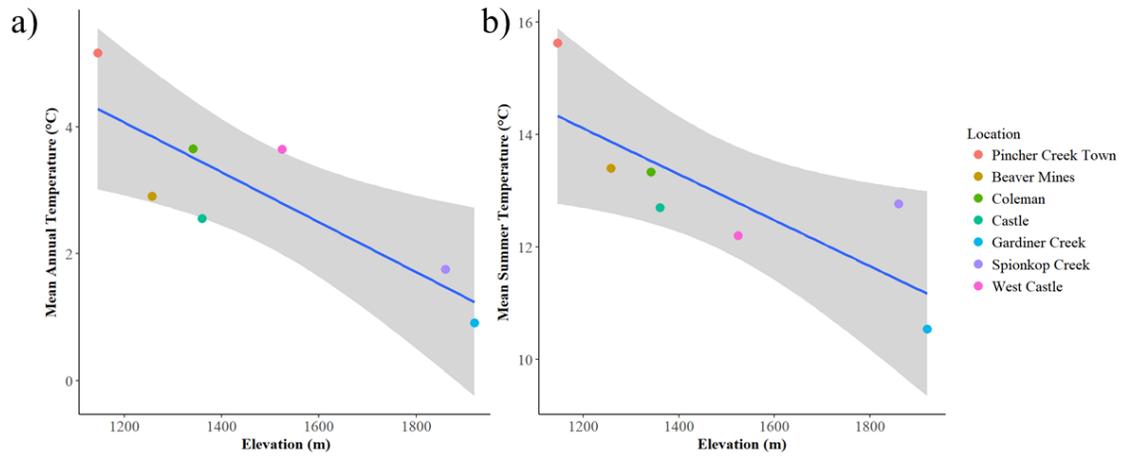


Figure 5 Annual and Summer (JJA) Precipitation Enhancement.

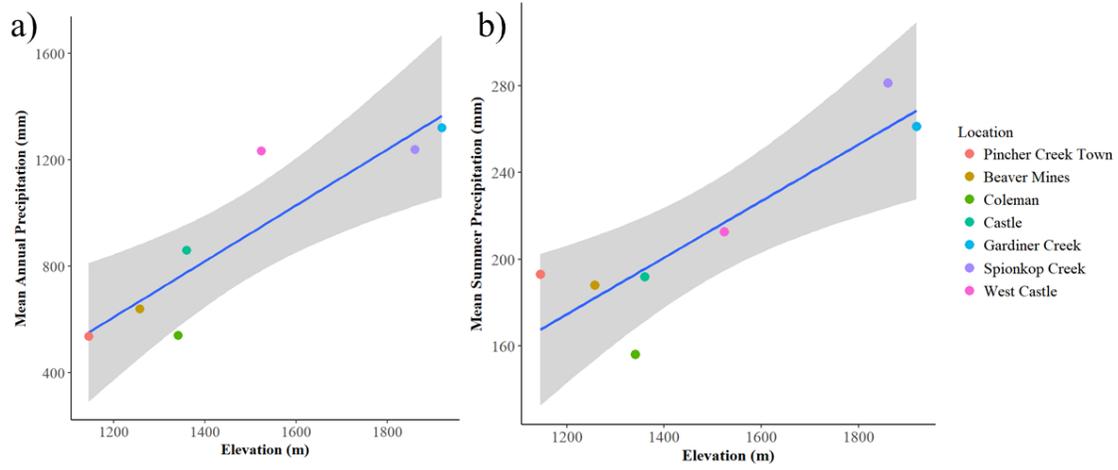


Figure 6 Nine variables used in the random forest model: a) elevation, b) compound topographic index, c) standard curvature, d) slope, e) summer solar radiation, f) aspect, g) fire occurrence, h) bedrock geology, i) surficial geology.

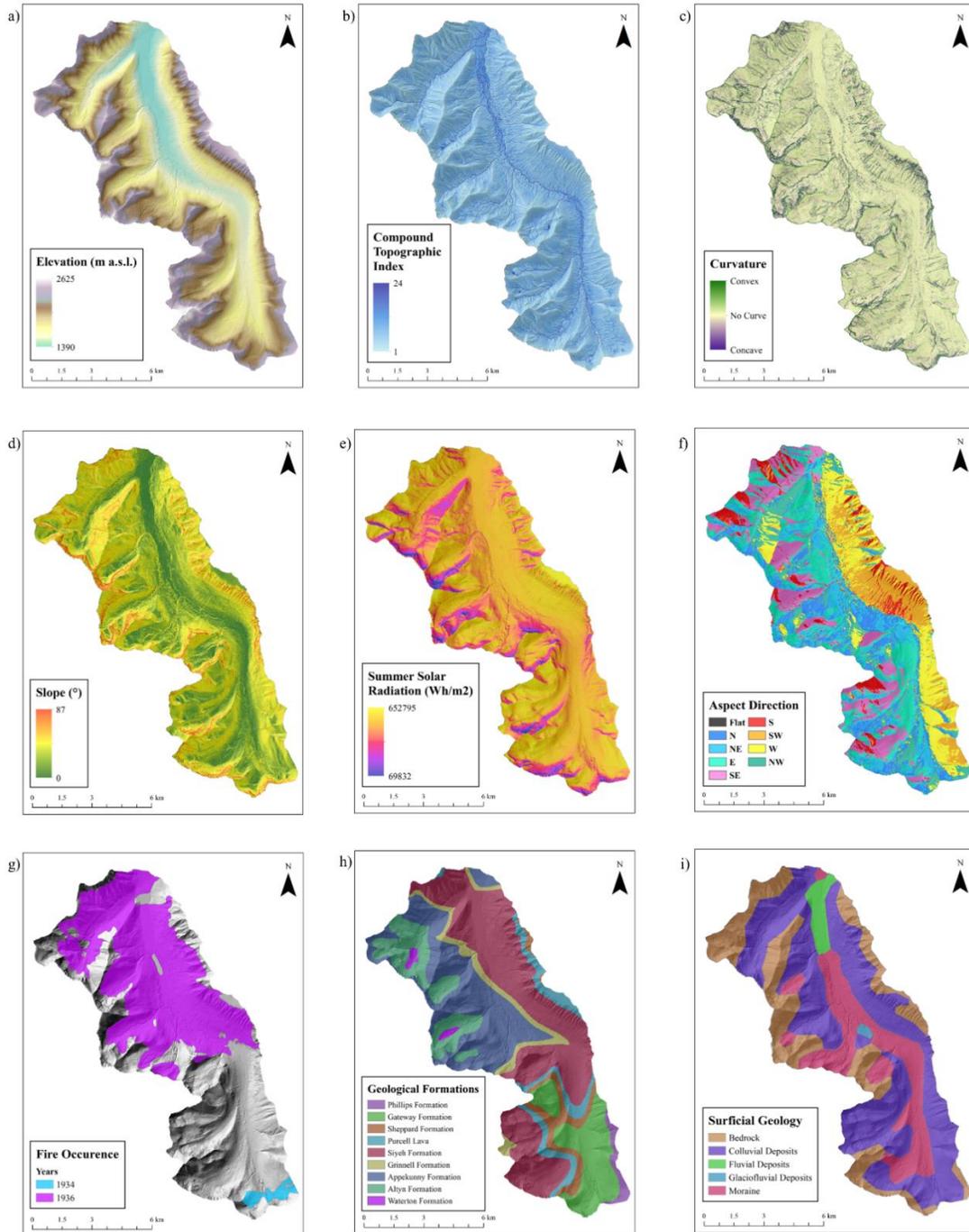


Figure 7 Correlation matrix of continuous variables.

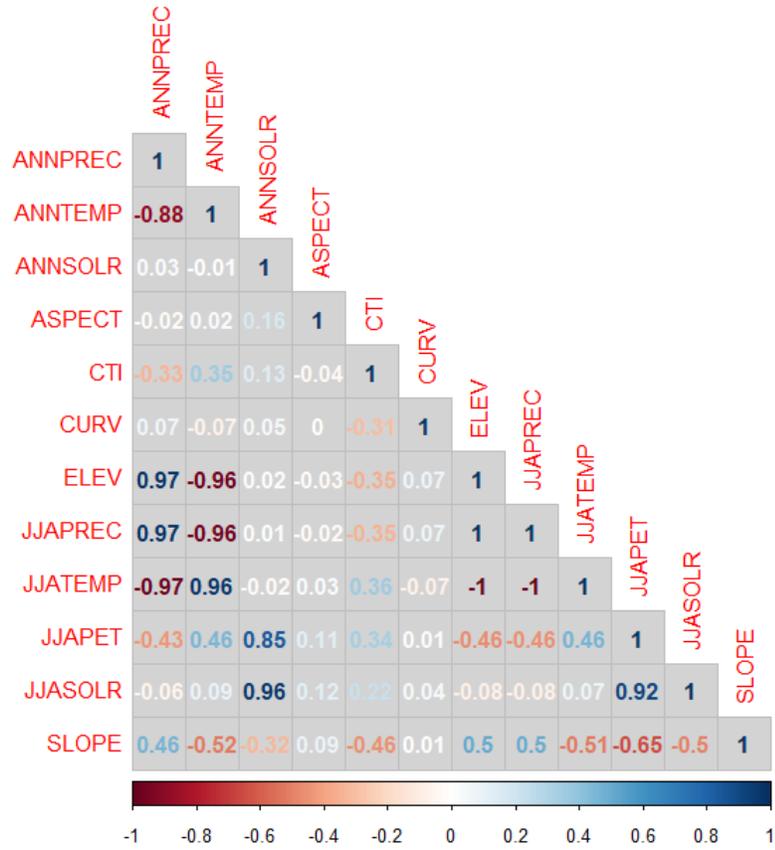


Figure 8 Result of canopy change spatial extension into unobserved areas of WCW. Grey areas are omitted anthropogenic disturbance.

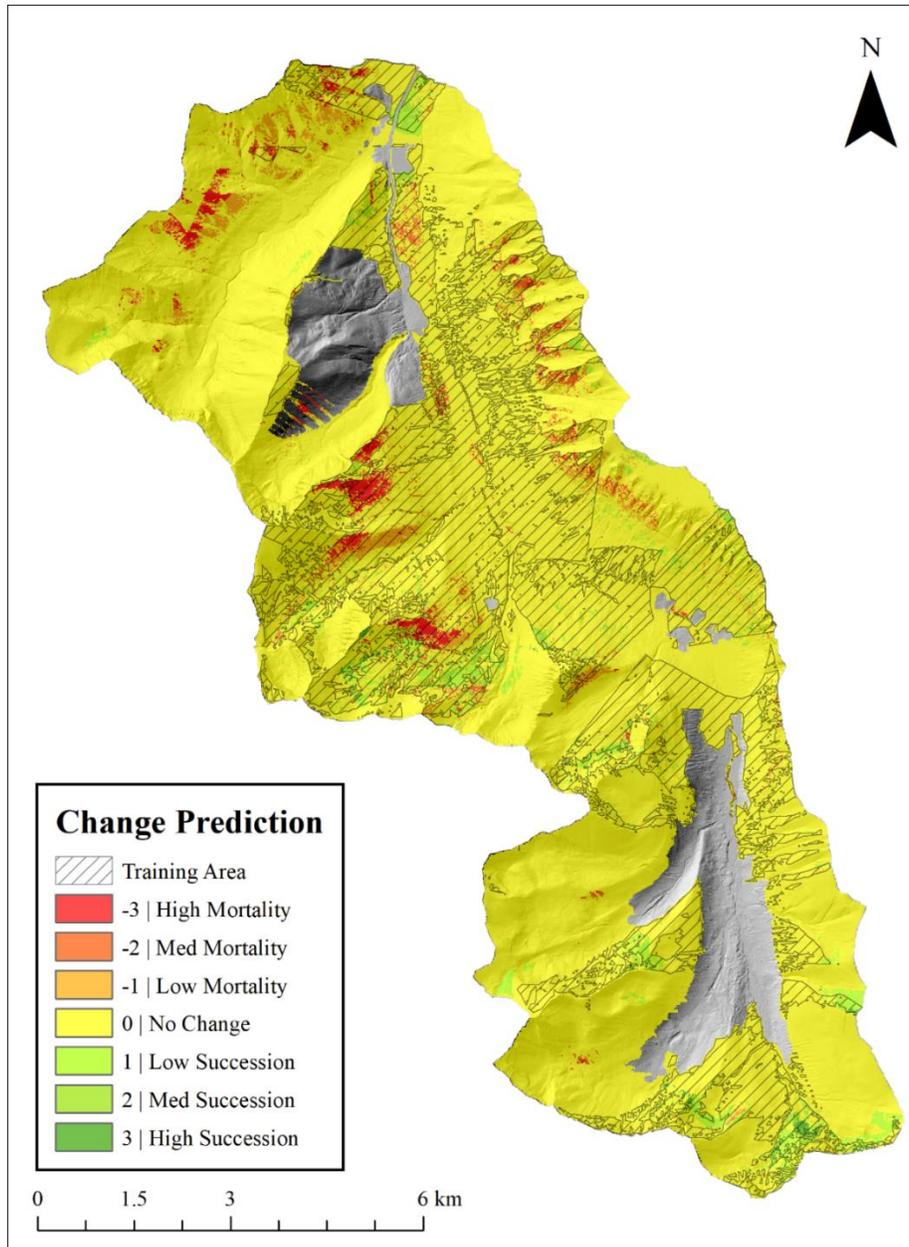
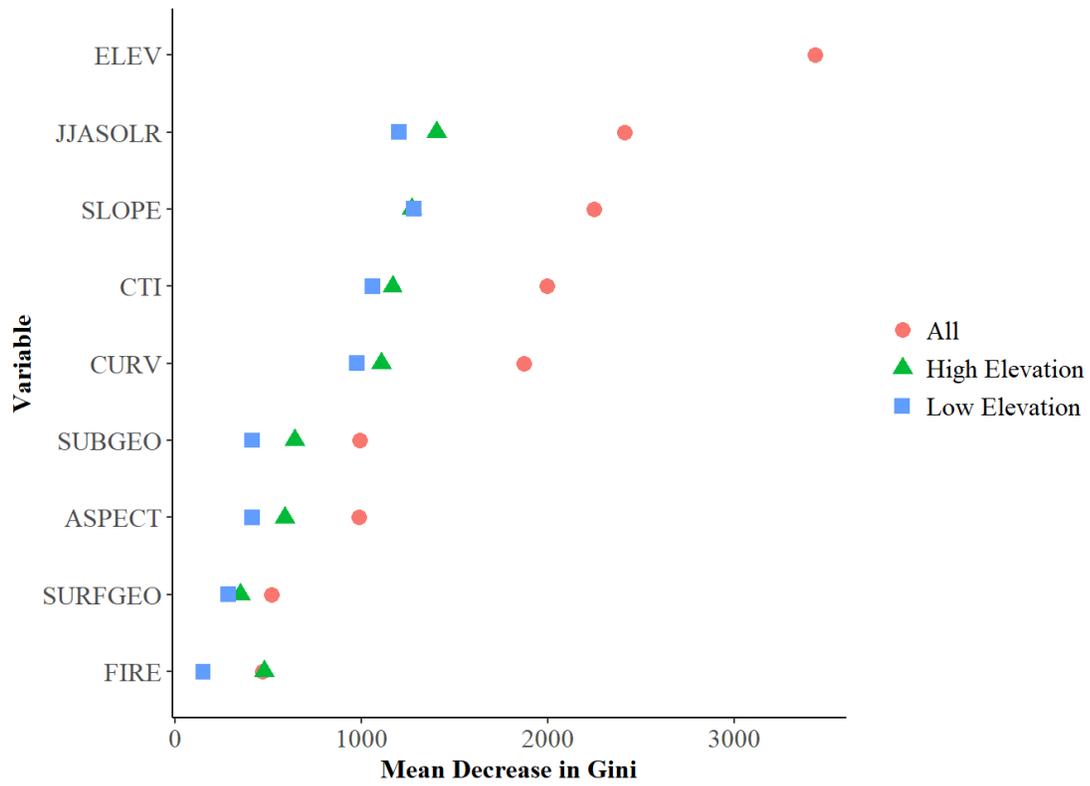


Figure 9 Variable Importance Ranked by MDG.



Appendix 4: Evaluating Lidar-estimated snow depth drivers and their consistency during mid winter and melt onset for a Rocky Mountain watershed

Cartwright, K., Hopkinson, C., Kienzle, S., Rood, S. EVALUATING LIDAR-ESTIMATED SNOW DEPTH DRIVERS AND THEIR CONSISTENCY DURING MID WINTER AND MELT ONSET FOR A ROCKY MOUNTAIN WATERSHED. *Hydrological Processes*. (In Review)

Evaluating Lidar-estimated snow depth drivers and their consistency during mid winter and melt onset for a Rocky Mountain watershed

Cartwright, K., Hopkinson, C., Kienzle, S., Rood, S

Abstract

Collecting spatially representative data over large areas is a challenge within snow monitoring frameworks. Identifying consistent trends in snow depth enables increased sampling efficiency by minimizing field collection time and costs associated with remote sensing methods. Seasonal snowpack depth estimations during mid-winter and melt onset conditions were derived from airborne Lidar over the West Castle Watershed in the southern Canadian Rockies over three years. Each data set was divided into five snow depth driver classes: elevation, aspect, topographic position index, canopy cover and slope. Although mid-winter class trends for each driver were similar, depth distributions within each class were significantly different due to recent snowfall events, redistribution and settling processes. Trends in driver class during late season melt onset were also similar to mid-winter but varied with seasonality. This is due to the differing stages of accumulation or ablation and the upward migration in the 0°C isotherm during spring, when snow depth can be declining in valley bottoms while still increasing at higher elevations. The observed consistency in depth driver controls and identification of zones of characteristically high and low snow storage can be used to guide future integrated snow monitoring frameworks.

Keywords: snow depth distribution, accumulation and ablation, Lidar, remote sensing, mountain hydrology, snow monitoring.

1. Introduction

Seasonal snow pack accumulation in the Rocky Mountains and other mountain watersheds is crucial to water security in western North America, and in many other regions worldwide. The snow pack accumulates over the winter season and subsequently melts in the spring through autumn providing the primary contribution to river flows, which provide the primary water supplies for municipal and industrial use, and irrigation. Estimates of headwater snow pack volumes are essential for water budget calculations that influence downstream cropping plans and provide advance warning of water shortages. Mountain region snow packs also influence spring flood risk, with rapid snow melt causing high river flows, and rain-on-snow events resulting in potentially severe flooding (Fang and Pomeroy, 2016).

Early snow pack monitoring involved periodic measurements of snow depths in snow course sites and subsequently, snow pillows with automated measurements of snow water equivalent (SWE) typically at higher elevation sites that accumulate deep snow packs. However, snow course or snow pillows provide snow pack measurements at single locations, the choice of which being confounded by the inherent variation in snow depth across mountain regions. Snow pack monitoring with airborne Lidar (light detection and ranging) can provide depth observations over whole watersheds (Hopkinson et al. 2012), and has the potential to improve inventories of snow pack extent and volume to enhance water supply and flood risk forecasts (Deems et al. 2013). Understanding the spatiotemporal variation of snow depth with respect to terrain and land

cover drivers is necessary to developing an effective spatially-explicit headwater snowpack monitoring framework that integrates depth, density and SWE observations from snow courses, snow pillows and airborne sampling . While basin-wide SWE is the desired goal of many water resource monitoring efforts, the focus of this study is snow depth, as of the two variables (depth and density) required to estimate SWE, depth is the more spatially variable and thus more critical in SWE determination (Dickinson and Whiteley, 1972; Steppuhn 1976).

Upper alpine zones tend to contain the most exposed and steep terrain in a catchment which promotes snow redistribution by wind and gravity, therefore flatter terrain is expected to collect more snow. The effects of orographic precipitation on snow distribution varies from windward to leeward slopes (Pomeroy and Bruin, 2001), and Roe and Baker (2006) suggest that maximum precipitation rates have high spatiotemporal variability which makes it difficult to assign a static location for maximum precipitation inputs. However, snowpack depth distributions with elevation are not solely due to orography and lapse rates but also the redistribution of snow by wind and gravity, followed by interception by vegetation (Greene et al. 1999; Hopkinson et al, 2010; 2012; Grunewald et al. 2014). Wind and gravity also promote the settlement of snow in areas of topographic transition from exposed to concave terrain (Revuelto et al. 2014; Lopez-Moreno et al. 2017). Therefore, it is expected that deeper snowpack will occur within the treeline and within topographic depressions.

Aspect is another well-researched variable that impacts snow depth distributions. Short and long wave radiation provide the energy for melt which can create preferential depth losses and potential density gains in aspects with high radiative inputs versus those with little (Golding and Swanson 1986; Anderson et al. 1958). This effect holds true at all elevations (D'Eon, 2004). Many studies have demonstrated higher snowmelt and sublimation rates in northern hemisphere south-facing relative to north-facing slopes (Haupt, 1951; Anderson and West, 1965; Hendrick et al. 1971; Rowland and Moore, 1992; D'Eon, 2004; Jost et al. 2007). Greater mean depths should (all else being equal) occur on north facing terrain with the lowest depths on south facing terrain where northern hemisphere shortwave radiative inputs are greater.

Studies of vegetation influences that examined both snow accumulation and melt rates in forested areas have found that melt rates in south facing sites with and without forest cover exceed those of clearcuts with a northerly aspect, confirming that aspect has a greater effect on melt than forest cover (Murray and Buttle, 2003; Haupt, 1951; Anderson and West, 1965; Hendrick et al. 1971; Rowland and Moore, 1992; D'Eon, 2004; Jost et al. 2007). Vegetation, which has a lower albedo than surrounding snow, can enhance melt by absorbing solar energy, which heats up stems and foliage and then re-emits energy as longwave radiation over nearby snowpack (e.g. Ellis et al. 2011). Vegetation also promotes ablation by intercepting snow and leaving it exposed on branches to wind and radiation that may drive sublimation losses (Pomeroy et al. 1998). Considering these effects, closed canopies should have the lowest depths, especially later in the season as day length and shortwave energy inputs increase. More open canopies should exhibit greater snow accumulation due to wind trapping, reduced longwave radiation and lower amounts of intercepted snow being subject to sublimation prior to falling to the ground.

1.2 Objectives

To support optimal Lidar-based mountain snow depth sampling that delivers spatially representative results, the priority in snow depth drivers and the variation in their distributions needs to be known for a given region and time of year. The over-arching goal of this analysis is

to determine if well-established snow depth distribution trends, with respect to various terrain and land cover-based snow depth controls, exist in the headwaters of the Oldman River Basin, southern Alberta. This is assessed by examining the seasonal and interannual consistency in class-based snow depth drivers during mid-winter and at melt onset. Lidar-based snow depth data are compared to publicly available in situ snow records, to place the government monitoring data into the watershed context. This information will ultimately support decisions over the optimal timing and sampling strategy required for an airborne Lidar-based snow monitoring framework. For example, if a high proportion of watershed-scale snow volume is found to be associated with certain landcover and/or terrain types or combinations thereof, then ensuring representative sampling of such watershed regions would be desirable in an operational aerial monitoring framework. Evaluating the temporal consistency of high and low snow storage zones can determine if aerial sampling needs to adapt to seasonal factors, such as snow accumulation- vs. ablation-dominated periods. By examining spatial and temporal trends in snow depth and relationships with various drivers at the watershed scale over multiple years, a foundation for future sampling requirements is established under mid-winter (accumulation) and melt onset (ablation) snowpack conditions.

Thus, to better understand the processes that influence Rocky Mountain snowpacks and to explore methodologies to improve snow pack monitoring, this study was undertaken with two primary objectives:

- (1) to assess the relative influences of different physical environmental drivers on snow depth over a full watershed; and
- (2) to investigate the consistency of snow pack patterns and drivers within and across years.

1.3 Study Area

The West Castle watershed (WCW) is a mountainous (~1400 to 2600 m a.s.l. elevation range) sub-watershed of the Oldman River Basin in Southern Alberta near the common border between Alberta and British Columbia, and Montana, USA. This region is referred to as the Crown of the Continent and it includes the hydrographic apex of North America, the common headwater zone for rivers that flow to the Gulf of Mexico and Atlantic Ocean, Pacific Ocean, and Hudson Bay and Arctic Ocean (Rood et al., 2005). The WCW is relatively natural but with land use impacts that are typical of mountain regions, including some forest harvesting and livestock grazing, roadways and linear clearings for power lines, seismic lines and pipelines. It is sparsely settled but does include a substantial (~1,000 m vertical) ski hill resort. The scale of these human impacts is sufficient for assessment with Lidar analyses, but that was not the emphasis of this initial investigation of environmental determinants of snow pack patterns.

The Oldman River Basin is a transboundary basin with headwater zones to the south of the WCW in the Waterton-Glacier International Peace Park, the world's first such park (1932), which has been designated as a UNESCO World Heritage Site. The Oldman River is Canada's most extensively utilized river for irrigation and consequently water supply forecasting is especially critical. Within the Oldman headwaters ~8.4% of annual runoff yield is derived from the Castle Watershed, which includes the West and South Castle rivers, (Kienzle and Mueller, 2013), making it the 3rd highest yielding sub-basin. The ~103 km² watershed is characterized by a large valley that runs north and northwest-ward, and includes a sequence of ecoregions

extending upward from the aspen parkland, to montane forest, subalpine forest and to the exposed rocky alpine summits (Figure 1; Samuelson and Rood, 2004). A snow course exists within the watershed, which is manually surveyed every month throughout the winter season, and an automated snow pillow is situated in an adjacent mountain valley, with very similar biogeographical conditions. The WCW provides an ideal case study relative to snow pack analyses, and the findings should be highly applicable to other watersheds in the Crown of the Continent and elsewhere in the Rocky Mountains, which provide critical water resources for much of western North America.

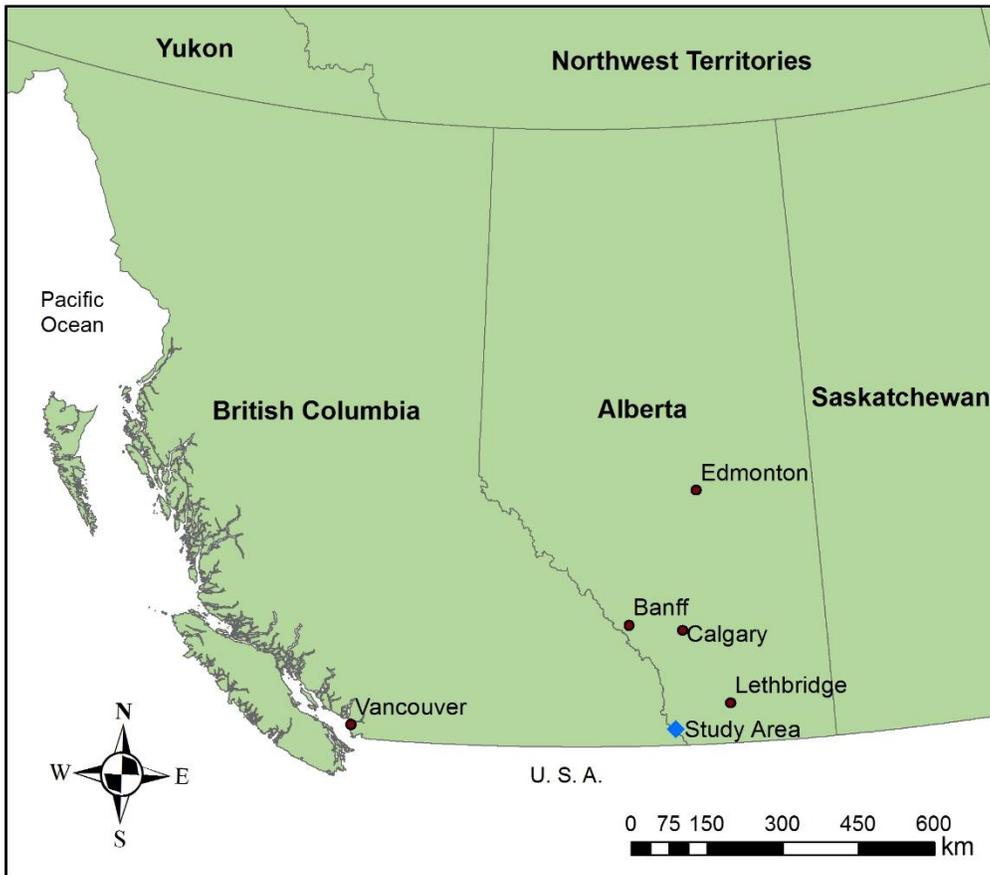


Figure 1. Study area, which is located near the Alberta/British Columbia border in southwestern Alberta, Canada.

2. Methods

2.1 Lidar Analysis

Winter flights to generate digital snow surface models (DSSMs) during snow- on conditions and another survey in summer to produce a digital elevation model (DEM) are required for Lidar snow studies. A LSDM (Lidar snow depth model) was created at a 1m grid cell resolution by subtracting the DEM from the DSSM (Eq 1, Hopkinson et al. 2004):

$$\text{LSDM} = \text{DSSM} - \text{DEM} \tag{Eq. 1}$$

The snow-off DEM was generated from a September 2014 survey of the same area when perennial snow was at its minimum. To produce snow-on DSSMs, airborne Lidar surveys were conducted in February 24, 2014 (mid-winter 2014), April 7, 2016 (melt onset 2016), and February 13, 2017 (mid-winter 2017). A small validation dataset exists for mid-winter 2014, which was limited to valley floor elevations <1450 m a.s.l. The root mean square error (RMSE) of this dataset is 0.27 m (n = 43), which was measured using survey grade differential GPS positioning of depth measurements using graduated depth probes. A snow depth validation dataset was collected along an elevation gradient in 2016 where the RMSE was 0.25 m (n = 165). Due to safety and logistical constraints, snow depth validation data were not collected for all surveys but all Lidar datasets were calibrated to flat runway surfaces 35 km west of the study site at the start and end of each survey flight, so the depth accuracy is expected to be consistent across the surveys.

A well-calibrated Lidar snow survey can produce a decimeter-level root mean square error, especially in zones of open canopy (Hopkinson et al. 2004; Grunewald et al. 2010; Lehning et al. 2011) with many studies reporting values ranging up to 0.3m and varying with landcover and terrain morphology (Geist and Stotter, 2008; Moreno Banos et al. 2009; DeBeer and Pomeroy, 2010; Hopkinson et al. 2011; Grunewald et al. 2013). Field observations were within the range of uncertainty found in other similar studies, confirming the suitability of our data for a more in depth assessment of drivers and their consistency. All datasets were gridded at a 1m spatial resolution using triangulation interpolation after being quality controlled and filtered following standard cleaning and classification methods described in Hopkinson et al (2012).

2.2 Snow Depth Driver Classes

Data collected in September 2014 was used to create a DEM, which was the input for creating snow depth control variable classes. Given the almost 1400 m of elevation range in the WCW and high local variation in snow depth, elevation was divided into 100m increments to identify broad elevation patterns, while mitigating against local terrain and landcover-induced noise. Slope, aspect and TPI (Topographic Position Index) are controls that require neighbourhood functions to derive the surfaces from a DEM. To ensure that cells along the outer edge of the area of interest had a sufficient neighbourhood, they were computed on a DEM that extended beyond the area of interest and then clipped to the watershed extent. Neighbourhood size for slope and aspect calculations is limited to the eight grid cells immediately surrounding a given cell, whereas the neighbourhood settings for TPI are user-defined. Using ArcGIS 10.4 (ESRI, Redlands, California), slope (Eq 2) and aspect were calculated with the aspect and slope tools.

$$\text{Slope } (^{\circ}) = \arctan(\text{rise} / \text{run}) \quad (\text{Eq } 2)$$

TPI is calculated by comparing a cell's elevation to that of its neighbours within a user defined window (Eq 3; Weiss, 2001).

$$\text{TPI} = E - M \quad (\text{Eq } 3)$$

Where E is the elevation of an individual grid cell and M is the mean elevation of grid cells within the window. TPI values are unique to the DEM used as an input so they are

generally stratified into three classes: depressions, uplands and transitional terrain which can be flat or sloped. Too small of a window may create unnecessary detail within the three TPI classes whereas too large of a window will include complex terrain features that don't represent the simple ridge / depression type of relief or the snow depth driver processes that we are aiming to represent. Previous research has found that study areas with larger elevation ranges have snow depths that correlate best with larger TPI windows (Revuelto et al. 2014). They used a 25m window in a small Spanish Pyrenees catchment with a 400m elevation range whereas another study (Lopez-Moreno et al. 2017) over a ~1400m elevation range found that a 200m TPI window produced the best correlations with snow depth, suggesting that as study area and terrain complexity increase, so should window size. Therefore, we chose a mid range for the window size of 101m based on local field observations that snowpack deposition and scour features tend to associate with terrain features in the 10m to 100m scale range, and to be compatible with the 100m sampling scale associated with elevation bins.

Aspect was divided into eight cardinal directions. "Flat" is a class produced by the aspect tool in ArcGIS, but it was not considered in this analysis as it contains only 7,191 data points whereas the next smallest aspect class (north) has 5,218,741 points. Exclusion of small or outlying classes (<0.5% of total area) will have no to negligible influence on class means and has been implemented in other snow studies (Grunewald et al. 2014).

Canopy Cover was calculated as a ratio of Lidar canopy to total returns (Barilotti et al., 2006) (Eq 4).

$$\text{Fractional Cover} = \text{Canopy Returns} / \text{All Returns} \quad (\text{Eq 4})$$

TPI and canopy cover were stratified using a three-class quantile approach, to ensure comparable class population sizes, while maximizing the chance of class separability. Slope also had three classes based on evenly distributed physical slope breaks at 30 degree increments instead of statistical sample breaks. As with aspect and elevation terrain classes, it was assumed that slope influences would be identified through physical slope attribute equal increments, rather than quantile stratification. The basis for the chosen class stratifications was to support the identification of trends in or separability of snow depth values within each driver class, and to mitigate the high level of noise in snow depth when mapped at a high resolution. This requirement for statistical comparison necessitated a categorical binning of the driver classes. The reclassified rasters were converted to class polygons and used to extract within-class snow depth sample statistics.

2.3 Lidar Snow depth Quality Control

The main sources of error with Lidar are vertical and horizontal errors as well as incorrect classification of returns from the forest canopy, buildings or other obstructions, which may be labeled as ground returns (Hodgson, 2004). The LSDM is derived from two Lidar surface models, so errors can be compounded, highlighting the need for rigorous quality control (QC). For example, it is intuitive that snow depth cannot be a negative number, so if negative values result, then this indicates either a raw laser ranging or model registration problem. Multipath (deflection of laser pulse energy from a single travel path due to multiple reflections

from high reflectance surfaces) leads to low points or pits in Lidar snow surface data that lie beneath the true surface (Hopkinson et al. 2004). This occurs because the ground to aircraft portion of the two-way travel time exceeds the time it took for the pulse to reach the ground after emission from the sensor (Hopkinson et al. 2006). Multipath is more likely to occur over a frozen snow surface than bare ground due to increased reflectivity, and this is likely to produce occasional negative snow depth values in the LSDM. Such erroneous values can be set to 0 (to limit their overall impact) or removed entirely from analysis (set to 'nodata'). For further information on Lidar as a snow depth measurement tool, and for specific examples of its use in cryospheric applications, see the review papers by Deems et al. (2013) and Bhardwaj et al. (2013).

Steep slopes and horizontal DEM uncertainty can produce erroneous depth values that exceed plausible snow depths (Hodgson et al. 2005). Horizontal coregistration errors may arise where the location of a steep cliff face is offset by a few decimeters in the x or y direction in the DSSM compared to DEM. If horizontal coregistration is imperfect, when the DEM is subtracted from the DSSM the result may be an implausibly large negative or positive value because the elevations associated with the top or bottom of the cliff face are in different cells between datasets. One approach to quality controlling slope induced errors is to eliminate steep areas from the datasets altogether (e.g. Zheng et al. 2016). At the scale of West Castle watershed, steep terrain and cliff faces ($>60^\circ$ slope) represent $<1.8\%$ of the total area. To evaluate the impact of QC measures, pre- and post-QC mean depths were compared for slope and elevation classes. It is expected that depth values in steeper terrain classes will be most affected by QC-based removal or thresholding of depth values but the overall impact on watershed-scale estimates is unknown.

Revuelto et al. (2014) employed a routine using the field-measured depth of a topographic depression as their maximum depth value while others (Kirchener et al. 2014; Grunewald et al. 2010) have employed knowledge of the area or canopy-based metrics to determine a maximum depth. However, the abundance of data in a LSDM makes removal of the upper 1% tail of the depth distribution a statistically reasonable approach to limit the upper range of snow depth. Eliminating these outlying cells entirely instead of simply thresholding implausible snow depths excludes questionable zones from the analysis (e.g. eliminating data from a snow free slope with an apparent snow depth of 60m is preferable to assigning a thresholded value to a cell where snow doesn't exist). A 99th percentile-based QC method implicitly addresses steep slope errors and provides a statistically consistent, reasonable, and automated means of setting a limit on maximum depth. The 99th percentile LSDM values were calculated after eliminating negative values. Overall, these QC measures removed ~1% of total 1m grid cells from each of the three datasets analysed.

2.4 Temporal Snow Depth Distribution Analysis

Determining how snow depth distributions and drivers vary temporally at the watershed scale supports future sampling design. For example, knowledge of snow depth driver variable spatial extent and which drivers are most important in controlling snow depth can inform driver

class sampling prioritization and location. The consistency of depth distributions was first assessed at the watershed scale using Kendall’s Tau correlation, given its non-parametric status and the non-normal distributions of the datasets. Snow depth distributions and storage estimates were derived at the watershed scale for two different stages in winter: mid-winter (MW) and melt onset (MO). Creating classes for each snow depth driver allowed a more refined test of mid-winter depth distribution consistency for which an ANOVA (analysis of variance) was employed at a 95% confidence level.

From the snow pillow records at Gardiner Creek (AEP, 2017), the February MW 2014 and 2017, as well as April MO 2016 survey times demonstrated SWE conditions, all falling in the range of 490 mm to 590 mm. During the days leading up to both mid winter surveys, there had been a period of fresh snowfall exceeding 100 mm SWE, while the melt onset mission followed a few days of early melt exceeding 10 mm SWE (Figure 2). An early melt ensued shortly after SWE peaked in 2015-16, compared to the other survey years where SWE reached its maximum in May. Leading up to the 2014 and 2017 surveys, the average temperatures of the 30 days prior were -14°C and -11°C respectively, whereas for 2016 it was -2°C. The SWE and temperature data confirm that 2014 and 2017 data can be considered to represent “mid-winter” and 2016 indicates “melt onset” conditions. The data extents and Government of Alberta weather station sites utilized in this study. Except for the March 2017 flight lines, the digital elevation model extent represents the area covered for all winter Lidar (Light detection and ranging) surveys.

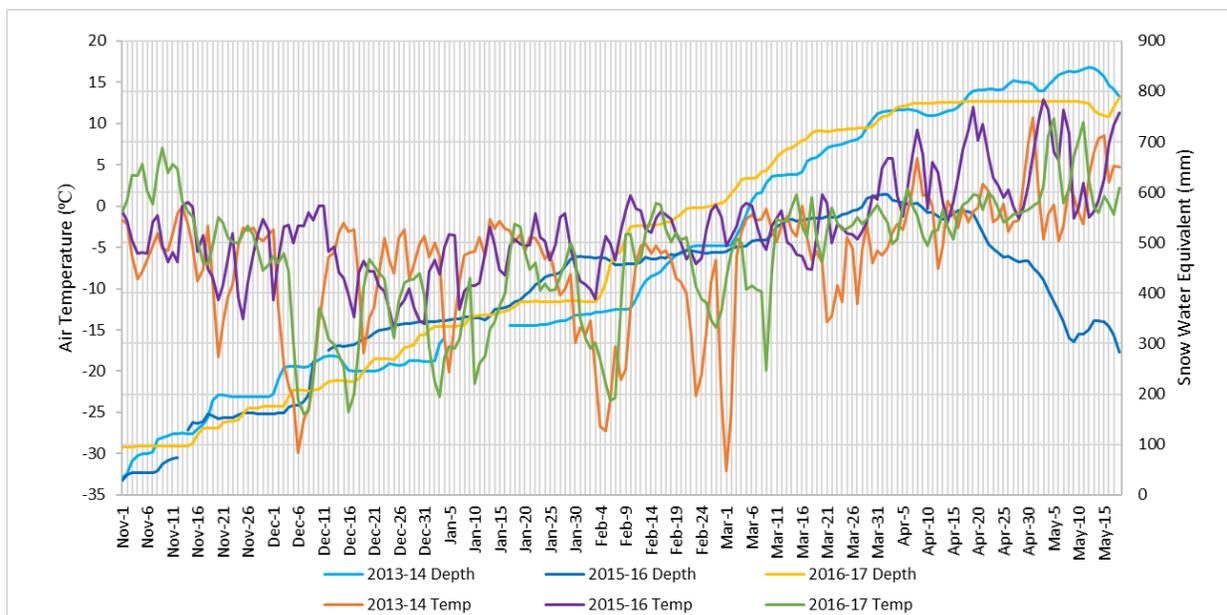


Figure 2. Daily snow water equivalent values from the Gardiner Creek site (1924 m a.s.l.), for 200 days starting on November 1st through to mid-May of the following year (AEP, 2017).

Snow depth statistics were generated for each of the snow depth driver class polygons using a python script which interrogated each LSDM and outputted a summary table. Snowpack

storage volume was estimated by multiplying mean depth of each driver class by the number of grid cells (excluding 'nodata' values) in that class. To detect statistically significant differences between depth distributions at mid-winter, ANOVAs were implemented on a random sample of each class' populations at a 95% confidence level. Sample n values of each snow depth driver class were as follows: elevation 10,000, aspect 25,000, canopy cover 30,000, TPI 30,000 and slope 30,000. This resulted in 90,000 < 200,000 samples used in each comparative test.

2.5 Government Snow Monitoring Comparison

The Government of Alberta Gardiner Creek hydrometeorological station (1970 m a.s.l.) measures snow depth and snow water equivalent, and is located just beyond the northern edge of the WCW boundary, making it the closest continuous automatic snowpack monitoring site. Depth is measured with a SR-50 and SWE data is collected with a snow pillow. The West Castle snow course (WCSC) sampling station (1524 m a.s.l.) also exists within the WCW, and is manually surveyed every month, depending on weather conditions, throughout the winter season. To give context for how the publicly available snow data relates to watershed-wide snow conditions, average snow depth from the Lidar data was compared to the SR-50 depth observed at Gardiner Creek (GC). Mean Lidar and SR-50 snow depths were both multiplied by the number of grid cells (excluding no data points) in each LSDM to estimate watershed-scale snow volume. Average depth from the LSDM was calculated by finding the mean of all cells (excluding no data points). Similarly, the depth at the West Castle snow course (WCSC) was also used to estimate volume, although the only date on which snow course measurements coincided with a Lidar survey was for the February 2014 mid-winter LSDM. For the other survey dates, the WCSC was surveyed 2+ weeks after the Lidar surveys, which precluded a viable comparison given the dynamic nature of the snowpack (Figure 2). These comparisons illustrate how a simple extrapolation of snow depth from a single regional location may lead to under- or over-estimates of watershed-scale snow volume; i.e. do these snow monitoring stations accurately represent the region they occupy?

3. Results

3.1 Snow Depth Driver Classes

Watershed land surface classifications derived from the Lidar data produced 14 elevation classes, 8 aspect classes, and 3 classes for slope, TPI and canopy cover (Figure 3).

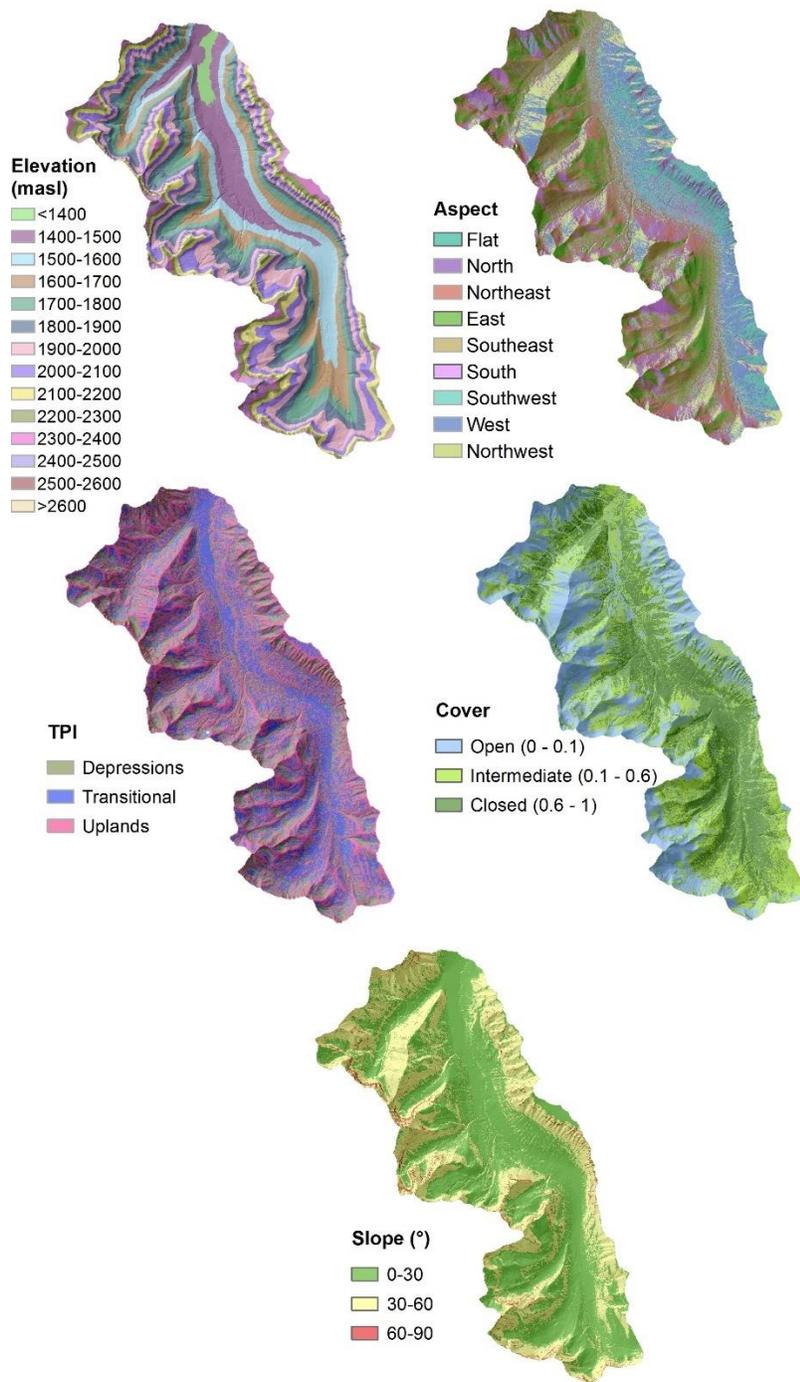


Figure 3. Snow depth driver classes.

3.2. Impact of LSDM Quality Control

Quality control reduced watershed scale depth by <0.1m (or one increment of measurement resolution) for MW 2014 and MO 2016 due to the removal of extreme depth outliers while MW 2017's mean depth was unchanged at 1.5m. Mid-winter 2017 elevation means were the least affected by QC, whereas half of the melt-onset elevation class means changed by >0.1m. The largest differences occur at upper elevations, whereas mean depths below 1800m showed minimal impact (<0.05m change in mean) across all years. As slope increased, so did the magnitude of difference between pre- and post- quality control datasets. Although depth changes were larger in steeper and higher terrain, QC only had a small effect on mean depth at the watershed scale. After QC, elevation and aspect class area distributions remained within 0.1% of the 2014 DEM areas. However, slope, TPI and canopy cover class area distributions differed by up to 0.6% relative to the DEM following QC.

Table 8 Mean depth at the watershed scale prior to and following QC.

| Dataset | Snow Depth (m) | |
|---------|----------------|----------------------|
| | No QC | 99th percentile clip |
| MW 2014 | 1.6 | 1.5 |
| MW 2017 | 1.5 | 1.5 |
| MO 2016 | 1.3 | 1.2 |

Table 2. Mean snow depth change as a result of quality control.

| Snow Depth Driver Class | | Δ Depth (m) | | |
|-------------------------|-----------|--------------------|-----------------|-----------------|
| | | Mid-winter 2014 | Mid-winter 2017 | Melt Onset 2016 |
| Elevation (m a.s.l.) | <1400 | 0.00 | 0.00 | 0.00 |
| | 1400-1500 | 0.00 | 0.00 | 0.00 |
| | 1500-1600 | 0.00 | 0.00 | 0.00 |
| | 1600-1700 | 0.00 | 0.00 | 0.00 |
| | 1700-1800 | 0.01 | 0.00 | 0.01 |
| | 1800-1900 | 0.06 | 0.01 | 0.06 |
| | 1900-2000 | 0.11 | 0.03 | 0.12 |
| | 2000-2100 | 0.09 | 0.02 | 0.12 |
| | 2100-2200 | 0.13 | 0.05 | 0.20 |
| | 2200-2300 | 0.07 | 0.03 | 0.11 |
| | 2300-2400 | 0.08 | 0.03 | 0.12 |
| | 2400-2500 | 0.09 | 0.03 | 0.17 |
| 2500-2600 | 0.13 | 0.04 | 0.23 | |

| | | | | |
|-----------|-------|------|------|------|
| | >2600 | 0.01 | 0.00 | 0.03 |
| Slope (°) | 0-15 | 0.01 | 0.00 | 0.01 |
| | 15-30 | 0.03 | 0.01 | 0.04 |
| | 30-45 | 0.08 | 0.02 | 0.10 |
| | 45-60 | 0.13 | 0.05 | 0.18 |
| | 60-75 | 0.20 | 0.17 | 0.31 |
| | 75-90 | 0.85 | 0.95 | 1.35 |

Table 3. Percent area of snow depth control classes after quality control.

| Snow Depth Driver Class | | Fall 2014 (DEM) | Mid- winter 2014 | Mid- winter 2017 | Melt Onset 2016 |
|-------------------------|-------------|--------------------|------------------------|------------------------|--------------------|
| Elevation (m a.s.l.) | <1400 | 1.3 | 1.3 | 1.3 | 1.3 |
| | 1400-1500 | 9.5 | 9.6 | 9.5 | 9.6 |
| | 1500-1600 | 11.5 | 11.6 | 11.5 | 11.6 |
| | 1600-1700 | 11.3 | 11.4 | 11.3 | 11.4 |
| | 1700-1800 | 11.7 | 11.8 | 11.7 | 11.8 |
| | 1800-1900 | 11.6 | 11.6 | 11.6 | 11.6 |
| | 1900-2000 | 12.6 | 12.5 | 12.6 | 12.5 |
| | 2000-2100 | 11.3 | 11.2 | 11.3 | 11.2 |
| | 2100-2200 | 8.7 | 8.6 | 8.7 | 8.6 |
| | 2200-2300 | 6.0 | 6.0 | 6.0 | 6.0 |
| | 2300-2400 | 3.4 | 3.4 | 3.4 | 3.3 |
| | 2400-2500 | 0.9 | 0.9 | 0.9 | 0.9 |
| | 2500-2600 | 0.2 | 0.2 | 0.2 | 0.2 |
| | >2600 | 0.0 | 0.0 | 0.0 | 0.0 |
| Aspect | Flat | 0.0 | 0.0 | 0.0 | 0.0 |
| | N | 10.0 | 9.8 | 9.9 | 9.8 |
| | NE | 12.7 | 12.5 | 12.7 | 12.5 |
| | E | 17.7 | 17.6 | 17.7 | 17.6 |
| | SE | 15.4 | 15.4 | 15.4 | 15.4 |
| | S | 7.9 | 7.9 | 7.9 | 7.9 |
| | SW | 11.5 | 11.6 | 11.6 | 11.6 |
| | W | 14.1 | 14.2 | 14.1 | 14.2 |
| | NW | 10.8 | 10.8 | 10.8 | 10.8 |
| Slope | 0-30 | 57.1 | 57.4 | 57.1 | 57.4 |
| | 30-60 | 41.2 | 40.9 | 41.1 | 40.9 |
| | 60-90 | 1.8 | 1.7 | 1.7 | 1.7 |
| TPI | Depressions | 33.1 | 32.5 | 33.0 | 32.6 |

| | | | | | |
|-------|----------------------------|------|------|------|------|
| | Transitional | 33.8 | 34.1 | 33.9 | 34.1 |
| | Uplands | 33.1 | 33.4 | 33.1 | 33.4 |
| Cover | Open (0 - 0.13) | 33.2 | 32.7 | 33.1 | 32.7 |
| | Intermediate (0.13 - 0.57) | 33.1 | 33.3 | 33.1 | 33.3 |
| | Closed (0.57 - 1) | 33.7 | 34.1 | 33.8 | 34.0 |

3.3. Snow Depth Distributions

3.3.1. Watershed Scale Snow depth Correlation

Snow depth distributions for MW 2014 and 2017 are visibly more similar to one another than MO 2016, which displays a distinct transition from low to high snow. All correlations depths around valley sides (Figure 4). This is confirmed with Kendall’s tau results indicating a higher correlation between mid-winter datasets (0.76) relative to the mid-winter and melt onset LSDM comparisons. Of the two mid-winter datasets, 2014 has a slightly higher correlation with melt onset 2016 (0.68 for MW 2014; 0.65 for MW 2017). All correlations are significant ($p < 0.01$) and demonstrate that at watershed-scale, there is a level of consistency in snow depth patterns across years and even between accumulation and ablation periods.

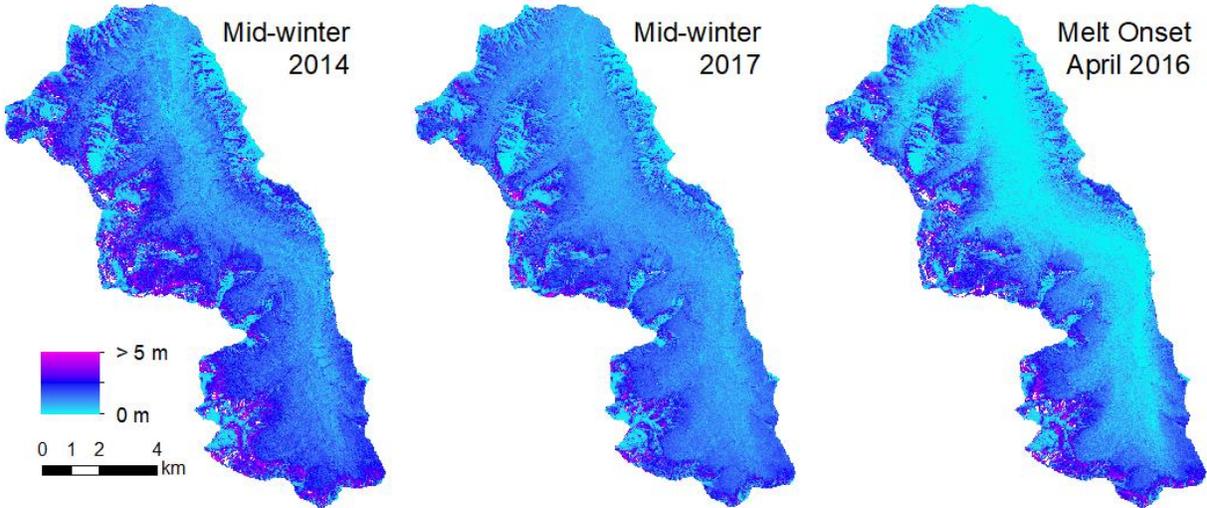


Figure 4. Lidar snow depth models within the West Castle Watershed extent.

3.3.2. Elevation

For each dataset, class-based depth distributions increased with elevation gains for the first half of the range (Figure 5). Peak depth snow depths consistently occur in the 1800-2200 m a.s.l. range. Although the exact elevation of peak depths varied by data collection, they were within the WCW treeline ecotone for all years. Above 2100 m a.s.l., mid-winter 2014 and mid-winter 2017 depths declined more sharply than melt onset 2016, illustrating continued snowpack accumulation in the upper alpine elevation ranges late in season after melt has commenced across much of the watershed. Mid-winter 2014 is the only dataset in which a consistent decline in depth occurs above the maximum depth at treeline (Figure 5). Despite the visible similarity, both of the mid-winter depth distributions are significantly different across all of the elevation bins below 2600 m a.s.l, above which the small class sizes result in no discernible difference.

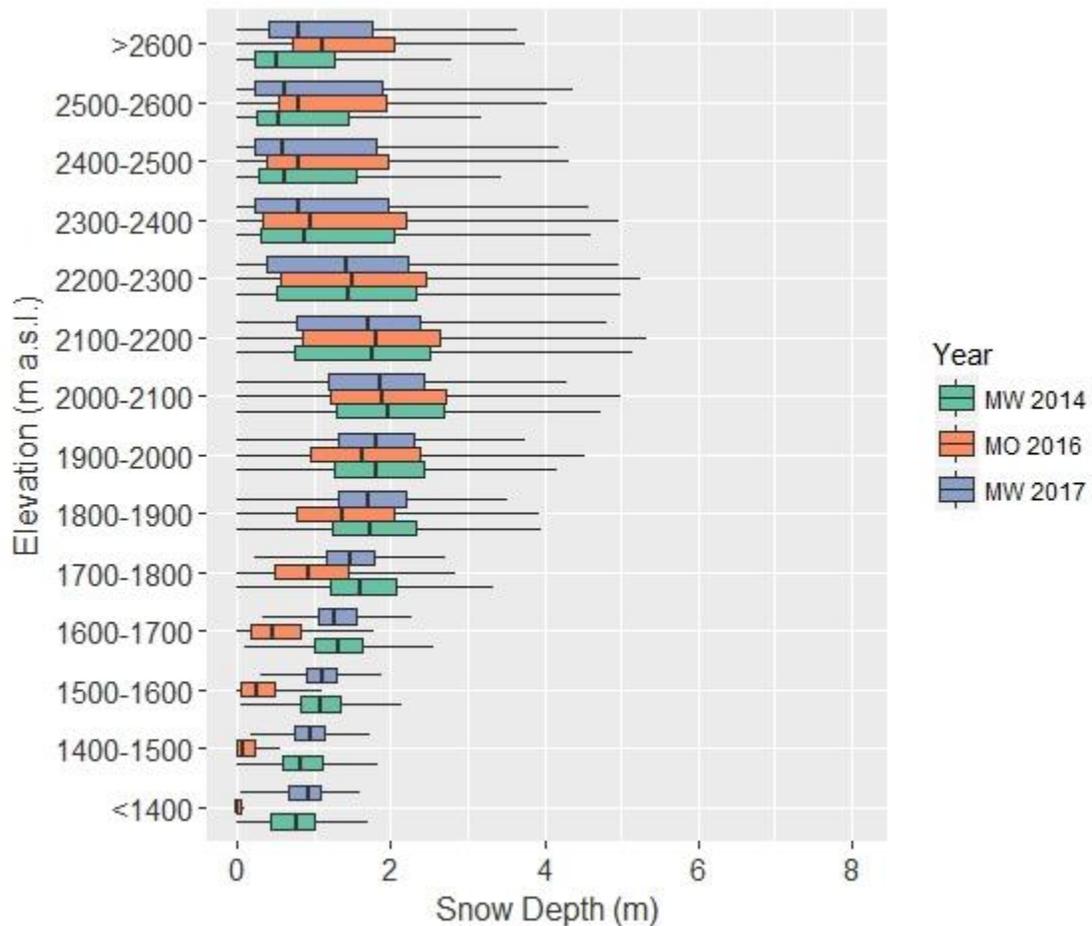


Figure 5. Boxplots of snow depth with each elevation class across the WCW. Median is shown by the bar, with the entirety of each coloured box representing the interquartile range.

Similar to depth, maximum snow storage (depth x area) occurred in the 1900-2000 m a.s.l. elevation class at mid-winter in both years (Table 4). At melt onset, maximum storage

shifted upwards to the 2000-2100 m a.s.l. class. The uppermost elevation classes (>2500 m a.s.l.) stored the least amount of snow for all years and seasons of data due to diminishing land area at alpine elevations. During mid winter and melt onset periods, the treeline ecotone (1900-2100 m a.s.l.) consistently represents the watershed zone of greatest snowpack depth and overall storage.

Table 4. Snowpack storage volume with elevation classes at the WCW scale.

| Elevation Class (m a.s.l.) | Storage ($\times 10^6 \text{ m}^3$) | | |
|----------------------------|---------------------------------------|---------|---------|
| | MW 2014 | MO 2016 | MW 2017 |
| <1400 | 1.0 | 4.5 | 1.1 |
| 1400-1500 | 8.5 | 1.5 | 9.5 |
| 1500-1600 | 13.3 | 3.8 | 13.4 |
| 1600-1700 | 15.8 | 6.7 | 15.1 |
| 1700-1800 | 20.0 | 12.4 | 18.7 |
| 1800-1900 | 22.5 | 18.5 | 21.8 |
| 1900-2000 | 24.6 | 23.1 | 24.2 |
| 2000-2100 | 22.6 | 23.2 | 22.2 |
| 2100-2200 | 15.8 | 17 | 15.9 |
| 2200-2300 | 9.4 | 10.4 | 9.3 |
| 2300-2400 | 4.3 | 5.0 | 4.2 |
| 2400-2500 | 1.0 | 1.3 | 1.2 |
| 2500-2600 | 0.3 | 0.4 | 0.3 |
| >2600 | 0 | 0 | 0 |

Aspect

West facing terrain had the lowest mean depth for all years, although the median depth value of the southwest class was slightly lower in 2016 (Figure 6). The lowest zones of storage were south and south-west zones (Table 5). The deepest mean snow depth occurred on north and northeast facing terrain (Figure 6), although the east class had the highest storage value (Table 5) due to a larger east-facing land area. Due to high sample n values, all classes had significantly different mid-winter mean depths. Visibly, however, the two mid-winter aspect-stratified depth distributions were indistinguishable from one another, while melt onset demonstrated the most marked range in class means (Figure 6).

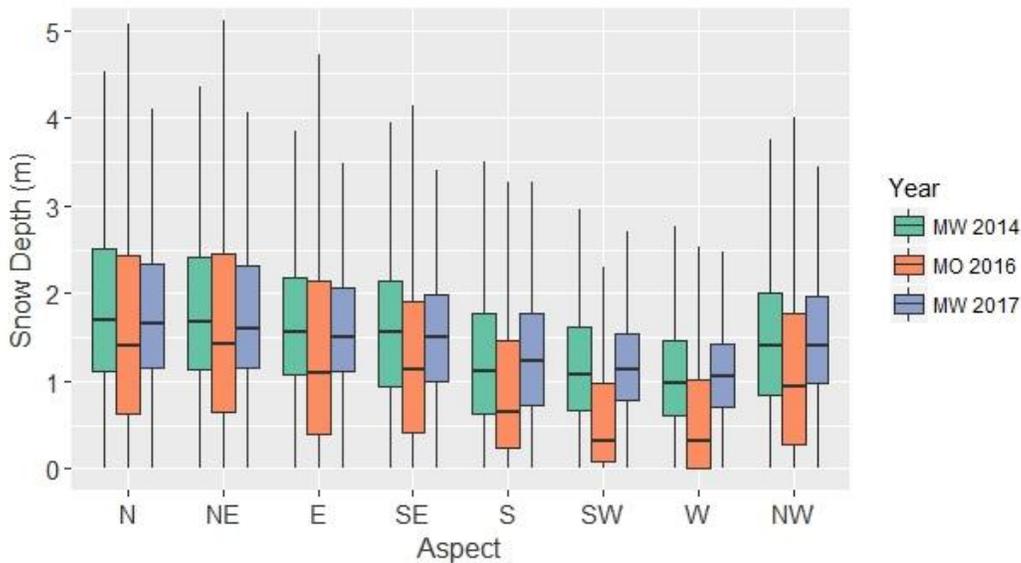


Figure 6. Boxplots of aspect class snow depth distributions. Median is shown by the bar, with the entirety of each coloured box representing the interquartile range.

Table 5. Total snowpack storage volume by aspect class across the WCW.

| Aspect Class | Storage ($\times 10^6 \text{ m}^3$) | | |
|--------------|---------------------------------------|---------|---------|
| | MW 2014 | MO 2016 | MW 2017 |
| N | 19.4 | 17.3 | 19.2 |
| NE | 24.9 | 21.6 | 24.4 |
| E | 31.6 | 25.3 | 30.8 |
| SE | 25.6 | 20.7 | 24.6 |
| S | 10.6 | 7.7 | 10.6 |
| SW | 13.9 | 7.9 | 14.1 |
| W | 15.9 | 9.3 | 16.3 |
| NW | 17.1 | 13.3 | 16.8 |

Topographic Position Index

Across all datasets, topographic depressions demonstrate the highest mean snow depths, and overall storage, with uplands displaying the lowest depths (Table 6 and Figure 7). For all three datasets, TPI class depth means were significantly different, with the two mid-winter distributions visibly almost identical

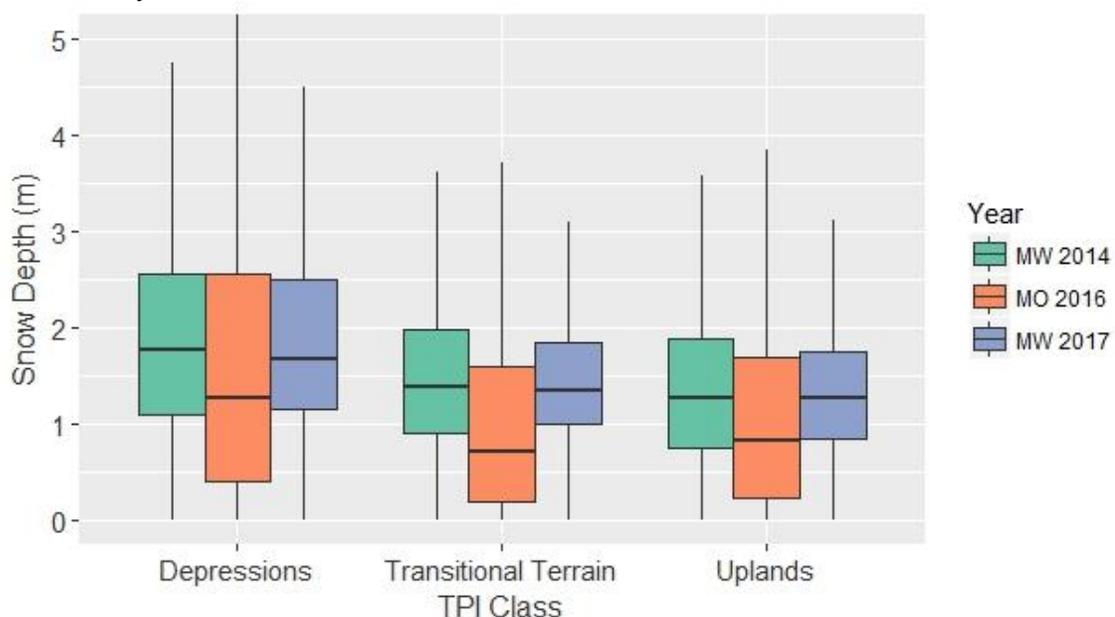


Figure 7. Topographic position index class snow depth distribution boxplots. Median is shown by the bar, with the entirety of each coloured box representing the interquartile range.

Table 6. Total snowpack storage volume by TPI class across the WCW

| TPI Class | Storage ($\times 10^6 \text{ m}^3$) | | |
|--------------|---------------------------------------|---------|---------|
| | MW 2014 | MO 2016 | MW 2017 |
| Uplands | 47.7 | 37.5 | 44.3 |
| Transitional | 52.2 | 34.8 | 48.8 |
| Depressions | 64.8 | 54.6 | 64.8 |

Canopy Cover

In mid-winter and melt onset, the highest mean snow depth occurs in the intermediate quantile (0.13 - 0.57) canopy cover class with the lowest mean depth under closed canopy (Figure 8). The lowest snow depths were experienced in closed canopy covers at melt onset, presumably because closed canopy conditions dominate valley floor elevations where early season snow melt occurs first. Due to a high variance in snow depth data, the open canopy class mean depths demonstrated no significance difference across all three datasets. However, an ANOVA yielded significantly different mean depths across the three canopy cover classes for both mid winter and melt onset conditions. During mid winter, snow storage was greatest in areas of intermediate cover (Table 7), and open areas stored more snow than intermediate and closed canopies. At melt onset, open canopy areas displayed slightly larger storage volumes than intermediate canopy covers due to the increase of snow cover and depth in alpine areas above treeline (Figure 5).

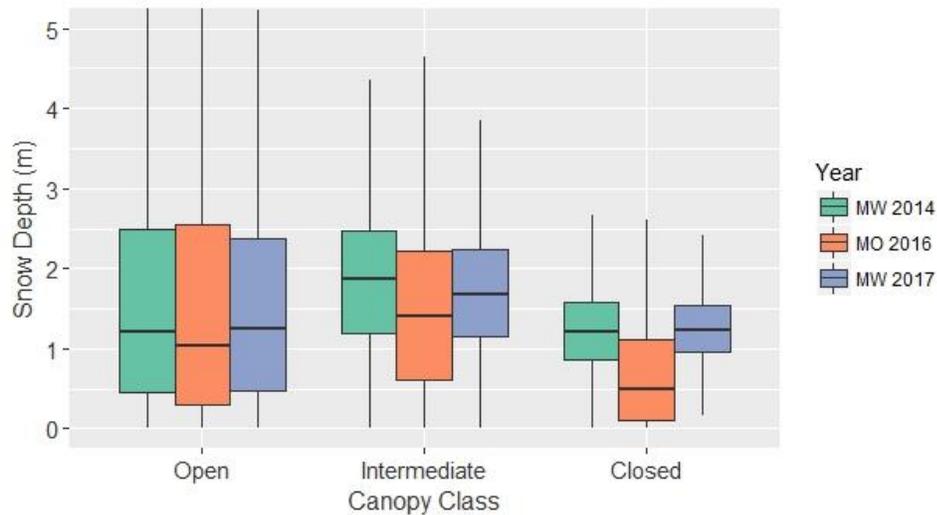


Figure 8. Canopy cover snow depth distribution boxplots. Median is shown by the bar, with the entirety of each coloured box representing the interquartile range.

Table 7. Total snowpack storage volume by canopy cover class across the WCW.

| Cover Class | Storage ($\times 10^6$ m ³) | | |
|--------------|--|---------|---------|
| | MW 2014 | MO 2016 | MW 2017 |
| Open | 52 | 53.1 | 54.1 |
| Intermediate | 62.4 | 52 | 58.7 |
| Closed | 44.2 | 24.3 | 44.9 |

Slope

Mean snow depth was lowest above 60° in 2014, with intermediate sloped terrain (30-60°) displaying the greatest depth (Figure 9). At melt onset, snow depth tends to slightly increase with terrain steepness (Figure 9). Despite this, steep slopes stored the least amount of snow at melt onset and mid-winter due to these classes representing a very small (~1.8%) portion of the watershed (Table 8). Flatter terrain stored the most snow for all time periods sampled (Table 8) though the difference from the 0-30° and 30-60° classes is less pronounced during melt onset compared to mid-winter.

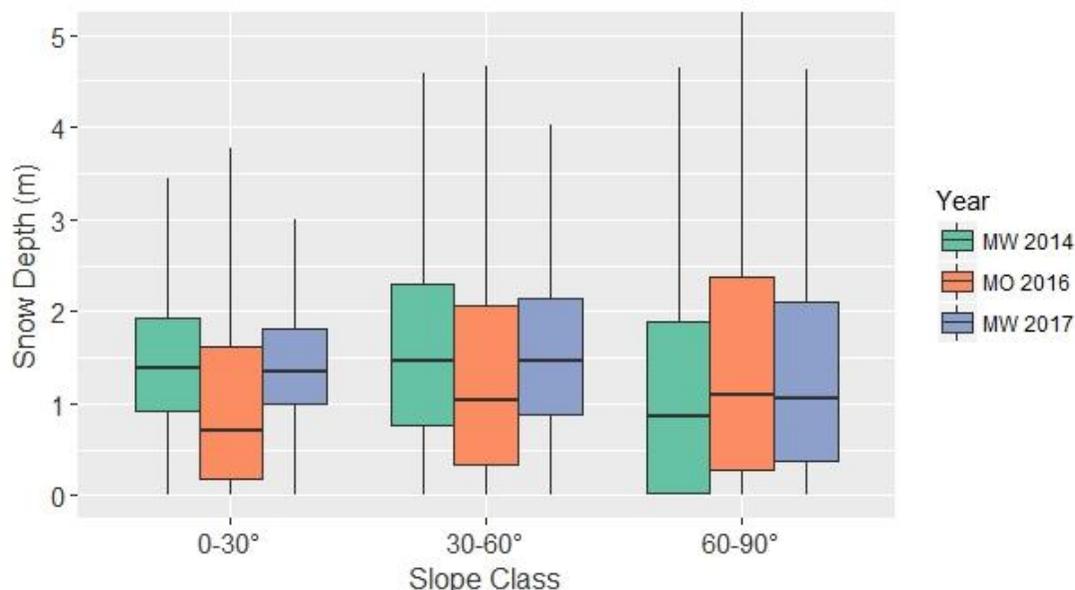


Figure 9. Slope class snow depth distribution boxplots. Median is shown by the bar, with the entirety of each coloured box representing the interquartile range.

Table 8. Total snowpack storage volume by terrain slope class across the WCW.

| Slope Class | Storage (x10 ⁶ m ³) | | |
|-------------|--|---------|---------|
| | MW 2014 | MO 2016 | MW 2017 |
| 0-30 | 87.5 | 60.9 | 86.3 |
| 30-60 | 69.3 | 59.4 | 67.9 |
| 60-90 | 2.1 | 2.6 | 2.8 |

Public and Lidar Snow Data Comparison

From all three data sources (GC, WCSC, WCW LSDM) West Castle snow course displays the lowest average depth for the single time available for comparison, while the Gardiner Creek SR50 automatic monitoring station displays snow depths that are 0.4 to 0.5 m

deeper than the overall WCW mean (Table 9). Consequently, applying these public monitoring depths to the watershed as a whole produce an overestimate of snowpack volume for GC and an under-estimate from WCSC. (Table 9). Given the WCSC sampling elevation is in the valley bottom at 1524 m a.s.l. a low depth estimate relative to the WCW LSDM is to be expected. However, the sampling elevation for Gardiner at 1970 m a.s.l. is near the middle of the elevation range for WCW (1400 to 2600 m a.s.l.) and might be expected to produce a reasonable estimate of the overall watershed-scale depth. Therefore, the consistently high snow depth relative to the watershed mean is a function of the GC station being located in the zone of highest snowpack depth near treeline (Figure 5). This observation serves as a reminder that caution is required if point-based snow depth monitoring data are extrapolated to larger spatial domains, such as when setting up hydrological models that rely on snow depth or SWE data as inputs.

Table 9. Mean depths from Lidar and provincial snow sampling data within or adjacent to WCW. West Castle snow course data for 2016 and 2017 are not included due to ~ 2 weeks of disparity between field and Lidar sampling.

| Lidar Survey Date | Mean Depth (m) | | |
|-------------------|----------------|------|-------|
| | GC | WCSC | Lidar |
| Mid-winter 2014 | 1.91 | 1.15 | 1.53 |
| Mid-winter 2017 | 1.90 | - | 1.51 |
| Melt Onset 2016 | 1.70 | - | 1.18 |

Discussion

Snow Depth Drivers and Quality Control

For all datasets, the grid cells excluded by quality control are associated with steeply sloped terrain. LSDM extreme outlier values in upper elevation range steep slope areas ranged from 62-110m across the three datasets prior to quality control. These depth values occurred either in individual grid cells or small linear clusters along cliff edges and were associated with small class sizes. At the watershed scale, mean snow depth was either unchanged or adjusted downwards by ~0.1m using the upper 99th percentile rejection approach. This suggests that co-registration errors do not affect the mean snow depth at the watershed scale but given steep sloped areas account for <1.8% of the watershed area, erroneous depth values in this class could confound accurate interpretation. Thresholding values using either the 99th percentile or 10m (an arbitrary but reasonable upper depth value) produces values within 0.1m of the pre- or post-QC data (Table 10). Snow depth values to use as an upper threshold have been determined with field-based measurements of known depressions and zones of wind loading during the season of Lidar data collection in other studies (Revuelto et al. 2014). Overall, any of these methods are sufficient if depth means are required in watershed-scale resource monitoring efforts. While this automated QC workflow does not substantively alter watershed scale depths, a more manual or

field observation-based QC method may be required if accurate depths are desired in the steep upper regions of alpine environments. For example, a detailed slope-based assessment of upper elevation snowpack on north-facing aspects might require additional QC, as these areas are generally the last to melt and are important in understanding late season runoff. Ultimately there are several approaches to Lidar snow depth QC, and exclusion of the upper 1% is expedient and has negligible impact on overall results.

Table 10. Mean snow depth sensitivity to QC method.

| Dataset | Mean WCW Snow Depth (m) | | | | |
|---------|-------------------------|---|-----------------------|---------------------------------------|---------------|
| | No QC | 99 th percentile elimination | 60° Slope elimination | Threshold 99 th percentile | Threshold 10m |
| MW 2014 | 1.6 | 1.5 | 1.6 | 1.6 | 1.6 |
| MW 2017 | 1.5 | 1.5 | 1.5 | 1.5 | 1.5 |
| MO 2016 | 1.3 | 1.2 | 1.3 | 1.2 | 1.3 |

Inter-annual Depth Distributions

Elevation and Treeline

Elevation is a well-researched explanatory variable of snow depth distributions. Peak depth for all years occurred within the elevation range of the treeline ecotone. Treeline is a transitional zone from predominately forested hill slopes to more rocky, complex open terrain areas devoid of vegetation. As snow is redistributed from exposed upper elevations, the increase in canopy cover with descending elevation (Figure 10) results in greater potential for canopy interception. Moeser et al. (2015) used Lidar data to show that the variety of canopy traits (height and cover) at the interfaces between open and forested areas promote complex and variable snow accumulation and ablation patterns relative to zones with more homogenous canopy traits. Grunewald et al. (2014), Hopkinson et al. (2012), Kirchner et al. 2014 and Zheng et al. (2016) have all demonstrated that zones of maximum snow depth occur at elevation ranges below maximum local elevations. This suggests that downward redistribution of snowpack and vegetation trapping in the subalpine treeline ecotone is a universally important snowpack distribution process and thus needs to be considered both in snowpack monitoring and in mountain watershed snowpack simulations. Moreover, it has been shown that treeline in the WCW has varied markedly over the last century (McCaffrey and Hopkinson, In Review), which suggests that snowpack distributions are also likely to vary in response to long term shifts in treeline.

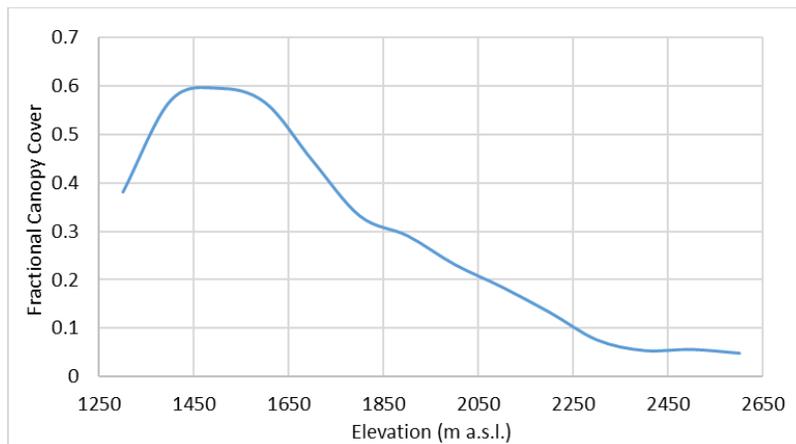


Figure 10. Mean fractional canopy cover within elevation classes in the area immediately surrounding WCW.

Declining snow depth below treeline can be explained by hydrometeorological conditions. Environmental lapse rates are the thermal gradients of decreasing air temperatures with elevation gains, which vary widely in alpine settings (Pigeon and Jiskoot, 2008). Lapse rates frequently result in temperatures $>0^{\circ}\text{C}$ at the base of mountains while below freezing conditions exist at upper elevations. Low elevation melt can also be enhanced by rain, which adds thermal energy to the snowpack (Garvelmann et al. 2015). The topographic relief created by mountain ranges promotes orographic precipitation events, where precipitation is enhanced on the leeward sides of mountains. As air masses descend the leeward side of a mountain, they can become exhausted of precipitation prior to reaching the valley bottoms, thus depositing more snow in alpine and sub-alpine zones than mid and lower elevations (Mott et al. 2015). Elevations above the zone of greatest snow depth did not produce a consistent pattern across the years (Figure 5). While linear decreases in snow depth with elevation are sometimes reported above the zone of maximum depth (Grunewald et al. 2014; Hopkinson et al. 2012), seasonal and interannual distributions influenced by a dynamic phenomenon such as wind are unlikely to be consistent at all times. Lidar scans of multiple mountainous study areas have confirmed a variety of upper elevation site-specific micro-scale effects (Grunewald et al. 2014). Peak erosion and redistribution of snowpack typically occurs during and immediately after snowfall while the surface layers are less dense than deeper settled layers (Pomeroy et al. 1997). How far unconsolidated snow is transported depends on shelter, wind speed and potential obstructions (changes in topography or vegetation) along the wind vector after snow crystals are entrained (Schirmer and Lehning, 2011).

Aspect

Aspect-delineated mean snow depths consistently show south and south-west facing terrain with the lowest mean depth, with north and northeast mean depths being the largest. Significantly high snowmelt rates in south facing areas relative to those with a north aspect have been confirmed in many studies (Haupt, 1951; Anderson et al. 1965; Hendrick et al. 1971;

Rowland and Moore, 1992; D'Eon, 2004; Jost et al. 2007). Shortwave radiation provides the energy for melt and refreezing which leads to preferential depth loss and potential density gains in aspects with high radiative inputs (south) versus those with little (north) (Golding and Swanson 1986; Anderson et al. 1958a). This effect holds true at all elevations (D'Eon, 2004). The slight east-west deviation from the expected north-south stratification in respective maximum and minimum depths suggests two additional causal mechanisms are influencing snow depth distribution in the WCW: i) the addition of afternoon sensible heat and longwave radiant inputs to the snowpack maximizes the net energy receipt, and therefore melt or sublimation, when solar radiation originates west of solar noon. The combined increased energy receipt will reduce snowpack depth on south-west facing slopes but it does not readily explain increased depths on the north-east facing slopes; ii) of the Rockies of SW Alberta are prone to westerly dominated Chinook winds. The presence of strong winds incident on west facing slopes has the potential to both elevate sublimation losses, while depositing blowing snow on the east-facing leeward slopes.

Topographic Position Index

Enhanced snow accumulation was expected to occur in topographic depressions and decline as terrain became more exposed. This occurred in each dataset, however the difference between Upland and Transitional snow depth means were small. These results suggest that the transport of snow to depressions is a reliable and predictable process, which was established by Lopez-Moreno et al. (2014) in a study of intra- and inter-annual snow persistence where topographic position index received more attention as an explanatory variable of snow depth than in previous studies.

A challenge of using TPI in snow depth and modelling studies is that depressions vary in depth throughout a given winter season. Once a depression is full of snow, a relatively smooth surface is created. After this point the terrain feature will continue to have a high depth, but new snow loading on top of the smooth surface will not be subject to the same localized shelter that earlier season snow was. With a variety of topographic depressions throughout a large study area, they will likely fill at different rates as a function of depression size, availability of snow for redistribution, and other factors that may influence or intercept the snow while being transported across the landscape. Schirmer and Lehning (2011) confirmed with variogram analyses that scale breaks increased after a snow storm as small scale terrain features were filled. While distinct stratifications in depths were apparent for each of the terrain classes (Figure 7), the systematic decrease in depth expected from depression to transition to upland classes only occurred in mid winter, with melt onset illustrating slightly decreased mean depth relative to uplands. It is likely, therefore, that TPI may be an adaptive snow depth driver variable, which is more important at the onset of snowpack accumulation in early winter but which diminishes in importance as depth increases across the landscape. Unfortunately, our data only observed mid winter and melt onset conditions where the snowpack was already established, so more work is needed to better understand the dynamic nature of TPI as a snow depth driver.

Canopy Cover

As with topographic depressions, some vegetative features eventually become buried after which topographic smoothing occurs (Schirmer and Lehning, 2011). If all or some of the short vegetation gets buried during the winter and the surface is smoother than it is in snow-off conditions, subsequent snow accumulation may no longer be influenced by short canopy in that area falling within the Open class. Tree geometry and opening patch sizes influence snow accumulation patterns (Varhola et al. 2010) by creating localized microclimates with unique wind flows and vectors. Open canopy has greater potential for enhanced redistribution and ablation by wind and radiation. Total canopy cover is lowest at upper elevations (Figure 10), but the mean snow depth for open cover classes (Figure 8) is usually higher than the mean snow depth for elevation classes above treeline (Figure 5). Much of the terrain with intermediate canopy cover has an elevation within 1900-2100 m a.s.l. (Figure 11), an approximation of the alpine treeline ecotone within the WCW (McCaffrey and Hopkinson, In Review). Previous studies of treeline and snow accumulation using Lidar (Moeser et al., 2015) found depths to be most variable at treeline, and this is confirmed visually over a sample area at mid-winter in 2014 (Figure 12). With aspect also influencing treeline elevation (McCaffrey and Hopkinson, In Review), the distribution of watershed-scale class sizes should be reflected in the sampling while attempting to also capture a representative sampling of elevations and canopy cover strata within each aspect class.

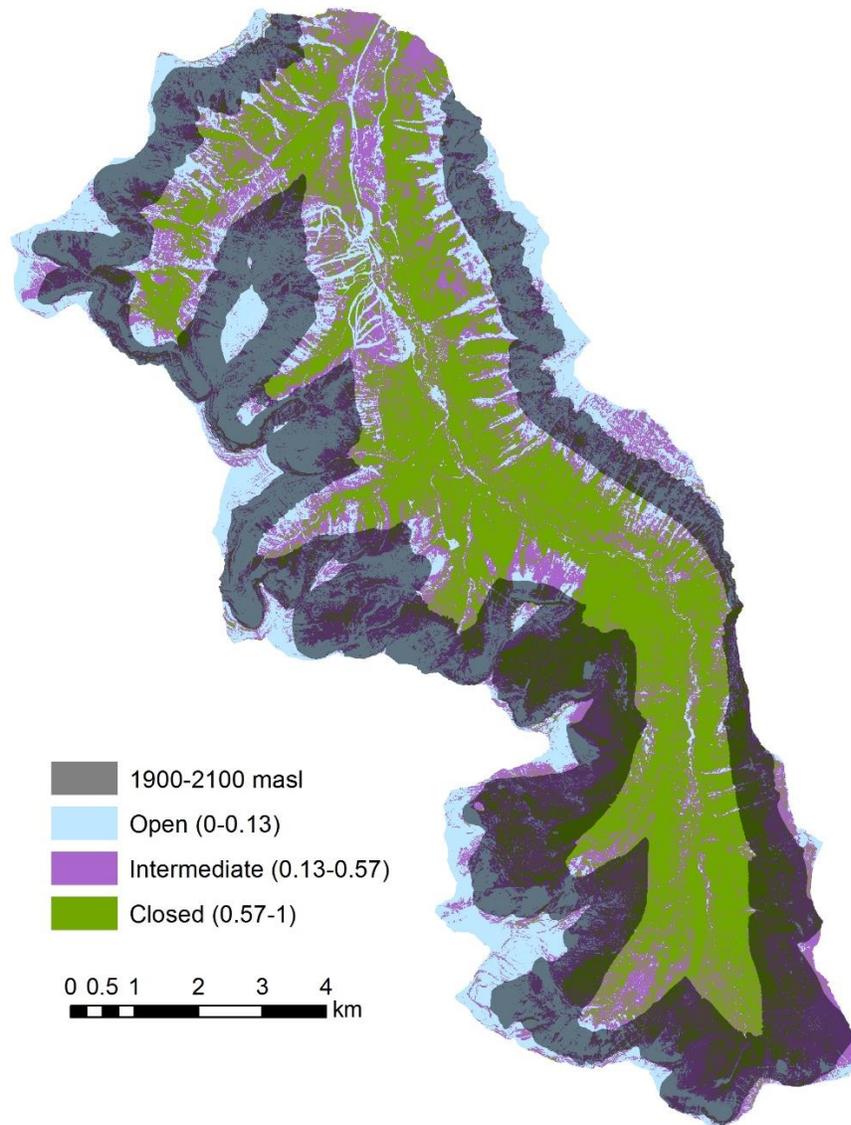


Figure 1 Approximate treeline ecotone elevation zone (1900-2100 m a.s.l.) and canopy cover classes.

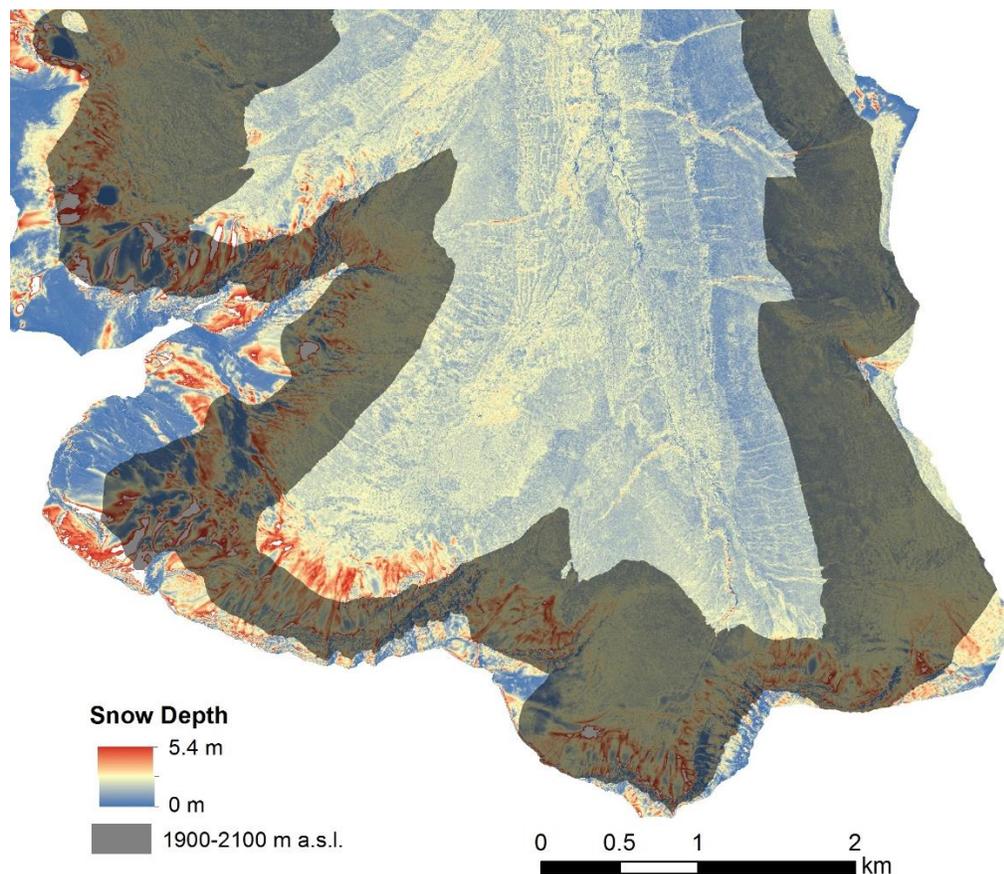


Figure 2. Mid winter 2014 snow depth variability near treeline.

Mean snow depth was lowest in the closed canopy class for all datasets. The closed canopy supports greater interception of snow, increasing the time available for sublimation losses. There is potential to lose 60-80% of intercepted snow in conifer canopies to sublimation during favourable hydrometeorological conditions (low humidity and wind) (Hedstrom and Pomeroy, 1998; Pomeroy et al. 1998). Lodgepole pine (*Pinus contorta*) conifers are the dominant tree type in the study area. If conditions are not favourable for sublimation, snow accumulation can increase below closed canopies as the branch loading capacity is exceeded and intercepted snow reaches the ground. Increased longwave radiation emitted from tree stems and branches can enhance melt, which is likely why the closed class mean is low during melt onset when temperatures were higher than either of the mid-winter surveys. Canopy cover's relation to snow accumulation and ablation suggest that this variable may be less temporally consistent than other snow depth drivers and may require further research or an adaptive sampling approach in an integrated framework.

Slope

Considering potential slope-induced errors, the higher depth means in the 60-90° class are likely a mix of genuinely large depths from snow drifts adjacent to near-vertical cliff edges, as well as outlying artefacts from over hanging cornices and/or depth amplification due to horizontal misalignment. The steep slope class represents <2% of the total WCW while low and intermediate slopes contain 57% and 41%, respectively. Many of the steepest class cells were eliminated in the quality control step. However, remaining cells in the steep class may still possess slope-induced errors, so the potential for Lidar snow depth measurements bias is highest in the steep slope class. To some unknown degree, such depth amplification will be mitigated by negative depth errors that have been set to zero in the initial QC stage. However, further research is needed to quantify the proportion of cells influenced by such positive and negative bias, as well as the overall accuracy of depth estimates and in steep sloping terrain. Given the potential for bias in the steep slope class and a general lack of depth differentiation in the intermediate and low slope classes, these empirical results suggest slope is not a consistent or useful driver of grid-cell level snow depth.

Public and Lidar Snow Data Comparison

Grunewald & Lehning 2015, and Hopkinson et al. 2012 suggested that public monitoring sites do not correspond well to Lidar based regional estimates of snowpack volume. Grunewald and Lehning (2015) state “It appears that representative cells are rather randomly distributed and cannot be identified a priori” in reference to identifying areas that are representative of static monitoring sites where hydrometeorological data is collected. While, placement of monitoring stations like Gardiner Creek within the treeline ecotone maximizes the seasonal record due to snow depths being high relative to warm valley bottoms or exposed alpine slopes, such locations are not designed to represent snow depths or SWE that are characteristic of their parent watersheds.. A network of monitoring stations would ideally capture the full variation in terrain and landcover attributes, such that they could support accurate spatialization of depths or SWE across the domains they represent. However, the size and cost of a network that would produce spatially explicit snowpack information for physically-based water resource assessments would likely be prohibitive, hence providing further justification for exploring alternative remote sensing-based monitoring.

Conclusion

New approaches to integrating in situ monitoring sites and high resolution spatial data are needed to advance the goals of physically-based and distributed water resource and flood hazard assessment. Airborne Lidar is a viable means of monitoring spatially explicit snow depth at high resolution over critical mountainous source water regions. Lidar snow depth quality control methods were tested and found to have negligible overall impact on watershed-scale mean snow depth. The upper 99th percentile depth-rejection approach advocated in this study utilizes a consistent metric that can be applied without the need for calibration, field observation or expert knowledge. If watershed scale depth is used for modelling or water supply forecasting, this quality control method has negligible impact on overall snow storage estimates due the emphasis on removing statistical outliers. However, if stratified classes are utilized for local scale estimates, volumes could be skewed at upper elevation and steeper slope sites where

uncertainties are high and estimates of upper elevation snowpack could influence late season runoff simulation.

Well-established spatial snow depth distribution trends that have been observed with manually collected and remotely sensed data are confirmed through the analyses presented here. For example, the elevation distribution of snow depths above and below treeline, primarily driven by orographic processes, lapse rates and wind redistribution, compliments the work of previous Lidar mountain watershed snowpack distribution studies (Hopkinson et al. 2012; Grunewald et al. 2014). The occurrence of peak snow depths within the treeline ecotone was observed across all years sampled during mid winter and melt onset. Reduced snow depth on south and southwest-facing terrain subject to the highest amounts of solar radiation and frequent winds, occurred consistently across all datasets. TPI results demonstrated depressions contain more snow than other terrain types but the strength of TPI as a driver of snow depth is expected to diminish throughout the winter season as depressions fill. Closed canopies consistently displayed the shallowest snowpack in mid-winter, likely due to interception sublimation losses and higher longwave radiation fluxes. Slope results were inconclusive and subject to bias in the high slope class, suggesting that slope is an unreliable explanatory variable of grid cell-level snow depth.

The observed watershed-scale correlation between the mid winter snow depth datasets (Kendall correlation of 0.76, $p < 0.01$) demonstrates that dominant spatial snow depth patterns and drivers during the accumulation phase of snowpack development tend to be consistent across years. The correlations between mid winter and melt onset (0.68 and 0.65, $P < 0.01$) are lower but also indicate that dominant drivers and patterns act throughout the winter season regardless of the accumulation or ablation trend. These are encouraging results that suggest that a Lidar-based snow depth sampling and imputation routine could produce a viable and time efficient snow depth monitoring framework. Integrating Lidar and continuous, automated data streams has the potential to enhance monitoring networks by increasing temporal and spatial availability of data in highly complex and hydrologically productive landscapes.

The spatial, seasonal and interannual consistency and variability in snow depth driver classes presented here provide the basis for: a) designing a more time- and cost-effective Lidar snowpack sampling strategy; and b) extrapolating watershed-wide snow depths from Lidar sample datasets using spatial imputation techniques. These two objectives, combined with spatially explicit watershed-scale snow density modeling, will be the next steps in building this research into an operational snowpack monitoring framework.

Acknowledgements.

Funding for the operational lidar snow depth monitoring study was provided by the Government of Alberta Environment and Parks (AEP), Alberta Innovates Energy and Environmental Solutions – Water Innovation Program, and the Natural Sciences and Engineering Council Discovery Grants program. Hopkinson also acknowledges laboratory equipment funding from the Canada Foundation for Innovation. Field logistical and data collection support was provided Clean Harbors Airborne Imaging, Castle Mountain Resort and AEP. Reed Parsons and Dr. Craig Mahoney are gratefully acknowledged for assistance with field data collection.

References

- AEP, 2017. Data obtained directly from Alberta Environment and Parks.
- Anderson, H.W., Rice, R.M. and West, A.J. (1958). Snow in forest openings and forest stands. *Proceedings of the Society of American Foresters*, 46–50.
- Anderson, H.W. and West, A.J. (1965). Snow accumulation and melt in relation to terrain in wet and dry years. *Proceedings of the 33rd Western Snow Conference*, 73–82.
- Barilotti, A., Turco, S., and Alberti, G. 2006. LAI determination in forestry ecosystems by LIDAR analysis. Workshop on 3D Remote Sensing in Forestry, BOKU Vienna.
- Breiman, L., Friedman, J., Stone, C. J., and Olshen, R.A. (1984). *Classification and Regression Trees*. Oxford, Taylor and Francis.
- Bhardwaj, A., Sam, L., Bhardwaj, A. and Javier Martin-Torres, F. (2016). Lidar remote sensing of the cryosphere, Present applications and future prospects. *Remote Sensing of the Environment*, 177, 125-143.
- DeBeer, C. M. and Pomeroy, J. W. (2010) Simulation of the snowmelt runoff contributing area in a small alpine basin. *Hydrology and Earth Systems Science*, 14, 1205-1219.
- Deems, J., Painter, T. and Finnegan, D. (2013). Lidar measurement of snow depth, a review. *Journal of Glaciology* 59(215), 467-479.
- D'Eon, R. (2004). Snow depth as a function of canopy cover and other site attributes in a forested ungulate winter range in southeast British Columbia. *BC Journal of Ecosystem Management*, 3, 1–9.
- Dickinson, W.T., and H.R. Whiteley, (1972). A sampling scheme for shallow snowpacks. *IAHS Bulletin*, 17(3): 247-258.
- Ellis, C.R., Pomeroy, J.W., Essery, R.L.H., Link, T.E. 2011. Effects of needleleaf forest cover on radiation and snowmelt dynamics in the Canadian Rocky Mountains. *Canadian Journal of Forest Research* 41 (3), 608-620
- Fang, X. and Pomeroy, J.W. (2016). Impact of antecedent conditions on simulations of a flood in a mountain headwater basin. *Hydrological Processes* 30, 2754-2772.
- Garvelmann, J., Pohl, S. and Weiler, M. (2015). Spatio-temporal controls of snowmelt and runoff generation during rain-on-snow events in a mid-latitude mountain catchment. *Hydrological Processes*, 29, 3649-3664.
- Geist, T. and Stotter, J. (2008). Documentation of glacier surface elevation change with multi-temporal airborne Lidar laser scanner data – Case Study: Hintereisferner and Kesselwandferner, Tyrol, Austria. *Z. Gletscherkd. Glazialgeol*, 41, 77-106.
- Golding, D. L. (1974). The correlation of snowpack with topography and snowmelt runoff on Marmot Creek Basin, Alberta. *Atmosphere*, 12, 31–38.
- Greene, E. M., Liston, G. E., and Pielke, R. A. (1999). Simulation of above treeline snowdrift formation using a numerical snow-transport model. *Cold Regions Science and Technology*, 30, 135-144.
- Grunewald, T., Buhler, Y. and Lehning, M. 2014. Elevation dependency of mountain snow depth. *The Cryosphere* 8, 2381-2394.
- Grunewald, T. and Lehning, M. 2015. Are flat-field snow depth measurements representative? A comparison of selected index sites with areal snow depth measurements at the small catchment scale. *Hydrological Processes*, 29(7), 1717-1728.
- Grünewald, T., Schirmer, M., Mott R. and Lehning, M. (2010). Spatial and temporal variability of snow depth and ablation rates in a small mountain catchment. *The Cryosphere*, 4, 215-225.

- Grunewald, T., Stotter, J., Pomeroy, J.W., Dadic, R., Banos, I. M., Marturia, J., and Lehning, M. (2013). Statistical modelling of the snow depth distribution in open alpine terrain. *Hydrology and Earth Systems Science*, 17(8), 3005-3021.
- Haupt, H.E. (1951). Snow accumulation and retention on ponderosa pine lands in Idaho. *Journal of Forestry*, 49(12), 869–871.
- Hedstrom, N. and Pomeroy, J. (1998). Measurements and modelling of snow interception in the boreal forest. *Hydrologic Processes*, 12, 1611–1625.
- Hendrick, R.L., Filgate, B.D., and Adams, W.M. (1971). Application of environment analysis to watershed snowmelt. *Journal of Applied Meteorology*, 10, 418–429.
- Hodgson, M. E., and Bresnahan, P. (2004). Accuracy of Airborne Lidar-Derived Elevation. *Photogrammetric Engineering and Remote Sensing*, 70(3), 331-339.
- Hodgson, M.E., Jensen, J., Raber, G., Tullis, J., Davis, B.A., Thompson, G., and Schuckman, K. (2005). An evaluation of Lidar-derived elevation and terrain slope in leaf-off conditions. *Photogrammetric Engineering and Remote Sensing*, 71(7), 817- 823.
- Hopkinson, C., Pomeroy, J., DeBeer, C., Ellis, C., and Anderson, A. (2011), Relationships between snowpack depth and primary Lidar point cloud derivatives in a mountainous environment. *Proceedings of Remote Sensing and Hydrology Symposium*, 352.
- Hopkinson, C., Sitar, M., Chasmer, L.E. and Treitz, P. (2004). Mapping snowpack depth beneath forest canopies using airborne LIDAR. *Photogrammetric Engineering and Remote Sensing*, 70, 323-330.
- Hopkinson, C., and Demuth, M.N. (2006). Using airborne Lidar to assess the influence of glacier downwasting on water resources in the Canadian Rocky Mountains. *Canadian Journal of Remote Sensing*, 32(2), 212–222.
- Hopkinson, C., Collins, T., Anderson, A., Pomeroy, J. and Spooner, I. (2012). Spatial Snow Depth Assessment Using Lidar Transect Samples and Public GIS Data Layers in the Elbow River Watershed, Alberta. *Canadian Water Resources Journal*, 37(2), 69-87.
- Jost, G., Weiler, M., Gluns, D. and Alila, Y. (2007). The influence of forest and topography on snow accumulation and melt at the watershed-scale. *Journal of Hydrology* 347:101–115.
- Kienzle, S. W., and Mueller, M. (2013). Mapping Alberta’s surface water resources for the period 1971-2000. *Canadian Geographer*, 57(4), 506–518. <https://doi.org/10.1111/j.1541-0064.2013.12050.x>
- Kirchner, P. B., Bales, R. C., Molotch, N. P., Flanagan, J. and Guo, Q. (2014). Lidar measurement of seasonal snow accumulation along an elevation gradient in the southern Sierra Nevada, California. *Hydrology and Earth System Sciences* 18(10), 4161-4275.
- Lehning, M., Grünewald, T. and Schirmer, (2011). M. Mountain snow distribution governed by an altitudinal gradient and terrain roughness. *Geophysical Research Letters*, 38, 5 pp.
- Marchand, W. D. and Killingtonveit, A. (2005). Statistical probability distribution of snow depth at the model sub-grid cell spatial scale. *Hydrological Processes*, 19, 355–369.
- McCaffrey, D. R. and Hopkinson, C. Terrain and Fire-Exposure Controls on Century-Scale Canopy Cover Change in the Alpine Treeline Ecotone. *Remote Sensing*. Submitted.
- Moeser, D., Stahli, M., and Jonas, T. (2015). Improved snow interception modelling using canopy parameters derived from airborne Lidar data. *Water Resources Research*, 51, 5041-5059.

- Mott, R. Scipion, D., Schneebeli, M., Dawes, N., Berne, A. and Lehning, M. (2015). Orographic effects on snow deposition patterns in mountainous terrain. *Journal of Geophysical Research, Atmospheres*, 119, 1419-1439.
- Murray, C., and Buttle, J. (2003). Impacts of clearcut harvesting on snow accumulation and melt in a northern hardwood forest. *Journal of Hydrology*, 3, 197–212.
- Pigeon, K. E., and Jiskoot, H. (2008). Meteorological Controls on Snowpack Formation and Dynamics in the Rocky Mountains. *Arctic, Antarctic and Alpine Research*, 40(4), 716-730.
- Pomeroy, J.W., Parviainen, J., Hedstrom, N., and Gray, D.M. (1998). Coupled modelling of forest snow interception and sublimation. *Hydrologic Processes*, 12, 2317–2337.
- Pomeroy, J.W., Marsh, P., and Gray, D.M. (1997). Application of a distributed blowing snow model to the Arctic. *Hydrological Processes*, 11(98), 1451-1464.
- Lopez-Moreno, J., Revuelto, J., Alonso-Gonzalez, E., Sanmiguel-Vallelado, A., Fassnacht, S. R., Deems, J., and Moran-Tejeda, E. (2017) Using very long-range terrestrial laser scanner to analyze the temporal consistency of the snowpack distribution in a high mountain environment. *Journal of Mountain Science*, 14(5), 823-842.
- Revuelto, J., Lopez-Moreno, J. I., Azorin-Molina, C., and Vicente-Serrano, S. M. (2014). Topographic control of snowpack distribution in a small catchment in the central Spanish Pyrenees, Intra- and inter- annual persistence. *Cryosphere Discussions*, 8(2), 1937–1972.
- Roe, G.H., and Baker, M. B. (2006) Microphysical and Geometrical Controls on the Pattern of Orographic Precipitation. *Journal of the Atmospheric Sciences*, 63, 861-880.
- Rood, S. B., Samuelson, G. M., Weber, J. K., & Wywrot, K. A. (2005). Twentieth-century decline in streamflows from the hydrographic apex of North America. *Journal of Hydrology*, 306(1-4), 215-233.
- Rowland, J.D., and Moore, R.D. (1992). Modelling solar irradiance on sloping surfaces under leafless deciduous forests. *Agricultural and Forest Meteorology*, 60, 111–132.
- Samuelson, G. M., & Rood, S. B. (2004). Differing influences of natural and artificial disturbances on riparian cottonwoods from prairie to mountain ecoregions in Alberta, Canada. *Journal of Biogeography*, 31(3), 435-450.
- Schirmer, M., and Lehning, M. (2011). Persistence in intra-annual snow depth distribution, Fractal analysis of snow depth development. *Water Resources Research*, 47, doi, 10.1029/2010WR009426.
- Steppuhn, H.W. (1976). Areal water equivalents for prairie snowcovers by centralized sampling. *Proceedings of the Western Snow Conference*, 44, 63-68.
- Varhola., A., Coops, N.C., Weiler, M., and Moore, R.D. (2010). Forest canopy effects on snow accumulation and ablation, An integrative review of empirical results. *Journal of Hydrometeorology* 392, 219-233.
- Weiss, A. 2001. Topographic Position and Landforms Analysis. Retrieved from: http://www.jennessent.com/downloads/TPI-poster-TNC_18x22.pdf
- Zheng, Z., Kirchner, P.B., and Bales, R. C. (2016). Topographic and vegetation effects on snow accumulation in the southern Sierra Nevada, a statistical summary from Lidar data. *The Cryosphere*, 10, 257-269.

Appendix 5: Spatial and Temporal Consistency of machine learning-based imputation of snowpack depth within an operational lidar sampling framework

Cartwright, K., Mahoney, C. Hopkinson, C. SPATIAL AND TEMPORAL CONSISTENCY OF MACHINE LEARNING-BASED IMPUTATION OF SNOWPACK DEPTH WITHIN AN OPERATIONAL LIDAR SAMPLING FRAMEWORK. *Hydrological Processes*. (In Review)

SPATIAL AND TEMPORAL CONSISTENCY OF MACHINE LEARNING-BASED IMPUTATION OF SNOWPACK DEPTH WITHIN AN OPERATIONAL LIDAR SAMPLING FRAMEWORK

Cartwright, K., Mahoney, C., Hopkinson, C.

Abstract

Airborne Lidar can support high resolution watershed-scale snow depth mapping that provide the spatial coverage necessary to support water supply forecasts for mountainous headwaters. This research utilized Lidar and machine learning to evaluate snow depth drivers and assess the feasibility of sampling datasets for the spatial imputation of snow depth to the watershed scale under mid and late winter conditions, as well as at the onset of melt. Snow depth raster models in continuous integer format proved to be the fastest type of input data to implement a random forest depth simulation. Models of watershed-scale snow depth developed from spatially constrained flight-line training samples correlated with more spatially widespread Lidar snow depth test data but were outperformed by models generated from training data sampled across the entire watershed. All random forest simulations produced r^2 values ranging from 0.41 to 0.61 and RMSE values from 0.7m to 1.0 m ($p < 0.05$). By evaluating the performance of snow depth drivers within each of the random forest simulations, it was found that aspect and topographic position index were consistently ranked as important drivers regardless of seasonality. While Lidar sampling and machine learning imputation is shown to be a viable framework for mountain snowpack monitoring over extended source water regions, work is needed to optimise flight path sampling configurations.

Introduction

Snowpack Monitoring

Mountainous terrain is highly complex and the seasonal snow that accumulates in these regions exhibits large spatial variability as a result. This variability can limit the accuracy of spatial snow depth models (Erxleben et al., 2002; Pomeroy et al., 1997). Models can be valuable prediction tools based on a small amount of input data to fill in spatial data gaps over a broad area. Understanding water storage by seasonal snow is especially important in regions where major rivers start in the mountains (Byrne et al. 2006). Snow water equivalent (SWE) is the variable desired for modelling water resource availability and it is the product of snow depth and density. Density can be sampled in the field and has the potential to be modelled from hydrometeorological data (Jonas et al. 2009). The localized range in snow depth can exceed four times that of density (Dickinson and Whiteley, 1972; Steppuhn 1976), thus SWE values are more sensitive to snow-depth variations than density. It is therefore important to prioritize a time- and cost-efficient means of acquiring snow-depth data over density.

In-situ data collection is challenging in mountains due to logistical access and safety constraints. In Alberta, sonic rangefinders measure continuous seasonal snow depth, while SWE is measured by snow pillows. These technologies are automated and typically record values at time increments of 15, 30 or 60 minutes. Both depth and SWE values are also collected manually once every 4-5 weeks, at snow course sites, which are often co-located or at lower elevations than the automated stations. Snow data are currently measured at approximately 12345 sites across the mountain headwaters of Alberta, equating to a single measurement per 54321 km², which limits the production of reliable water supply forecasts. Should climate warming and headwater snowpack trends continue to change, the current snowpack monitoring network in Alberta may not be sufficient to inform future water supply predictions accurately.

Lidar and Snowpack Measurement

The use of airborne Light Detection And Ranging (Lidar) data could improve the current availability of headwaters snow-depth data. Lidar is often used to collect baseline terrain data and can be an effective environmental monitoring data source that is gaining popularity as a snow depth measurement tool (Grunewald et al. 2010, Banos et al. 2011, Lehning et al. 2011, Hopkinson et al. 2012, Grunewald et al. 2013, Bhardwaj 2016, Zheng et al. 2016, Lopez-Moreno et al. 2017). Considering the high variability of snow depth distributions and the infeasibility of collecting spatially explicit and landcover-representative data manually, coupling Lidar sourced snow depth models with density data will provide more spatially complete and accurate headwater SWE estimates than is currently possible using traditional methods.

Lidar data can provide dense point clouds, which effectively sample the Earth's surface, from which high spatial resolution (≤ 1 m) snow depth can be derived with decimeter accuracy (Hopkinson et al. 2004). As Lidar datasets are generally acquired over large-scales (relative to field campaigns) they are capable of providing tens or hundreds of millions of depth measurements from a LSDM (Lidar Snow Depth Model), significantly greater than the hundreds or thousands of points from an intensive field sampling campaign (Elder et al. 1998).

Lidar's ability to make observations quickly and safely are advantages when spatial coverage and measurement resolution are considered, as in-situ measurements can introduce spatial bias through safety and ease of access constraints. However, Lidar data acquisitions over large areas require favorable sky conditions for the mission duration, with costs increasing for longer and larger survey missions. Although Lidar data have demonstrated reliability for terrestrial surface mapping, field measurements remain invaluable for calibrating and validating observed products. Deems et al. (2013) and Bhardwaj et al. (2013) both provide further explanation of Lidar as a cryospheric measurement tool and the inherent challenges associated with the use of this technology in mountainous zones.

Typically Lidar data are acquired wall-to-wall over the entirety of a study site, however, recent studies note the potential to use a Lidar sampling (i.e. not covering the whole study area) approach to model the complete extent of a study area. This concept was first addressed in Alberta by Hopkinson et al. (2012) who adopted a snow accumulation unit approach to modelling, where terrain and land cover class means were determined with Lidar transect sample

datasets and extended over larger, unsampled areas based on the physical attributes of the terrain. Such an approach, in conjunction with sophisticated machine learning models, has the potential to reduce the costs of Lidar surveys whilst providing wall-to-wall snow-depth products with minimal reduction in overall accuracy.

Snow Depth Modelling

Since the inception of snow depth and SWE modelling, primary methods used have been multiple linear regression (MLR; Golding, 1974; Elder et al., 1991) and binary regression trees (BRT; Elder et al. 1998; Balk & Elder, 2000). These techniques have a history of variable success in the prediction of snow information (

Table 9) and can determine the order of snow depth predictor variable importance (as measured by the contribution of the variable to overall model accuracy) and apply those statistical relationships in a predictive modelling context. This has been achieved primarily through using different combinations of variables, power functions of variables and/or interactions with variables to see which produces the best model, rather than determining a ranked order of predictor variables ability to explain snow depth distribution.

Early multiple linear regression (MLR) studies typically focused on SWE, instead of depth. The first documented regression model for predicting snow accumulation was developed in 1960 (Ku'zmin, 1960) using canopy cover in a Russian boreal forest study area. Subsequent studies utilized MLR with multiple ancillary datasets as predictors (elevation, topographic position index, aspect, slope and forest density) to produce a model that explained 48% of the variability in SWE (Golding, 1974). A similar approach was applied by Lopez-Moreno and Nogues-Bravo (2006) who used elevation, elevation range, radiation and two location parameters to explain 50% of SWE variability.

Binary regression trees (BRTs), a type of decision tree, are another common statistical method applied in snowpack modelling. Further details on BRTs and their use in snow depth and water equivalent modelling are described by Molotch et al. 2005, Breiman et al. 1984, Elder et al. 1995 and 1998, Balk and Elder 2000. Individual BRT's typically yield better results for distributed snow depth and water equivalence estimates than other methods, including MLR (Molotch et al. 2005). An advantage of decision trees over MLR is their interpretability, as well as their capacity to handle more predictor variable distributions (e.g. skewed, normal, etc.) and formats (e.g. continuous or categorical) than MLR, for which a primary assumption of the technique is that distributions are continuous and normal (Kuhn and Johnson, 2013).

These studies represent a cross-section of MLR and BRT applications to snow pack modelling, using different variable combinations, with varying success (e.g. Hosang and Dettwiler, 1991; Elder et al. 1991. Chang and Li, 2000; Erxelben et al 2002; Winstral et al. 2002; Molotch et al. 2005; Jost et al. 2007). Across a subset of the body of literature, MLR and BRT approaches have explained 18-90% of variation in predicted snow depth or SWE products (Table 1). Lidar (terrestrial or airborne) featured in only three of the reviewed studies, all of which were published in 2010 or later, suggesting the adoption of Lidar for such work is still in its infancy. One Lidar study employed MLR techniques in tandem with Airborne Lidar snow depth data acquired from around the world (Grunewald et al. 2013). Lidar snow data were favoured over interpolated raster surfaces from relatively small field data sets for snow depth modelling. With LSDMs from around the world, the authors compared snow depth drivers in a variety of mountainous zones across the globe. They concluded that universal, globally applicable snow depth models do not exist and that predictor variables interact in such a complex manner that predictor importance is often site-specific, corroborating the research of others (Molotch et al., 2005).

The complex nature of snow, terrain and microclimatology is perhaps one of the greatest challenges to establishing replicable snowpack modelling routines, despite impressive technological advances in data collection, processing and analysis over recent decades.

Considering the logistical challenges associated with collecting robust datasets manually, optimized Lidar sampling has the potential to support distributed watershed- or regional-scale snow models of sufficient detail and accuracy for meaningful water supply predictions.

Table 9. Summary of snow modelling studies.

| Author and Year | Data | Method | Variables | Results (r ²) |
|-------------------------------------|---------------|--------|---|-----------------------------------|
| Elder et al., 1998 | Field | BRT | solar radiation, slope angle, elevation | 0.6-0.7 (snow depth) |
| Balk and Elder, 2000 | Field | BRT | net solar radiation, elevation, slope, vegetation cover type | 0.54-0.65; 0.60-0.85 (snow depth) |
| Erxleben et al., 2002 | Field | BRT | elevation, slope, aspect, net solar radiation, vegetation | 0.18-0.30 (snow depth) |
| Winstral et al., 2002 | Field | BRT | upwind index, drift delineator, elevation, solar radiation, slope | <0.4; 0.50-0.60 (snow depth) |
| Marchand and Killingtveit, 2005 | Field, Geodar | MLR | elevation, aspect, curvature, slope | 0.5-0.48 (snow depth) |
| Lopez-Moreno and Nogues-Bravo, 2006 | Field | BRT | elevation, elevation range, solar radiation, slope, location (to seas/oceans, elevational divide) | 0.15-0.70 (snow depth) |
| Lopez-Moreno and Nogues-Bravo, 2006 | Field | MLR | elevation, elevation range, radiation, slope, location (to seas/oceans, elevational divide) | 0.51-0.58 (snow depth) |
| Lehning et al., 2011 | ALS | MLR | elevation, fractal roughness | >0.70 (snow depth) |
| Grunewald et al., 2013 | ALS | MLR | elevation, slope, northing, wind, surface roughness | 0.27-0.90 (snow depth) |
| Golding, 1974 | Field | MLR | elevation, TPI, aspect, slope, forest density | 0.48 (SWE) |
| Elder et al., 1991 | Field | MLR | elevation, slope, radiation | 0.27-0.40 (SWE) |
| Molotch et al., 2005 | Field | BRT | elevation, slope, aspect, northness | 0.28-0.41 (SWE) |
| Plattner et al., 2006 | Field | MLR | elevation, curvature, distance from ridge, shelter | 0.41 (SWE) |
| Jost et al., 2007 | Field | MLR | elevation, aspect, forest cover, solar radiation, temperature | 0.83-0.88 (SWE) |
| Litaor et al., 2008 | Field | BRT | upwind index, elevation, slope, aspect, slope-aspect topoclimatic index, solar radiation, plant biomass, species richness | 0.85-0.91 (snow depth and SWE) |
| Grunewald et al., 2010 | TLS | MLR | elevation, slope, aspect, radiation/elevation, slope, max SWE, wind speed | 0.30-0.40 (daily ablation rates) |

Random Forest Models

In mountainous environments, there are many driving mechanisms which interact at a variety of spatiotemporal scales to produce an observed snow depth distribution. Relationships between snow depth and independent variables are often non-linear, which complicates traditional predictive modelling approaches (Anderton, 2000; Nogues-Bravo, 2003). Random Forest (RF) is a non-parametric algorithm that creates ensembles of decision trees. Predictive variable order importance of snow depth controls in the Spanish Pyrenees has been determined with RF

(Lopez-Moreno et al. 2017) and it has also been used to quantify snow volume and predict snow depth errors (Tinkham et al. 2014). Random Forest has recently been used to map snow cover extent with freely available synthetic aperture radar data (Tsai et al. 2019), at a 100 m grid cell resolution. Identification of hydrometeorological sensor sites for integration with Lidar snow depth data in three 1 km² areas has also been achieved recently (Oroza et al. 2016). Random Forest has and is currently utilized in other disciplines, in the pursuit of time and cost reductions, such as wetland mapping (Amani et al. 2019; Chignell et al. 2018; Wang et al. 2015) as well as crop yield monitoring (Jeong et al. 2016). The combination of RF and snow depth predictions remains relatively unexplored compared to the traditional methods of MLR and BRT. A novelty of this research is that we utilize extensive datasets (103 km²) at a high resolution (3m), compared to previous work where smaller study areas and/or lower resolutions were employed, for modelling snow depth rather than snow covered area. This paper also provides recommendations for optimal input data format to guide future modelling efforts.

Objectives

The current study aims to build a foundation for the operational production of large-scale snow depth models in a mountainous headwater environment through the application of innovative machine learning methods and airborne Lidar data. Moreover, this study will compare modelled snow-depth results from multiple years to demonstrate the approaches' efficacy and consistency. These are achieved through addressing the following objectives:

- 1) Use of RF to model snow depth and assess prediction accuracy under mid-winter and melt onset conditions
- 2) Assess the consistency of inter-annual and seasonal predictor variable importance; i.e. do watershed-scale snowpack depth controls vary through time?
- 3) Determine the feasibility of using Lidar snow depth sample datasets for high-resolution, spatial imputation
- 4) Optimise the RF snow depth modeling approach by exploring trade offs between computation time and output quality as a result of training data sample sizes, input raster type, raster resolution and input data quality control methods.

Study Site and Data

Study Site

Add general description of the site, and some characteristics.

The West Castle Watershed (WCW) study area is in southwestern Alberta, Canada, along the continental divide (Figure 1). Situated in Southern Alberta near the Canada – US border, the 103 km² WCW is a sub-watershed of the Oldman River Basin from which ~8.4% of annual runoff yield is derived when combined with the neighbouring South Castle watershed (Kienzle and Mueller, 2013). With the West and South Castle combined to create the Castle River, this system represents the 3rd highest yielding sub-basin in the ~26,000 km² Oldman watershed.

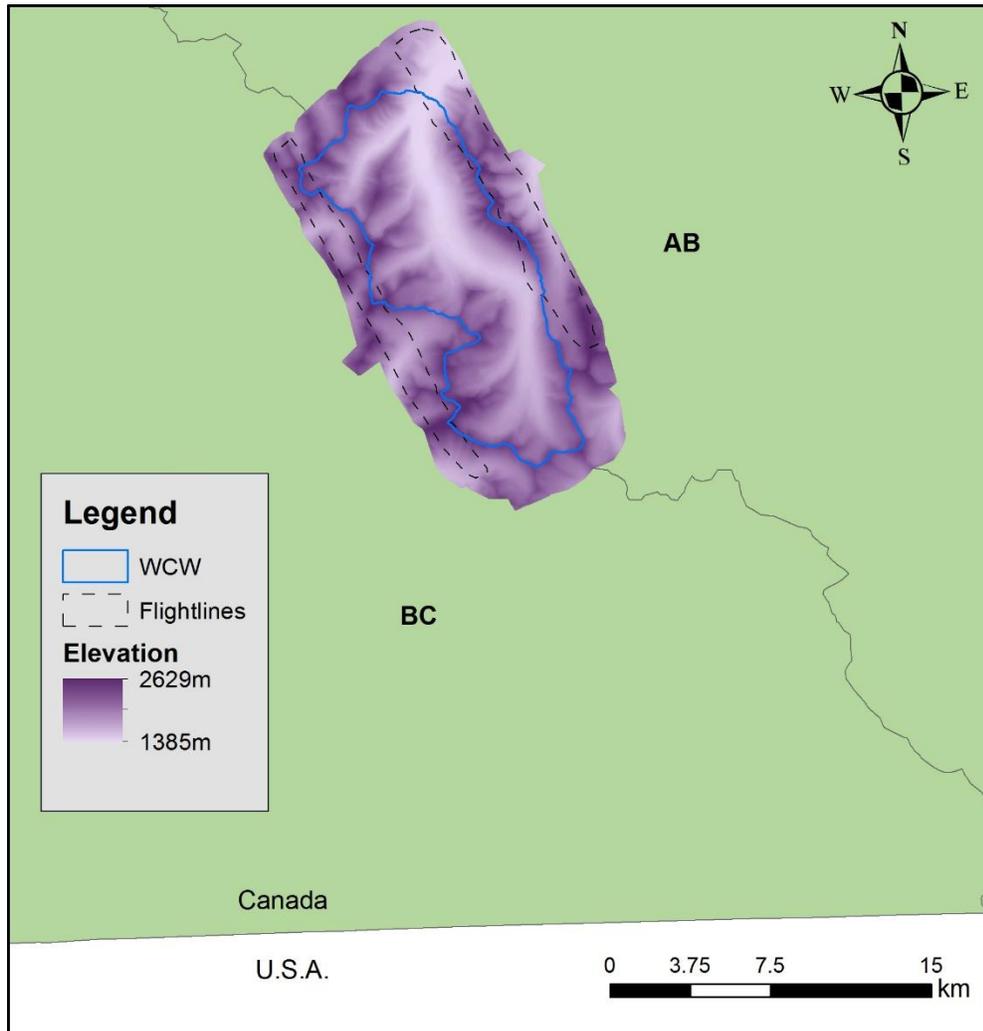


Figure 3. The WCW is southwest of Calgary, Alberta, with the west and south edges of the watershed following the continental divide between British Columbia and Alberta. All surveys covered the same extent except for the March 2017 flight lines.

Lidar

Multiple Lidar datasets were acquired over the West Castle Watershed (WCW) using ALTM 3100EA (2014) and Riegl Q1560 (2016 and 2017) sensors flown aboard a Piper Navajo twin engine survey aircraft. All missions were planned at a median altitude of 1000 m a.g.l. with 40 degree total field of view, 50% mean swath side lap, and sample point density >2 pts/m². A snow-off acquisition was conducted October 18, 2014 survey when perennial snow was at its minimum. To produce snow-on digital snow surface models (DSSMs), further airborne Lidar surveys were conducted February 24, 2014 (mid-winter), April 7, 2016 (melt onset), February 13, 2017 (mid-winter) and March 20, 2017 (late winter).

Methods

Lidar Snow-Depth Models

A 1m raster snow-off digital elevation model (DEM) was generated using a triangulated irregular network (TIN) procedure in Surfer (Golden Software, Colorado, USA) from the October 2014 survey following data cleaning and ground classification using Terrascan (Terrasolid, Finland). The same cleaning, classification and raster procedures were followed to produce the snow-on DSSMs. The DEM data were then subtracted from the DSSM data to generate a 1m raster Lidar snow depth model (LSDM) for each of the winter acquisitions (Eq 1). Further discussion of the processing and validation of these lidar snow depth datasets is provided in Cartwright et al. (In Review) and Cartwright (2019).

$$\text{LSDM} = \text{DSSM} - \text{DEM} \quad (\text{Eq 1})$$

The extent of the LSDMs from which snow depth training data were taken is shown in Figure 1. Three LSDMs covering 103 km² and a fourth dataset covering the extent of two flightlines (Figure 1) exist for the WCW. Two of the datasets were collected in February, corresponding to mid-winter (MW) conditions, and the third dataset with complete watershed coverage was surveyed at melt onset (MO). The sample ‘flight-lines’ dataset (FL) was the result of a planned full survey in March 2017 being interrupted by a rapid weather change, resulting in single flight line swaths over the east and west boundaries of the WCW in late winter (LW). ‘WCW’ here refers to data covering the entire watershed, whereas ‘FL’ refers to sample data from within the flight-line extent (Figure 1). Each 1m LSDM was processed as per Hopkinson et al. (2012) and the canopy cover raster was calculated as a ratio of canopy returns to all returns (Barilotti et al., 2006) (Eq 2).

$$\text{Fractional Cover} = \text{Canopy Returns} / \text{All Returns} \quad (\text{Eq 2})$$

Lidar snow survey root mean square errors (RMSE) of a decimeter have been reported (Hopkinson et al. 2004; Grunewald et al. 2010; Lehning et al. 2011), with accuracy generally highest in areas of open canopy but RMSE values ranging up to 0.3m have been published (Geist

and Stotter, 2008; Moreno Banos et al. 2009; Hopkinson et al. 2011; Grunewald et al. 2013). Field validation for mid-winter 2014 at elevations <1450 m a.s.l. produced a RMSE of 0.27 m while a slightly lower RMSE value of 0.25 m was observed in 2016 for depth validation across the 1700 to 2100 m a.s.l. elevation range. Combined with the well-documented accuracies of airborne Lidar for snow depth, we are confident these data represent typical contemporary lidar snow depth model quality.

Given the tendency for sample distribution tails to represent outliers, all LSDM data beyond the 99th percentile value were removed to mitigate against outlier influence in subsequent machine learning modeling steps (see Cartwright et al. In Review, for a more thorough description). This approach was adopted, as the objective is to develop a snow depth imputation routine that is generally applicable to the watershed as a whole rather than accurately characterising the behaviour of statistically outlying behaviour. Many outliers occur on steep slopes or along cliff bands where the topography causes problems due to an extreme incidence angle during data collection (Hodgson et al. 2005; Zheng et al. 2016). With the upper 1% of data eliminated, we are more confident in data quality, as we avoid assigning erroneous depth values to places where there may not be any snow, as could occur if thresholding by depth.

DEM, LSDM and Canopy Cover raster data layers were all generated at 1 m and then bilinearly resampled (post-quality control) to yield 3 m products. Executing modelling steps at the watershed-scale on 1 m resolution data was computationally intensive and unnecessary given the of developing a practical workflow for operational use in snowpack mapping and monitoring. With higher resolution data, the noise level is inherently increased and a larger training data sample is required to adequately represent study site characteristics. A 3 m resolution dataset mitigates noise and may better integrate with current hydrological models, given resampling from high to low resolution is more feasible than low to high. The use of 1 m datasets for snow modelling was not considered computationally efficient for watershed- or regional-scale (>100 km²) operational estimates of snowpack volumes.

Remote Sensing Derivatives

Resampling rasters derived from neighbourhood functions such as aspect, slope and Topographic position index (TPI) can change z-values in the process and can potentially create surfaces that misrepresent true terrain attributes (Kienzle, 2004). For this reason, aspect, slope and TPI were calculated after generating the resampled 3m DEM. Slope was calculated in degrees using Eq 3. TPI is a measure of terrain concavity (depression), convexity (upland) or transitional terrain. It compares each DEM grid cell to the mean of surrounding cells in a user-defined window, where the mean of within-window grid cells (M) is subtracted from the central grid cell elevation (E) (Eq 4). Small windows can create unnecessary detail within the TPI layer yet large windows simplify the relief (Lopez-Moreno et al. 2017). For this analysis, a TPI window sizes of 100m was used (See Cartwright et al. In Review for an in depth discussion of the TPI window choice).

$$\text{Slope } (^{\circ}) = \arctan(\text{rise} / \text{run}) \quad (\text{Eq. 3})$$

Random Forest Workflow

Random Forest (RF) is a non-parametric decision tree ensemble method that interrogates reference and corresponding predictor data units to maximise information gain (Breiman, 2001). RF can be employed in classification or regression modes, the former predicting labels to be assigned to any given target, where the latter predicts numerical outputs to be assigned. RF employs a sophisticated data splitting process that gives it machine learning status over MLR or BRT algorithms. BRT and RF methods are similar in that they both rely on decision trees with binary 'nodes' connected by 'branches'. However, whilst a BRT forms only a single tree, RF 'grows' multiple trees. RF models require training, employing a dataset that exhibits both reference and predictor units, a trained RF model is the basis of inferring predictions from other predictor units not employed in model training. Regression-based RF methods, employed here, impute estimation units for a given target based on stored reference units related to predictor units in a trained RF model.

To train a RF model, first, a parent dataset (containing reference and predictor units) is randomly sampled (with replacement, defined as bootstrapping) at a ratio of 2:1 to define training and testing datasets, respectively. The training data are used to grow a single decision tree through iteratively splitting predictor units based on information gained about corresponding reference units, known as node splitting. This is achieved by imposing a binary threshold across a single of multiple 'mtry' randomly sampled predictor units where the predictor that yields the greatest information gain is selected. Unused predictors at a given node do not influence results and increase computational efficiency. After splitting, each node produces two branches (left and right), both of which link to independent nodes where the splitting process is repeated. Node splitting is repeated until a predefined information gain threshold is not exceeded, producing a termination node (or leaf) that is split no further. Once all leaf nodes are finalised the decision tree is complete. The decision tree growing process is repeated 'ntree' times, where each tree is grown from a different bootstrap sample, to form a trained RF model (Cutler, 2007). The predictor units from each testing data bootstrap sample are passed through each decision tree to assess model prediction accuracies and predictor variable importance.

To obtain predictions from a RF model, predictor units from targets where no reference information is available are submitted to the trained model. These predictor units follow a pathway through each decision tree based on the splitting criterion at each node to arrive at leaf nodes. The mean of each estimation unit (one from each decision tree leaf node) provides the final predictive output for the given target. This is repeated for each target to yield a final prediction dataset.

A benefit of RF over BRTs and MLR is its ensemble approach which offers more stable models and variable importance assessments (Kuhn and Johnson, 2013). RF differs from other statistical modelling techniques in that it does not rely on statistical significance (as in MLR) to determine predictor variable importance (i.e. the contribution of each predictor to model overall accuracy). Importance in RF was determined using the R statistical software program using the 'randomForest' package (Liaw and Wiener, 2002). Variable importance was assessed based on mean squared error (MSE) outputs, which represent a comparison of how the MSE of the

original bootstrap sample changed when variables were permuted. When a variable is important to a model, permuting values for that variable over the dataset will negatively affect predictions. This metric is given as “% Increase MSE”, derived by calculating the MSE for the whole model before and after predictor permutation, then finding the percent change between the MSE values (sometimes greater than 100%) (Breiman, 2002). A larger ‘% increase MSE’ demonstrates how much the model accuracy has degraded due to the removal of a certain predictor, the greater the MSE, the more important the variable.

Five predictor variables were used in this study (elevation, aspect, slope, topographic position and fractional canopy cover), all of which exhibit importance in modelling snow accumulation in the Albertan Rockies (Golding, 1974b; Chang and Li, 2000; Lopez-Moreno and Nogues-Bravo, 2006; Marchand and Killingtveit, 2005; Grunewald et al., 2013). All RF models were trained, and predictions made using the ‘randomForest’ package for the R statistical programming software (Liaw & Wiener, 2018).

Model Optimization

The two main internal parameters ‘mtry’ and ‘ntree’ can be tuned to optimize RF models during model training. Additionally, the format of input predictor data (as rasters) has been noted to influence model training and prediction times, and accuracies (Mahoney et al. 2016). Raster format influences the amount of storage space required (Bonham-Carter, 1994), making it important to assess how using integer-based rasters as modelling inputs, which occupy less memory than floating point data, could affect processing time and results.

Before random forest models were applied to each dataset and extent (WCW and FL), the impact of training data format and sample size, and forest size (ntree) was assessed as a function of processing time and model statistics.. Summary statistics R^2 and RMSE (root mean square error) were calculated from the comparison of training data reference units and predicted snow depths (estimation units) from the same X, Y locations. These metrics were used to assess model performance, and provided an opportunity for more thorough assessment in the workflow optimization. A model was deemed optimal when highest R^2 values and lowest RMSE values were produced. As RMSE is an absolute measure of correspondence between two variables it was considered more important than R^2 in determining optimal models. In the event of multiple models yielding identical RMSE values, the model with the greater R^2 was deemed optimal. These trials were carried out with the Melt Onset 2016 data as later season snowpacks are more homogenous than earlier distributions, reducing variability in the input data (Lopez-Moreno, 2017). Table 2 summarises the model optimization trails conducted.

Table 2. Description of exploratory optimization trials conducted for snow depth modelling using random forest.

| Trial | Description |
|---------------------------|---|
| Ntree | The number of trees grown during RF model training can affect model accuracy and training time. Optimized model will yield high accuracies in a short amount of time. |
| Data format | All floating point variables were converted to integer by multiplying by 10 and removing decimals. Integer data require less memory than floating point equivalents and theoretically take less time to process during model training. Integer data were further split in to discrete classes, where each class represented a range of continuous data. This potentially improves computational time, but may adversely affect model accuracies. Continuous integer vs categorical class input data formats were evaluated. |
| Training data sample size | Smaller training data samples can be processed more quickly and efficiently. However, fewer samples yield models with reduced accuracies. |

For ntree analyses, values of 100, 250, 500 and 1000 were used to train an RF model based on (Green boxes in Figure 2), a parent dataset consisting of 50,000 samples from the 2016 WCW extent LSDM, in continuous integer format. The LSDM (reference unit) was multiplied by 10 to produce whole number snow depths and rounded to the nearest decimetre. A similar approach was applied to all predictor units with the exception of TPI, where the maximum negative value was added to all TPI cell values before the multiplication and rounding steps. This removed negative values and produced a common data format for all driver variables entered into subsequent R script implementation. A further trial was performed with discretely classified categorical data inputs (described in Cartwright et al. In Review), to assess whether the reduction in memory requirement (Bonham-Carter, 1994) would reduce processing time.

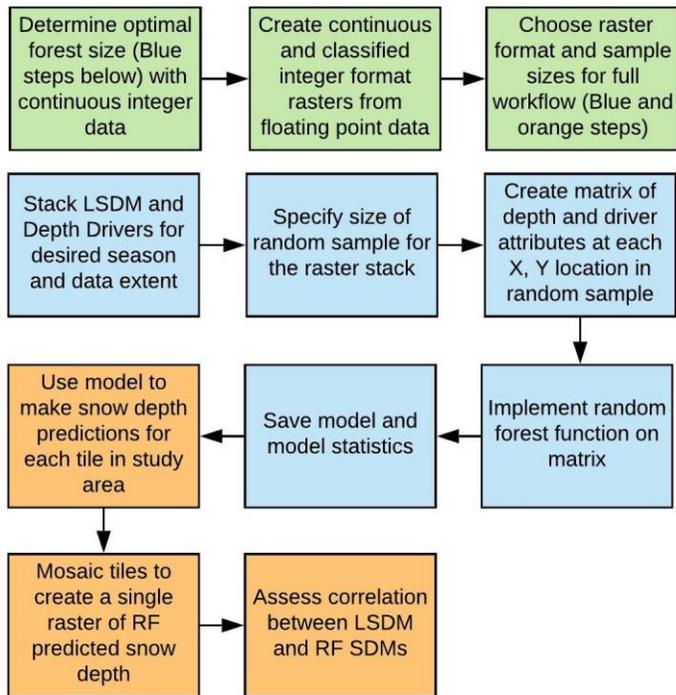


Figure 2. Random forest modelling workflow.

To determine an appropriate sample size of training points with respect to processing time and correspondence between training and predicted data, trials using 50,000 (0.44% of 3m grid cells in the WCW) and 100,000 (0.88% of 3m grid cells in the WCW) points from the WCW were conducted. Once a suitable sample size for the WCW was determined, the same proportion of WCW training points were used to identify the sample size of training points for the flight lines to ensure consistency. To sample the training data, rasters were stacked by X, Y coordinates and a random sample of points was taken across the entire data spatial extent. This step identifies a ‘seed’ value to ensure the same random X, Y locations are sampled for each data extent, and does not specify that sampling contains a certain amount of snow depth driver attributes (e.g. X % of sample must contain north facing terrain), to ensure a true random selection. The randomly sampled data points determine the range of values for each snow depth driver (Figure 3) that are used as inputs for creating the final random forest models. Predictions were made over a ~20% subset area of the watershed for exploratory optimization (number of trees, raster format and sample size) trials.

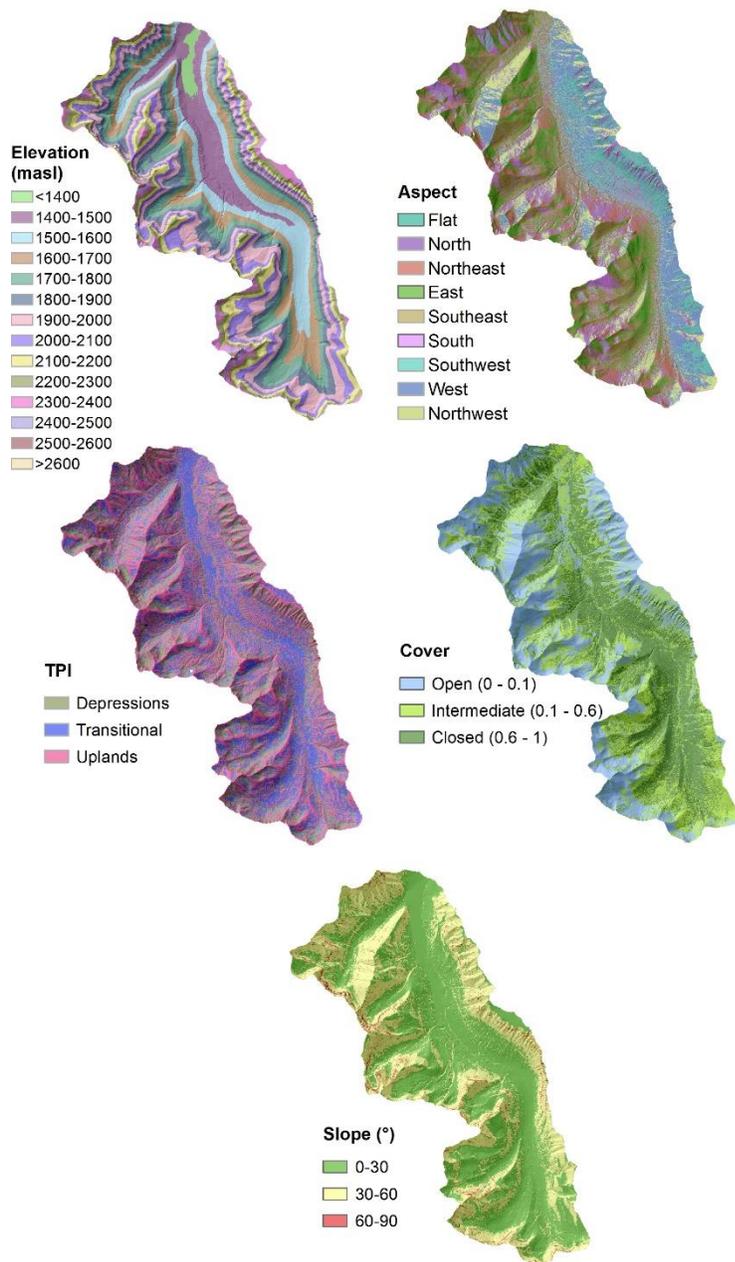


Figure 3. Snow depth driver classes for the WCW extent. Class stratifications illustrated here are categorical for illustrative purposes. Model simulations were performed using common continuous integer driver data formats.

Once optimal models for all trial were determined, they were used for predictive purposes. Predictions were made for tiled subsets of the full WCW extent, which were subsequently mosaicked to create a single, seamless predicted snow depth raster for the WCW, from which correlation and summary statistics (mean, maximum depth etc.) were determined. Model performance was assessed by three statistics at the watershed scale: R^2 , RMSE and Pearson Correlation Coefficient (PCC). R^2 and RMSE are calculated from training data and corresponding predictions (RF bootstrap samples), and PCC values are derived from independent samples sourced from the comparison of Lidar derived LSDMs and RF predicted SDMs (snow depth models) whilst maintaining the same sample size. Data points for LW 2017 PCC were limited to the zones of overlap between the WCW and LW 2017 flight-lines, whereas other PCC data were randomly sampled across the whole WCW as both the original Lidar and predicted layers exhibit the same extent.

Results

Ntree Optimization

The influence of ntree on model accuracy, as described by R^2 and RMSE, and computation time is assessed from multiple RF models, each trained with a different value of ntree. The R^2 and RMSE of each RF model (derived between training and corresponding predicted points) are identical regardless of the number of trees used in model training (Table 3). Thus for the current dataset, a RF model trained using 100 decision trees yields adequate accuracy with minimal time cost. As ntree increased, processing time for model training also increased. RMSE calculated between snow depth comparisons and is expressed in meters. As the precision of the derived Lidar LSDM products is accurate and rounded to ± 0.1 m, RMSE values were also rounded accordingly.

Table 3. Forest size (ntree) RF model optimization as a function of model accuracy and computational time expenditure for a subset of the watershed.

| ntree | R^2 | RMSE | Time (min) |
|-------|-------|------|------------|
| 100 | 0.61 | 0.5 | 83 |
| 250 | 0.61 | 0.5 | 159 |
| 500 | 0.61 | 0.5 | 202 |
| 1000 | 0.61 | 0.5 | 275 |

Data Format & Sample Size Optimization

50,000 and 100,000 randomly sampled points from continuous floating point and integer, and discretely classified categorical training datasets were used to train RF models and assessed for model parsimony. In general, smaller training datasets produce models more quickly with similar

measures of accuracy (Table 5). Similarly, models trained using integer data are completed faster than with floating point data, with broadly similar model accuracies. Although categorical data exhibits a faster run time than continuous integer data, model R^2 was reduced. The chosen model used a dataset of 50,000 points in continuous integer format, yielding a compromise between computational run time and most accurate model statistics (Table 4). Minimal variation in RMSE and R^2 was observed for trials with continuous data. While a slight improvement in model performance was realised with 100,000 training data points, this doubling of sample size resulted in a 4.3 fold increase in run time (>10 hrs) for a small subset of the watershed.

Table 4. Outputs and timing as a result of input data type and point count for a subset of the watershed.

| Raster Format | n | R^2 | RMSE | Time (min) |
|--------------------|---------|-------|------|------------|
| Floating Point | 50,000 | 0.61 | 0.5 | 140 |
| Continuous Integer | 50,000 | 0.61 | 0.5 | 79 |
| Categorical | 50,000 | 0.56 | 0.5 | 18 |
| Floating point | 100,000 | 0.63 | 0.5 | 606 |
| Continuous Integer | 100,000 | 0.62 | 0.6 | 602 |

WCW and FL Data Distributions

For the simulations at watershed-scale, the random sampling of the two data extents (WCW and FL) resulted in relatively similar proportions of each snow depth driver class in respective training datasets compared to the proportion of grid cells each class occupies at the WCW scale (Table 5). Elevation, TPI and slope classes are well represented in the WCW training data. Notable deviations from the watershed proportions are present in the WCW for the intermediate and closed canopy cover classes, where intermediate cover was over-sampled by 7.7% and closed canopies were under-sampled by 8.3%. The WCW sampled data contains fewer grid cells that face northeast (4.8%) and a greater amount of northwest facing terrain (4.9%) compared to the entire watershed. The FL sampled data exhibits a higher proportion of elevation classes above 1700 m a.s.l. (above sea level) as well as below 1400 m a.s.l. Open canopies were oversampled by the FL sample compared to the other cover classes, likely due to the location of the FL data, near the watershed boundaries, which exhibit higher elevations (near and above treeline) and are therefore less likely to exhibit dense canopies. Slope and aspect classes training data proportions corresponded best with watershed class percent areas. TPI was poorly represented by the FL data with 98.1% of the training data coming from uplands, again an artifact of the high elevations over which the FL data were acquired near the watershed boundaries.

Table 5. Relative sample coverage for each WCW snow depth driver class in the training data. Note, RF model applied at watershed scale using continuous integer driver data. Categorical bins presented below for comparative purposes.

| Driver | Class | % of WCW area | % of WCW Training Data | % of FL Training Data |
|-------------------------|--------------|---------------|------------------------|-----------------------|
| Elevation (m a.s.l.) | <1400 | 1.4 | 1.4 | 3.0 |
| | 1400-1500 | 9.5 | 9.6 | 6.0 |
| | 1500-1600 | 11.5 | 11.8 | 4.8 |
| | 1600-1700 | 11.3 | 11.2 | 9.3 |
| | 1700-1800 | 11.7 | 11.7 | 12.1 |
| | 1800-1900 | 11.6 | 11.7 | 12.6 |
| | 1900-2000 | 12.6 | 12.5 | 12.4 |
| | 2000-2100 | 11.3 | 11.1 | 12.2 |
| | 2100-2200 | 8.7 | 8.6 | 10.6 |
| | 2200-2300 | 6.0 | 6.0 | 8.9 |
| | 2300-2400 | 3.3 | 3.2 | 5.5 |
| | 2400-2500 | 0.9 | 0.9 | 2.1 |
| | 2500-2600 | 0.2 | 0.2 | 0.4 |
| Cover | Open | 33.2 | 33.9 | 46.4 |
| | Intermediate | 33.2 | 40.9 | 36.9 |
| | Closed | 33.6 | 25.3 | 16.7 |
| TPI | Depressions | 25.3 | 25.5 | 1.5 |
| | Transitional | 38.7 | 38.3 | 0.4 |
| | Uplands | 36.0 | 36.2 | 98.1 |
| Aspect | N | 5.1 | 5.4 | 5.3 |
| | NE | 17.6 | 12.4 | 12.8 |
| | E | 17.7 | 17.7 | 14.3 |
| | SE | 15.4 | 15.3 | 12.9 |
| | S | 7.9 | 8.0 | 12.2 |
| | SW | 11.5 | 11.9 | 12.4 |
| | W | 14.1 | 13.5 | 14.7 |
| | NW | 10.8 | 15.7 | 15.5 |
| Slope | 0-30 | 58.5 | 58.6 | 53.0 |
| | 30-60 | 39.8 | 39.7 | 44.7 |
| | 60-90 | 1.7 | 1.7 | 2.3 |

Inter and Intra-annual Predictor Importance

Predictor importance was obtained for each RF model using temporally distinct reference (Lidar) data and compared as a function of time to ascertain if certain predictors are consistently more important than others. Under mid-winter conditions, aspect was the most important predictor for both 2014 and 2017 (Figure 4a). TPI was among the top three variables, listed 3rd most important for 2014 and 2nd for 2017. Elevation had the 2nd highest importance for 2014, but was less important than 3rd ranked slope for 2017. Elevation and aspect's importance under melt onset conditions were much greater than other predictors.

For the flight line-based predictions, where data from all years were restricted to the spatial extents of the flight line dataset acquired during 2017, aspect and TPI were consistently among the top three drivers for each dataset (Figure 4b). Slope displayed higher importance than cover or elevation under mid-winter conditions whereas late winter and melt onset importance did not have slope in the top three. Elevation was only in the top three drivers for late winter and melt onset data when training data was extracted from the flight-lines. It is important to consider that for the FL based trials, the proportion of training points to grid cells in the LSDM was kept constant to the proportion utilized in WCW based trials. The use of fewer training points for FL trials as well as having different extent of data to sample from (Table 5; Figure 1) may influence importance results.

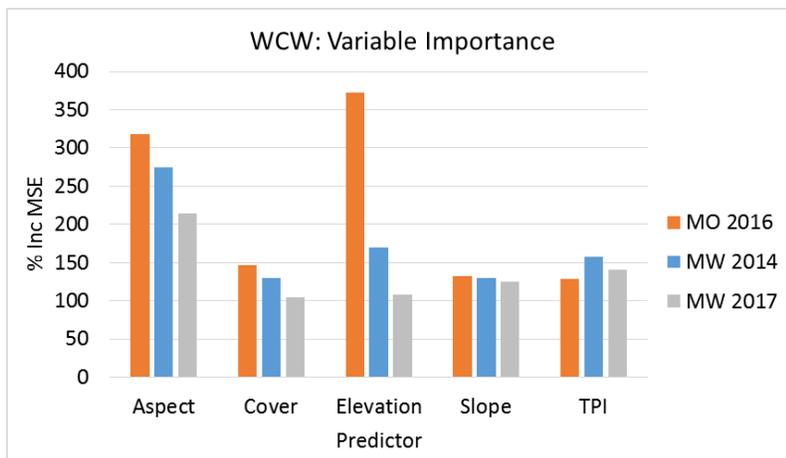


Figure 4a MSE values for input data from the watershed (WCW). Note: higher 'Inc MSE' represents better predictive capability (see text).

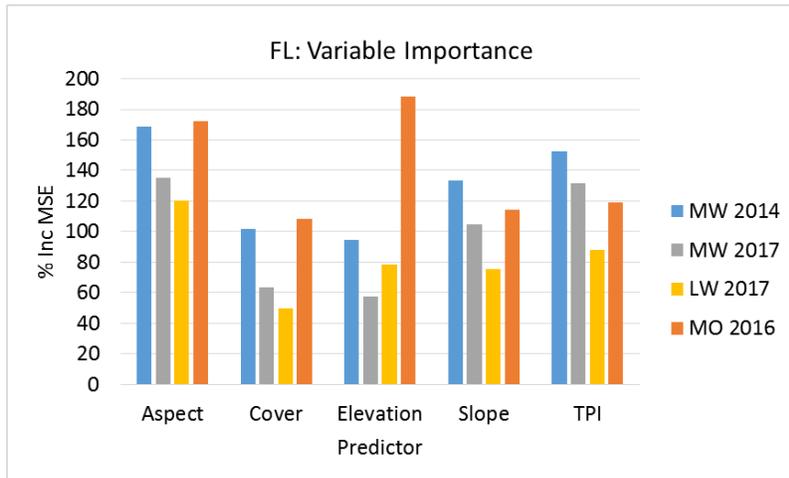


Figure 4b MSE values for input data from the flight line (FL) sample extent.

Random Forest Snow Depth Model Assessment

Snow depth model (SDM) R^2 (Figure 5) was consistently higher for WCW based training data compared to the flight-line data (Table 6). The best R^2 (0.61) values occur under melt onset conditions. RMSE values indicate the same absolute error (0.7m) across all WCW data extents, with similar values for MW and MO conditions to FL sample-based models. The R^2 for LW 2017 is low (0.41) and RMSE is high (1.0m).

Models based on FL sample data run almost 3.5 times faster than those based on full watershed training data (Table 6). However, as expected, model R^2 and RMSE statistics indicate that full watershed models yield superior results. In general, melt onset models from WCW training data yield more accurate models than mid-winter, however, this is not true for FL sample-based models. These observations indicate that sample location and timing are of great importance when constructing a Lidar-based snow pack imputation framework.

Table 6. Workflow timing, model performance and depth correlation data using Pearson Correlation Coefficients. All results are significant at the 95 % confidence level (i.e. $p < 0.05$).

| Data Extent | Season | R^2 | RMSE | Time (min) |
|-------------|---------|-------|------|------------|
| WCW | MO 2016 | 0.61 | 0.7 | 236 |
| | MW 2014 | 0.47 | 0.7 | 252 |

| | | | | |
|-------------|---------|------|-----|-----|
| | MW 2017 | 0.44 | 0.7 | 260 |
| Flight line | MO 2016 | 0.51 | 0.8 | 70 |
| | MW 2014 | 0.44 | 0.8 | 80 |
| | MW 2017 | 0.41 | 0.7 | 76 |
| | LW 2017 | 0.41 | 1.0 | 75 |

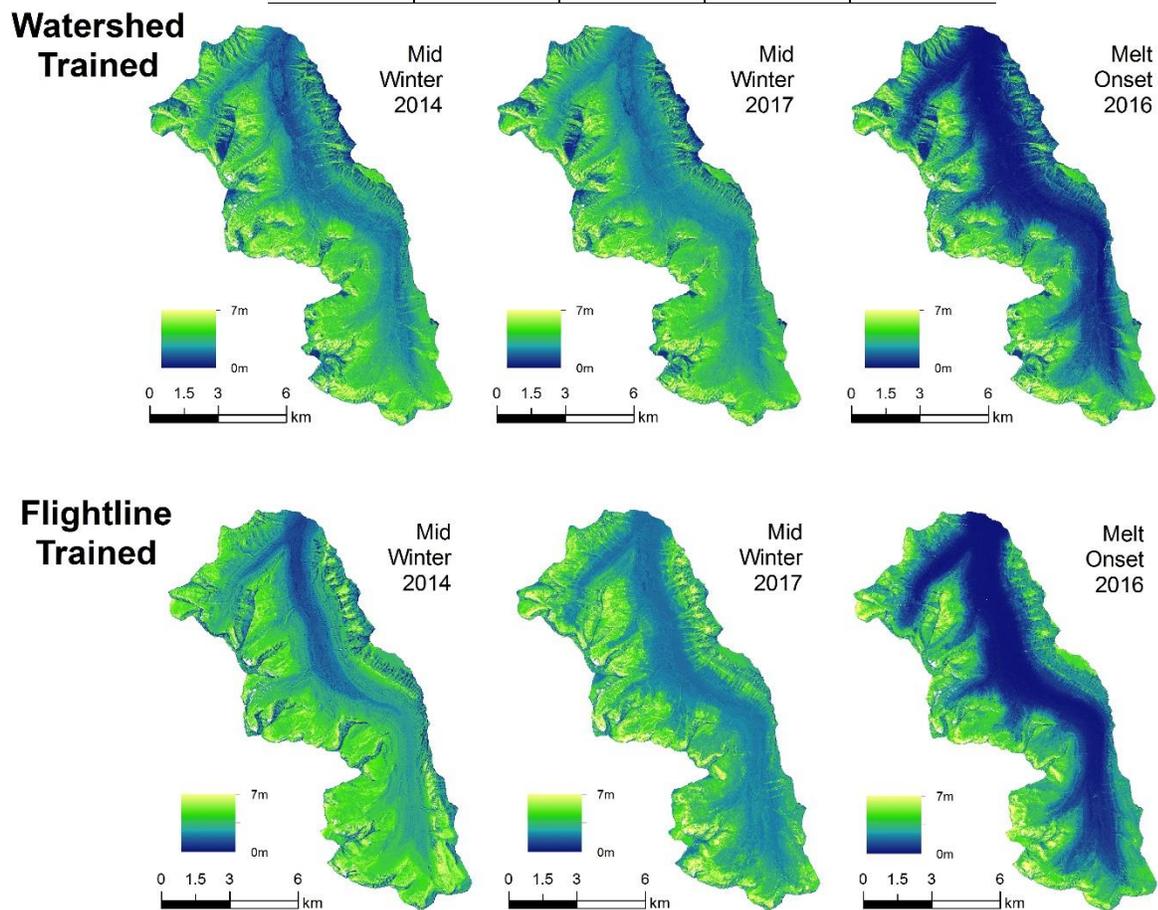


Figure 5. West Castle Watershed and Flight-line trained Random Forest Snow Depth Models.

Summary statistics (Table 7) illustrate that LSDMs and RF SDMs exhibit broadly similar summary statistics, never exceeding a difference greater than ± 0.5 m with few exceptions (excluding ‘Max’). Mean values aligned best between LSDM and WCW-based models, and less so with FL-based models, due to the sampling bias of FL training data towards high elevations where snow packs tend to be deeper (Cartwright et al. In Review). Flight-line based training points produced greater mean depths than observed in the LSDMs or WCW-based models. Smaller standard deviations were observed for FL-based models compared to LSDMs and WCW statistics. RF-predicted maximum depth values were consistently smaller than within LSDMs. This is an expected result, as RF models tend to systematically under-estimate extreme values. Taking the mean of many (leaf) estimation units will always yield a result less than the maximum value of all the leaf values unless all leaf values are equal, which is unlikely. Moreover, extreme values, by definition, represent only a small portion of the training dataset and are therefore unlikely to be selected in many bootstrap samples which are used to train individual decision trees in the RF model. If these values do not feature in training often, it is less likely that they will feature in model predictions..

Table 7. Summary statistics for Lidar and Random Forest based snow depth models (SDMs), predicted using training data points from the watershed (WCW) or flight-lines (FL).

| Season | SDM | Mean (m) | Stdev (m) | Max (m) |
|---------|--------|----------|-----------|---------|
| MO 2016 | LSDM | 1.2 | 1.2 | 6.3 |
| | WCW RF | 1.2 | 0.9 | 5.5 |
| | FL RF | 1.5 | 1.1 | 5.3 |
| MW 2014 | LSDM | 1.5 | 0.9 | 5.4 |
| | WCW RF | 1.5 | 0.7 | 4.7 |
| | FL RF | 2.2 | 0.7 | 4.5 |
| MW 2017 | LSDM | 1.5 | 0.9 | 8.1 |
| | WCW RF | 1.5 | 0.7 | 7.2 |
| | FL RF | 1.9 | 0.8 | 6.2 |
| LW 2017 | LSDM | 1.3 | 1.5 | 8.1 |
| | FL RF | 2.0 | 1.0 | 7.2 |

Discussion

Sampling and Model Optimization

Knowledge of snow depth predictor variable importance over multiple years can help guide decisions in sampling schemes if any terrain and landcover stratification is to be implemented. The temporal element of this analysis can also guide timing of data collection by illustrating which time during the winter yields better predictive results: mid-winter or melt onset. Mid-winter sampling might produce similar results, while melt onset may represent more homogeneous conditions in late season snowpacks (Lopez-Moreno et al. 2014). The potential viability of flight line sampling as a basis for RF snow depth model training is supported by the results in Tables 6 and 7. However, the disparities in depth and driver class proportions between flight line training data and the whole WCW is likely a function of sub optimal sampling locations. The flight lines (Figure 1) were surveyed on the east and west edges of the N/S oriented watershed. By sampling the watershed boundaries, high elevation (>2100 m a.s.l.) and upland terrain is more abundant in the training sample (Table 5). With much of the FL training data falling within treeline and above, higher depths in these zones will skew predictions if elevation is an important predictor. The WCW training data class proportions are similar to the actual WCW class sizes, although the WCW and FL training data included fewer closed canopy points than are present in the entire WCW (Table 5). Canopy cover decreases with elevation gains above treeline in the WCW, explaining why more cells with open and intermediate canopy attributes were present compared to closed canopy. Closed canopies in this area are known to display lower depth distributions compared to the other classes (Cartwright et al. In Review), and their under-sampling in the training data may inflate depth values. Mean depth values in Table 7 remained the same for WCW based predictions as they were in the original LSDM, suggesting that the training data was not influenced by the distribution of canopy characteristics in the input dataset. The FL training data was represented by 98.1% of upland terrain, which would favour lower depth values (Cartwright et al., In Review; Hopkinson et al. 2004). With varying driver importance across seasons and data extents (Figures 4a & 4b) as well as the potential for variables to interact, it is difficult to attribute the random forest modelling results solely to training data point location. The under or over representation of certain driver classes (Table 6) does however provide a basis to guide future flight line planning, especially when combined with RF-based driver importance rankings.

Optimized Lidar sampling with strategic flight line positioning could increase the viability of using Lidar to monitor snowpack. Efficiency is also a function of the data processing and analysis steps. One area to reduce computation time in the modelling workflow is observed with forest size (Table 3). Creating too many or too few trees in a random forest can lead to over or under-fitting of models (Kuhn and Johnson, 2013) so it was important to test a range of values. R^2 and RMSE were unchanged by forest size, but the computation times varied. It was decided that 250 trees would be a conservative choice when predictions were made over the 103 km² watershed, without expending the time required for 500 or 1000 trees. The faster production of models with integer formats and fewer points was expected, as it is an established concept in spatial and visual data processing to reduce storage space/memory with certain data formats or types of compression (Bonham-Carter, 1994). Table 4 illustrates a RAM limitation when 100,000 training points were used, as the difference in run time for integer and categorical input

data is much smaller compared to the 50,000 point trials. The RMSE value for the 100,000 training sample size does not produce lower RMSE values. Converting float data to integer format using a multiplier is an ideal approach to time-saving without compromising end results in term of both R^2 and RMSE. While a larger sample size may increase correspondence between the LSDMs and RF SDMs, the significant correlations between these datasets (Table 6) suggest that the sample size of 50,000 (or 0.44%) is sufficient. If a stronger correlation was desired at the cost of computational run time, training data point placement (i.e. ensuring a representative variety of terrain and cover attributes are sampled) and large sample sizes could be considered.

Snow Depth Driver Variables

Elevation

Elevation has consistently proven to be a valuable predictor of snow depth distribution (Revuelto, 2014b; Elder et al. 1998; Erxleben et al. 2002; Hopkinson et al., 2012; Molotch and Bales, 2005; Grunewald et al. 2015). Lopez-Moreno et al (2010) found that elevation increases in importance as grid size increases. Elevation was often a useful variable as indicated by MSE (Figures 4a & 4b), which has also been found in other RF snow depth studies (Tsai et al. 2019). Its utility was pronounced at melt onset, which is likely influenced by spring warming and environmental lapse rates (Pigeon and Jiskoot, 2008) creating a negative temperature gradient from valley bottoms to mountain summits. Lower elevations are also subject to rain on snow events in spring time. Rain adds thermal energy to the snowpack (Garvelmann et al. 2015), enhancing melt in areas that may already be prone to reduced snow inputs relative to upper elevations, where precipitation may still be adding to the snowpack.

Snowpacks become more homogenous as melt conditions progress (Lopez-Moreno et al. 2017), so a variable such as elevation which operates over larger scales (1000s of meters) rather than micro-features, such as changes in terrain or canopy cover, explains much of the overall watershed-scale snow depth distribution. Snow water equivalent records indicate that the mid and late winter datasets were surveyed while snow was still accumulating (Cartwright et al. In Review), meaning that snow from upper elevations may have limited gravitational redistribution down slope relative to the end of the season, making elevation a less important driver than it was at melt onset. Wind erosion and reduced shading by vegetation promotes lower and/or non-linear snow accumulation with elevation gains above treeline (Grunewald et al. 2014; Hopkinson et al., 2012; Zheng et al. 2016), and it is possible that these effects were more pronounced in the February 2014 and April 2016 datasets, increasing elevation's importance as a snow depth control. These observations suggest Lidar sampling and RF modeling is most successfully implemented late in the winter season, after a period of snowpack settling and homogenisation, several days after fresh snowfall.

Aspect

Aspect consistently performed well across the datasets. South facing terrain often accumulates the least amount of snow (Hopkinson et al. 2012; Kirchner et al. 2014; Zheng et al. 2016) primarily due to high solar radiation inputs relative to north-facing slopes. Higher accumulation on north facing terrain can be the result of reduced radiative inputs due to daytime shadowing and solar azimuth (Anderson and West, 1965; Haupt, 1951). In a watershed where strong wind gusts originate from south and south-westerly directions, the potential for complimentary wind and radiation loading effects exist. Wind erosion on the south and southwest slopes could enhance the ablation of snow that is already subject to more radiation-induced melt than on opposite facing terrain and therefore lower depth values result (Cartwright et al. In Review). A combination of spatially variable radiation budgets as well as wind transport (Hiemstra et al. 2002) could explain why aspect ranked highly, as both are well-established drivers of snow depth distributions. With aspect consistently in the top three important variables for all seasons and extents, the proportion of training points for aspect classes compared to the WCW class sizes may have more influence on the predicted depths than less important drivers.

TPI

Topographic position index is similar to curvature which is also used as a snow depth predictor in some studies (e.g. Marchand and Killingtveit, 2005; Lopez-Moreno and Nogues-Bravo, 2006; Plattner et al., 2006). These variables represent localized, potentially down to sub-meter terrain surface features where snow can either be trapped in depressions or scoured off convexities. Hopkinson et al. (2004) found that LSDM values were greater in valleys and shallower on ridge tops, with open area distributions closely related to topography. Topographic smoothing is the filling of depressions and the spaces between short vegetation (Schirmer and Lehning, 2011), and these processes increase with snow accumulation and persistence throughout the winter season. As discussed in the aspect section above, a consistent variable in terms of importance such as TPI may be more prone to bias model results if it is not proportionately represented in the training data.

Canopy Cover

The burial of low vegetation by wind redistributed and trapped snow, coupled with the decline in cover as elevation increases, could be responsible for canopy cover's low performance as a depth predictor. A study by D'Eon (2004) that examined snow accumulation in open and forested areas along an elevation gradient found that depth was significantly correlated with elevation, yet canopy and snow accumulation were only significantly correlated at lower elevations. The author attributed this to greater accumulation at higher sites with less canopy cover as a function of lower temperatures reducing ablation, suggesting that the importance of canopy cover varies with elevation (D'Eon, 2004).

Other recent Lidar-based research, using regression, has illustrated that up to 50% of snow depth distributions can be explained by canopy metrics (Zheng et al. 2018). The development of consistent relationships between cover and snow depth in this analysis may have been complicated by parameters that could not be adequately represented at the watershed scale. A likely contender in this environment is the effect of variable wind fields across complex terrain and canopies, as this is a source region for strong Chinook winds where speeds frequently exceed 100km/h (Pigeon & Jiskoot, 2008). Considering that high elevation relationships between depth and canopy are difficult to establish (D'Eon 2004), as well as the large amount of training data that falls within open and intermediate classes where depth distributions tend to be more variable in the WCW (Cartwright et al. In Review), these concepts may explain the higher mean depths produced by FL-trained models.

Slope

The West Castle Watershed is mostly forested along valley floor and sides, with a range of slope angles throughout the entire ~1200m elevation distribution. Steep slopes are often observed in alpine zones, where snow depth in the WCW is more variable than areas below treeline (Cartwright et al. In Review). Depending on the surrounding terrain and wind vectors, the top of a slope or cliff could be the location of a depth outlier due to the formation of a cornice (Schweizer et al. 2003) or an area of little snow due to wind erosion and exposure to solar radiation (Varhola et al. 2010). Areas where wind loading and/or cornice formation occur can eventually lose the built-up snow due to downward redistribution of snow by natural avalanches and sloughing (Schweizer et al. 2003). While an open area at the top of a steep slope may be bare due to these effects, the same level of canopy cover at a lower elevation may translate to enhanced accumulation (Varhola et al. 2010). If a study area is limited to open canopies and gentle/moderate slopes, slope can be an important predictor of snow depth (Grunewald et al. 2013). These interacting effects as well as the influence of wind and gravity make it difficult to establish consistent terrain-depth relationships in steep areas throughout the winter season.

When comparing slope surfaces derived from airborne Lidar it is also important to consider the amplification of vertical uncertainty over steeper slopes, as horizontal uncertainty is higher than in other areas and must be considered during quality control and analysis (Hodgson, 2004; Hodgson et al. 2005). The study area does not contain many grid cells with slopes $>60^\circ$ and most of these cells were eliminated by quality control. Future work may benefit from finding another study area with a different distribution of terrain slope values to determine if this driver should be kept in future modelling exercises. Including a wind parameter in modelling workflows may be another way to establish reliable snow depth trends in steep areas.

Watershed Snow Depth Predictions

Correlations between LSDMs and RF-predicted rasters (Table 6), indicate that using Lidar flight line sample datasets to impute watershed-scale snow depth is viable. Mid and late winter trials did not perform as well as melt onset in terms of R^2 values but all were significant and correspondence in watershed-scale mean depths is more important from a water resources inventory perspective (Table 6). Model performance results such as R^2 simply indicate the level of covariance but not absolute correspondence. Absolute error, as indicated by RMSE, was unchanged for the WCW trials though. With mid-winter datasets, more complex snow depth distribution patterns produce weaker correlations (Lopez-Moreno et al. 2017). Our lower R^2 mid-winter results as well as the conclusions of Lopez-Moreno et al (2017) could be influenced by snow that had yet to undergo further metamorphosis and redistribution relative to more settled snowpacks. By melt onset, much of the low-lying vegetation and topographic depressions are buried (Schirmer and Lehning, 2013), creating smoother surfaces for new snow to settle over top rather than being influenced by surface features. Once new snow stops accumulating, aspect and canopy-influenced ablation processes, such as radiation and sensible heat (Golding and Swanson 1986; Anderson et al. 1958a), might be more pronounced over a settled snowpack relative to wind distribution processes on a freshly accumulated snow surface. Static (or grid-cell level) features such as elevation, aspect and canopy cover, which were the top three WCW trained variables at melt onset, can therefore be more easily separated at melt onset than during mid-winter when localized turbulence and wind redistribution of fresh snow may mask their influence.

Under mid-winter conditions, WCW-based training points resulted in a RF SDM with the same mean depth as the LSDM (Table 7), despite weak but significant regression results ($R^2 < 0.5$; Table 6). Late-winter R^2 was the lowest but still significant. Standard deviations for the FL SDMs also suggest more variability from these models in most cases. Although RMSE values would generally be expected to show an inverse relationship with R^2 , this was not always the case, illustrating the need for further analysis of sample sizes and the best seed to use for all datasets. In order to objectively assess variable order importance to guide future sampling design, the X, Y locations of input data were constant across the datasets. This compromises the representativeness of the training sample compared to the depth distribution of the original LSDM as the locations of data to produce a good sample are not necessarily constant year to year. Statistical power and appropriate sample sizes could benefit from further exploration as more variety in depths might be captured from larger samples but increase the chance of type 2 statistical errors (Kaplan et al., 2014). R^2 is more sensitive to oddities in the dataset therefore it is possible that with a different seed and a more representative sample, R^2 values would be more reliable and possibly increase (Grace-Smith, 2019). However, direct comparisons of variable order importance would be less valid if varying depth driver attributes were present in the various training datasets as a result of using different seeds to sample the raster stack, as we would be unable to separate the effects of varying terrain attributes from seasonality when assessing variable order. For the purposes of this exploratory analysis, assessment of watershed scale means provides a simple metric that illustrates how a dense dataset with millions of data points can be modelled within reasonable accuracy of the original LSDMs.

It is clear from the mean depth results (Table 7) that the FL RF SDM outputs for the watershed are higher relative to the RF SDM or LSDM means. Given no such systematic over-estimation occurs for the watershed-scale RF SDM, this bias must be the result of the sampling configuration. While all landcover and terrain types are represented in the training datasets (Table 5), the centre of the valley, where elevations <1700 m a.s.l. and closed canopy covers are greatest, is under-represented by FL samples which likely explains the higher mean depths predicted from FL training data. Improved FL model results would be facilitated through more strategic flight line sampling to represent the range of land surface attributes and snow depth driver classes experienced within the watershed.

While many studies have included a radiation parameter (e.g. see Table 1), researchers have found that using a canopy cover or forest density parameter is an acceptable proxy for transmittance and better represents radiative inputs than parameters calculated from a DEM that excludes forest cover (Davis et al 1997). Erxleben et al (2000) found using the BRT methodology that model performance was improved when the radiation parameter was excluded, so only vegetation, elevation, aspect and slope were employed in model development. Considering these findings, and the ultimate goal of an efficient modelling workflow, we limited the predictors to the five chosen in situ (grid cell-level) properties. This is a simplification of reality, as other external (beyond the grid cell) snowpack controls, such as wind or radiation load are ignored due to such spatial and temporal continuous data not being readily available. Dynamic external drivers that vary in time and space can be simulated, however (Winstral et al. 2002; Molotch et al. 2005) but are expected to have variable predictive power and thus add non-deterministic complexity that may reduce the effectiveness of a machine learning snow depth extrapolation framework. Additional predictors may improve modelling results in future analyses.

Although Lopez-Moreno 2017 also used Lidar and random forest, they applied the algorithm to determine variable order importance, not to predict depths. With R^2 values ranging from 0.41-0.61 (Table 6), the RF results presented here exceed others' model performance in some cases (Table 1). Grunewald et al. 2013 reported R^2 values of 0.27-0.90 for MLR based models using Lidar data in various mountainous study areas. It is important to recognize that the grid cell resolution in this modelling study, 3m, is high and the results in Grunewald et al. (2013) include resolutions of 100m, 200m and 400m. With a high spatial resolution, it can be more difficult to draw a representative sample from the population because the population size is larger for a high-resolution dataset than it is with a coarser grid cell size. <1% of the available depth values were utilized in the 3m RF modelling process. Random Forest is advantageous for such dense datasets compared to MLR or other methods primarily because the algorithm splits the data and introduces randomness by variable permutation at tree nodes. The re-creation of watershed scale snow maps from spatially sampled datasets is a novelty of this work, and the results presented demonstrate there is potential for further refinement.

Conclusion

Random forest is a promising machine learning routine for modelling snow depth from airborne Lidar sample data, especially if strategic flight paths are chosen. Integration of this algorithm into snowpack monitoring frameworks may require an annually adaptive approach. Although the variables aspect, TPI and elevation were often chosen for random sampling at node splits, their inconsistent ranking suggests that one variable cannot be permanently prioritized over others when designing sampling routines using random forest. Other studies have suggested that importance measures are site-specific (Grunewald et al. 2013), but importance appears to be also sensitive to seasonal timing based on our results in the West Castle Watershed. While melt onset (April) yields the best combined model performance metrics (RMSE and R^2), similar watershed depth means at mid-winter suggest the ideal time for collection of partial Lidar datasets as inputs to integrated hydrological monitoring frameworks is not necessarily limited to a certain season if other metrics of model performance are considered. Modeling late-season snowpack is desirable as it is this snow, not what is present earlier in the accumulation season, which represents the stored freshwater resource that has the potential to influence subsequent downstream flood or drought hazards. However, if and when early or mid-winter watershed-scale snowpack conditions are desired, then Lidar sampling and RF modeling should be implemented after a period of snowpack settling and homogenisation, and not immediately following fresh snowfall.

Significant correlations between FL and WCW RF SDMs and LSDMS (Table 7) demonstrate that Random Forest has great potential for snow depth imputation as part of an overall operational Lidar snow depth monitoring framework. Collection of data over an optimally selected flight path that considers variable importance of drivers should improve model performance when trained from spatial sample data. This is an objective that needs to be addressed to operationalize Lidar-based snow monitoring in a cost- and time-effective manner. Further analysis is required to develop such a flight line sampling strategy. Using spatial sampling to train the RF routine is a viable approach to integrated snowpack/hydrological monitoring with airborne Lidar, but more research is needed to build landcover and terrain sample optimisation into a flight planning framework. This analysis shows that airborne Lidar in tandem with random forest produces results that are comparable to, or in some cases better than, previous work using multiple linear regression and binary regression trees to model snow depth over large spatial scales. Overall, Lidar is a valuable high resolution snowpack monitoring tool and this research demonstrates the potential for use in a more time- and cost-efficient operational sampling and imputation mode, as opposed to the more traditional wall-to-wall mapping mode.

Acknowledgements.

Funding for the operational lidar snow depth monitoring study was provided by the Government of Alberta Environment and Parks (AEP), Alberta Innovates Energy and Environmental Solutions – Water Innovation Program, and the Natural Sciences and Engineering Council Discovery Grants program. Hopkinson also acknowledges laboratory equipment funding from the Canada Foundation for Innovation. Field logistical and data collection support was provided

Clean Harbors Airborne Imaging, Castle Mountain Resort and AEP. Reed Parsons and Dr. Craig Mahoney are gratefully acknowledged for assistance with field data collection.

References

- Amani, M., Mahdavi, S., Afshar, M., Brisco, B., Huang, W., Mirzadeh, S., White, L., Banks, S., Montgomery, J., Hopkinson, C. 2019. Canadian Wetland Inventory using Google Earth Engine: The First Map and Preliminary Results. *Remote Sensing* 11 (7), 842
- Anderson, H.W., Rice, R.M. and West, A.J. (1958a). Snow in forest openings and forest stands. *Proceedings of the Society of American Foresters*, 46–50.
- Anderton, S.P. (2000). An analysis of spatial variability in snow processes in a high mountain catchment. PhD Thesis, University of Durham, UK, 251 pp.
- Balk, B., and Elder, K. (2000). Combining binary decision tree and geostatistical methods to estimate snow distribution in a mountain watershed. *Water Resources Research*, 36(1), 13-26.
- Baños, I. M., Garcia, A. R., Alavedra, J. M., Figueras, P. O., Iglesias, J. P., Figueras P. M. and Lopez, J. T. (2011). Assessment of Airborne LIDAR for Snowpack Depth Modelling. *Boletín de la Sociedad Geológica Mexicana* 63(1), 95–107.
- Barilotti, A., Turco, S., and Alberti, G. 2006. LAI determination in forestry ecosystems by LIDAR analysis. *Workshop on 3D Remote Sensing in Forestry*, BOKU Vienna.
- Bhardwaj, A., Sam, L., Bhardwaj, A. and Javier Martin-Torres, F. (2016). Lidar remote sensing of the cryosphere, Present applications and future prospects. *Remote Sensing of the Environment*, 177, 125-143.
- Bonham-Carter, D. (1994). *Geographic Information Systems for Geoscientists*, Volume 13. Ottawa, Pergamon. 416 pp.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- Breiman, L., Friedman, J., Stone, C. J., and Olshen, R.A. (1984). *Classification and Regression Trees*. Oxford, Taylor and Francis.
- Breiman, L., Friedman, J., Olshen, R. and Stone, C. (1984). *Classification and Regression Trees*. Chapman and Hall, Wadsworth, New York.
- Byrne, J. M., Kienzle, S. Johnson, D. G., Duke, Gannon, V., Selinger, B. and Thomas, J. (2006). Current and future water issues in the Oldman River Basin, Alberta, Canada. *Water Science and Technology*, 53, 327-334.
- Chang, K. T. and Li, Z. (2000). Modelling snow accumulation with a geographic information system. *International Journal of Geographical Information science*, 14(7), 693-707.

- Chignell, S. M., Luizza, M. W., Skach, S., Young, N. E and Evangelista, P. H. An integrative modeling approach to mapping wetlands and riparian areas in a heterogeneous Rocky Mountain Watershed. *Remote Sensing in Ecology and Conservation*, 2, 150-165.
- Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., and Lawler, J. T. (2007). Random Forests for Classification in Ecology. *Ecology*, 88(11), 2783-2792.
- Davis, R., Hardy, J., Ni, W., Woodcock, C., McKenzie, J., Jordan, R. and Li, X. 1997. Variation of snow cover ablation in the boreal forest: A sensitivity study on the effects of conifer canopy. *Journal of Geophysical Research*, 102, 29389-29395.
- Deems, J., Painter, T. and Finnegan, D. (2013). Lidar measurement of snow depth, a review. *Journal of Glaciology* 59(215), 467-479.
- D'Eon, R. (2004). Snow depth as a function of canopy cover and other site attributes in a forested ungulate winter range in southeast British Columbia. *BC Journal of Ecosystem Management*, 3, 1–9.
- Dickinson, W.T., and H.R. Whiteley, (1972). A sampling scheme for shallow snowpacks. *IAHS Bulletin*, 17(3): 247-258.
- Elder, K., Dozier, J. and Michaelsen, J. (1991). Snow accumulation and distribution in an alpine watershed. *Water Resources Research*, 27, 1541-1552.
- Elder, K., Michaelsen, J., and Dozier, J. (1995). Small basin modelling of snow water equivalence using binary regression tree methods. *Symposium on Biogeochemistry of Seasonally Snow-Covered Catchments*. IAHS Publication 228, 129–139.
- Elder, K., Rosenthal, W., and Davis, R. E. (1998). Estimating the spatial distribution of snow water equivalence in a montane watershed. *Hydrological Processes*, 12, 1793–1808.
- Erxleben, J., Elder, K. and Davis, R. (2002). Comparison of spatial interpolation methods for estimating snow distribution in the Colorado Rocky Mountains. *Hydrological Processes* 16(18), 3627-3649.
- Garvelmann, J., Pohl, S. and Weiler, M. (2015). Spatio-temporal controls of snowmelt and runoff generation during rain-on-snow events in a mid-latitude mountain catchment. *Hydrological Processes*, 29, 3649-3664.
- Geist, T. and Stotter, J. (2008). Documentation of glacier surface elevation change with multi-temporal airborne Lidar laser scanner data – Case Study: Hintereisferner and Kesselwandferner, Tyrol, Austria. *Z. Gletscherkd. Glazialgeol*, 41, 77-106.
- Golding, D. L. (1974b). The correlation of snowpack with topography and snowmelt runoff on Marmot Creek Basin, Alberta. *Atmosphere*, 12, 31–38.
- Golding, D., and Swanson, R. (1986). Snow distribution patterns in clearings and adjacent forest. *Water Resources Research*, 22(13), 1931–1940.

- Grace-Smith, K. (2019). Retrieved from: <https://www.theanalysisfactor.com/assessing-the-fit-of-regression-models/>
- Grunewald, T., Buhler, Y. and Lehning, M. (2014). Elevation dependency of mountain snow depth. *The Cryosphere* 8, 2381-2394.
- Grünewald, T., Schirmer, M., Mott R. and Lehning, M. (2010). Spatial and temporal variability of snow depth and ablation rates in a small mountain catchment. *The Cryosphere*, 4, 215-225.
- Grunewald, T., Stotter, J., Pomeroy, J.W., Dadic, R., Banos, I. M., Marturia, J., Spross, M., Hopkinson, C., Burlando, P. and Lehning, M. (2013). Statistical modelling of the snow depth distribution in open alpine terrain. *Hydrology and Earth Systems Science*, 17(8), 3005-3021.
- Hiemstra, C.A., Liston, G.E., and Reiners, W.A. (2002). Snow redistribution by wind and interactions with vegetation at upper treeline in the Medicine Bow Mountains, Wyoming, USA. *Arctic and Antarctic Alpine Research*, 34, 262–273.
- Hodgson, M. E., and Bresnahan, P. (2004). Accuracy of Airborne Lidar-Derived Elevation. *Photogrammetric Engineering and Remote Sensing*, 70(3), 331-339.
- Hodgson, M.E., Jensen, J., Raber, G., Tullis, J., Davis, B.A., Thompson, G., and Schuckman, K. (2005). An evaluation of Lidar-derived elevation and terrain slope in leaf-off conditions. *Photogrammetric Engineering and Remote Sensing*, 71(7), 817- 823.
- Hopkinson, C., Pomeroy, J., DeBeer, C., Ellis, C., and Anderson, A. (2011), Relationships between snowpack depth and primary Lidar point cloud derivatives in a mountainous environment. *Proceedings of Remote Sensing and Hydrology Symposium*, 352.
- Hopkinson, C., Sitar, M., Chasmer, L.E. and Treitz, P. (2004). Mapping snowpack depth beneath forest canopies using airborne LIDAR. *Photogrammetric Engineering and Remote Sensing*, 70, 323-330.
- Hopkinson, C., Collins, T., Anderson, A., Pomeroy, J. and Spooner, I. (2012). Spatial Snow Depth Assessment Using Lidar Transect Samples and Public GIS Data Layers in the Elbow River Watershed, Alberta. *Canadian Water Resources Journal*, 37(2), 69-87.
- Hosang, J., and Dettwiler, K. (1991). Evaluation of water equivalent of snow cover map in a small catchment area using a geostatistical approach. *Hydrological Processes* 5(3): 283-290.
- Jeong, J. H., Resop, J. P., Mueller, N. D., Fleisher, D. H., Yun, K., Butler, E. E., Timlin, D. J., Shim, K., Gerber, J. S., Reddy, V. R. and Kim, S. (2016). Random Forests for Global and Regional Crop Yield Predictions. *PlosOne*, 11(6), 15 pp.
- Jonas, T., Marty, C., and Magnusson, J. (2009). Estimating the snow water equivalent from snow depth measurements in the swiss alps. *Journal of Hydrology*, 378, 161-167.
- Jost, G., Weiler, M., Gluns, D. and Alila, Y. (2007). The influence of forest and topography on snow accumulation and melt at the watershed-scale. *Journal of Hydrology* 347:101–115.

- Kaplan, R., Chambers, D. and Glasgow, R. Big Data and Large Sample Size: A Cautionary Note on the Potential for Bias. *Clinical and Translational Science*, 7(4), 342-346.
- Kienzle, S. (2004). The effect of DEM raster resolution on first order, second order and compound terrain derivatives. *Transactions in GIS*, 8(1), 83-111.
- Kienzle, S. W., and Mueller, M. (2013). Mapping Alberta's surface water resources for the period 1971-2000. *Canadian Geographer*, 57(4), 506–518. <https://doi.org/10.1111/j.1541-0064.2013.12050.x>
- Kirchner, P. B., Bales, R. C., Molotch, N. P., Flanagan, J. and Guo, Q. (2014). Lidar measurement of seasonal snow accumulation along an elevation gradient in the southern Sierra Nevada, California. *Hydrology and Earth System Sciences* 18(10), 4161-4275.
- Kuhn, M. and Johnson, K. (2013). *Applied Predictive Modelling*. New York, Springer.
- Kuz'min, P. P. (1960). Snowcover and snow reserves. *Gidrometeorologicheskoe, Izdatel'sko*, Leningrad. [Translation], U.S. National Science Foundation, Washington, D.C.
- Lehning, M., Grünewald, T. and Schirmer, (2011). Mountain snow distribution governed by an altitudinal gradient and terrain roughness. *Geophysical Research Letters*, 38, 5 pp.
- Liaw, A. and Wiener, M. (2018). Retrieved from: <https://cran.r-project.org/web/packages/randomForest/randomForest.pdf>
- Lopez-Moreno, J. I. and Nogues-Bravo, D. (2006). Glacier development and topographic context. *Earth Surface Processes and Landforms* 31(12), 1585-1594.
- López-Moreno, J.I., Latron J., and Lehmann, A. (2010). Effects of sample and grid size on the accuracy and stability of regression-based snow interpolation methods. *Hydrological Processes* 24(14), 1914–1928.
- Lopez-Moreno, J., Revuelto, J., Alonso-Gonzalez, E., Sanmiguel-Valladolid, A., Fassnacht, S. R., Deems, J., and Moran-Tejeda, E. (2017) Using very long-range terrestrial laser scanner to analyze the temporal consistency of the snowpack distribution in a high mountain environment. *Journal of Mountain Science*, 14(5), 823-842.
- Marchand, W. D. and Killingtonveit, A. (2005). Statistical probability distribution of snow depth at the model sub-grid cell spatial scale. *Hydrological Processes*, 19, 355–369.
- Molotch, N. P. and Bales, R.C. (2005). Scaling snow observations from the point to the grid element, Implications for observation network design. *Water Resources Research* 41(11), W11421, doi: 10.1029/2005WR004229.
- Molotch, N. P., Colee, M.T., and Bales, R.C. (2005). Estimating the spatial distribution of snow water equivalent in an alpine basin using binary regression tree models, the impact of digital elevation data and independent variable selection. *Hydrological Processes* 19(7), 145901479.

- Nogues-Bravo, D. (2003) El estudio de la distribución espacial de la biodiversidad, conceptos y métodos. *Cuadernos de Investigación Geográfica*, 29, 67–82.
- Oroza, C., Zheng, Z., Glaser, S., Tuia, D. and Bales, R. 2016. Optimizing embedded sensor network design for catchment-scale snow-depth estimation using Lidar and machine learning. *Water Resources Research*, 52(10), 8174-8189.
- Pigeon, K. E., and Jiskoot, H. (2008). Meteorological Controls on Snowpack Formation and Dynamics in the Rocky Mountains. *Arctic, Antarctic and Alpine Research*, 40(4), 716-730.
- Plattner, C., Braun, L., and Brenning, A. (2006). Spatial variability of snow accumulation on Vernagtferner, Austrian Alps, in winter 2003/04, *Z. Gletscherk. Glazialgeol.*, 39, 43–57.
- Pomeroy, J.W., Marsh, P., and Gray, D.M. (1997). Application of a distributed blowing snow model to the Arctic. *Hydrological Processes*, 11(98), 1451-1464.
- Revuelto, J., Lopez-Moreno, J. I., Azorin-Molina, C., and Vicente-Serrano, S. M. (2014b). Topographic control of snowpack distribution in a small catchment in the central Spanish Pyrenees, Intra- and inter- annual persistence. *Cryosphere Discussions*, 8(2), 1937–1972.
- Schweizer, J., Jamieson, J. B., and Schneebeli, M. (2003). Snow avalanche formation. *Revised Geophysics*, 40(4), 1016, doi:10.1029/2002RG000123.
- Schirmer, M., and Lehning, M. (2011). Persistence in intra-annual snow depth distribution, Fractal analysis of snow depth development. *Water Resources Research*, 47, doi, 10.1029/2010WR009426.
- Steppuhn, H.W. (1976). Areal water equivalents for prairie snowcovers by centralized sampling. *Proceedings of the Western Snow Conference*, 44, 63-68.
- Tierney, L. T., Rossini, A. J., Li, N., and Sevcikova, H. (2018). Retrieved from: <https://cran.r-project.org/web/packages/snow/snow.pdf>
- Tinkham, W.T., Smith, A, Marshall, H. P., Link, T. E., Falkowski, M. J., and Winstral, A. H., (2014). Quantifying Spatial Distribution of Snow Depth Errors from Lidar Using Random Forest. *Remote Sensing of Environment*, 141, 105-115.
- Tsai, Y., Andreas, D., Oppelt, N. and Kuenzer, C. 2019. Wet and dry snow detection using Sentinel-1 SAR data for mountainous areas with a machine learning technique. *Remote Sensing*, 11(8), 29 pp.
- Varhola., A., Coops, N.C., Weiler, M., and Moore, R.D. (2010). Forest canopy effects on snow accumulation and ablation, An integrative review of empirical results. *Journal of Hydrometeorology* 392, 219-233.
- Wang, S., Deng, J., Chen, M., Weatherford, M. and Paugh, L. (2015). Random Forest Classification and Automation for Wetland Identification based on DEM Derivatives. Retrieved from:

https://www.researchgate.net/publication/305385776_Random_Forest_Classification_and_Automation_for_Wetland_Identification_based_on_DEM_Derivatives

Winstal, A., Elder, K., and Davis, R.E. (2002). Spatial Snow Modelling Of Wind-Redistributed Snow Using Terrain-Based Parameters. *Journal of Hydrometeorology*,3, 524-538.

Zheng, Z., Kirchner, P.B., and Bales, R. C. (2016). Topographic and vegetation effects on snow accumulation in the southern Sierra Nevada, a statistical summary from Lidar data. *The Cryosphere*, 10, 257-269.

Zheng, Z., Qin, M., Qian, K. and Bales, R. C. (2018). Canopy effects on snow accumulation: observations from Lidar, canonical-view photos, and continuous ground measurements from sensor networks. *Remote Sensing*, 10(11), 1796.

Appendix 6: In-Situ LED-Based Observation of Snow Surface and Depth Transects

Barnes, C., Hopkinson, C., Porter, T., Xi Z. 2020. IN-SITU LED-BASED OBSERVATION OF SNOW SURFACE AND DEPTH TRANSECTS. *Sensors* 20 (8), 2292

In-Situ LED-Based Observation of Snow Surface and Depth Transects

Celeste Barnes, Chris Hopkinson *, Thomas Porter, and Zhouxin Xi

Department of Geography, University of Lethbridge, Lethbridge, AB T1K 3M4, Canada; celeste.barnes@uleth.ca (C.B.); portta@uleth.ca (T.P.); zhouxin.xi@uleth.ca (Z.X.)

* Correspondence: c.hopkinson@uleth.ca (C.H.); Tel.: +1-403-332-4586

Received: 7 March 2020; Accepted: 15 April 2020; Published: date

Abstract: As part of a new snowpack monitoring framework, this study evaluated the feasibility of using an LED LIDAR (Leddar) time of flight sensor for snowpack depth measurement. The Leddar sensor has two additional features over simple sonic ranging sensors: (i) the return signal is divided into 16 segments across a 48° field of view, each recording individual distance-to-target (DTT) measurements; (ii) an index of reflectance or intensity signal is recorded for each segment. These two features provide information describing snowpack morphology and surface condition. The accuracy of Leddar sensor DTT measurements for snow depth monitoring was found to be < 20 mm, which was better than the 50 mm quoted by the manufacturer, and the precision was < 5 mm. Leddar and independent sonic ranger snow depth measurement showed strong linear agreement ($r^2 = 0.98$). There was also a strong linear relationship ($r^2 = 0.98$) between Leddar and manual field snow depth measurements. The intensity signal response was found to correlate with snow surface albedo and inversely with air temperature ($r = 0.77$ and -0.77 , respectively).

Keywords: Leddar; snowpack depth; LIDAR; LED LIDAR; sonic ranging device; intensity

1. Introduction

Wintertime snow accumulation and associated snowmelt provide a significant contribution to water resources in regions with seasonal snow packs [1] as surface water is “locked up” in a frozen storage state [2, 3]. Snow depth and density measurements are required to quantify the amount of snow water equivalent (SWE) that will be released from the snowpack at the time of melt. Of the two variables, snow depth is the major component of SWE [4]. Alberta Rocky Mountain headwater snowpack monitoring [5] has been operational for several decades, using both destructive and non-destructive techniques [6]. A snow probe and weighing tube are used to obtain field validation depth and SWE [7, 8]. This provides an accurate single point measurement but disturbs the snowpack in the process, making future measurements difficult to repeat at the same location. The second operational monitoring method in the headwaters uses non-destructive sonic ranging devices [9] to measure snowpack depth [10, 11]. Such units can be mounted on a tower pointing over the ground surface. Ultrasonic pulses are emitted and echoes are received by the sensor. The return signal is used to calculate snow depth by differencing distance-to-target (DTT) measurements when the snow surface is present from the DTT observation of the no-snow ground surface, following air temperature compensation. These sensors have a high degree of accuracy, but depth is an average obtained from the total sensor footprint (the further from the target, the larger the footprint) including undulations in the ground and snow surface. Snowpack depth is continuously recorded over the snow season. Sonic ranging sensors do not describe the snowpack morphology, structure, or density characteristics.

The intent of this study was to examine a low-cost (≈ 1000 Canadian dollars) and low-power (4 watt) alternative to the contemporary sonic ranging snow depth sensor. The low-cost low-power Leddar Tech IS16 (hereafter referred to as Leddar) sensor is a LIDAR (light detection and ranging) based “time of flight” device. It emits a single LED (light emitting diode)-diffused light source beam in the near infra-red (NIR) 940 nm wavelength. The return signal is divided into 16 segments. Both temperature compensated DTT and light intensity are observed and measured for each segment. The Leddar unit has added features over the sonic ranging sensor, those being a DTT and intensity response for each of the 16 segments of its footprint. There is limited light transmission into the surface of natural snowpacks [12], making LED technology an ideal candidate for snow depth monitoring by calculating the difference of the DTT snow surface from the bare earth DTT. Sonic ranging sensor pulses penetrate freshly fallen snow, producing an underestimation of snowpack depth [11, 13]. Beyond evaluating the potential to observe snow surface height and snowpack depth, a further goal of this study is to evaluate which, if any, snowpack features can be inferred from the additional intensity signal attribute collected by LED ranging sensors.

The snowpack surface texture and reflectance change over time as the pack evolves due to increasing and decreasing depths from snow accumulation, ablation, compaction, and wind/gravity-induced redistribution [14, 15]. As the snowpack evolves, variation in crystal grain size and structure [16-18], density, and the amount of water within the pack occurs [19, 20]. Ice lenses form within the pack during melt and refreeze cycles [6]. Impurities such as dust or dirt settle in the snowpack [19]. Coalesced and/or wet snow has a lower albedo than freshly accumulated snow [21]. As the snowpack ages and eventually becomes isothermal, the reflectance signature changes [22, 23].

The spectral response or albedo of snow is dependent on snow crystal grain size, age of the snow, and amount of water and impurities in the pack [12]. This causes the spectral response to be different at various stages of metamorphism. “Fresh” new snow produces the highest spectral reflectance over aging snow, soils, and vegetation [16] in the 940 nm band, in which the Leddar unit operates. At the 940 nm band, spectral albedo ranges from approximately 0.5 to 0.9. The continuous spatiotemporal intensity signal has the potential to be used to detect changes in snowpack surface characteristics. For example, it should be possible to use the intensity signature to detect when snow is falling on the pack’s surface, since fresh snow has a higher reflectance than older snow. These variations in the snowpack should be detectable in the Leddar’s DDT and intensity signals. The expectation is to see an increase in the spectral response for new snow accumulation and a decrease as the snowpack compacts and metamorphoses over time.

This study assessed the performance and features of the Leddar sensor. The objectives were to: (1) examine the Leddar sensor performance and the controls on the Leddar intensity signal; (2) compare the continuous temporal snowpack depth measurements of the Leddar with a standard sonic ranging sensor and infrequent field depth data observations to quantify sensor precision and accuracy for this particular application.

2. Materials and Methods

2.1. Leddar and SR50A Sensor Specifications

The Leddar and SR50A sonic ranging sensors used in this study each capture DTT measurements. An approximation of the respective Leddar and SR50A DTT sampling areas is shown in Figure 1a. The LeddarTech IS16 sensor (Figure 1b) is a solid-state, pulse-based, time-of-flight LED LIDAR ranging instrument that receives return signals on a 16-channel photodetector [24]. The unit’s internal processing chip performs a full waveform analysis on all segments of the return signal. When the “object demerging”

feature is enabled, it is possible for the sensor to detect multiple objects at varying distances and intensities within its field of view (FOV). LED technology is sensitive to ambient temperature [25], and an inverse relationship exists between external temperature and light output (illumination) [26]. A temperature sensor is located on the Leddar circuit board near the emitter optics [24, 27]. Proprietary algorithms compensate for this temperature sensitivity and are used to calculate distance measurements within the manufacturer's specified precision and accuracy [28].

The photodetector chip on the Leddar sensor has limited distance detection [27] and uses the intensity (amount of light captured by the receiver) property as part of the calculation for detecting multiple objects at varying distances within the FOV. The Leddar sensor range is zero to 50 m (Table 1). The sensor has a $48^\circ \times 8^\circ$ sampling FOV. The ground surface footprint (seen in Figure 1a) is dependant on the height of the sensor above the target and defined by the beam length and depth. The length of the beam on the ground is 0.89 multiplied by the distance to the target. Each segment is 1/16 of the total beam length. The depth of the beam is 0.14 multiplied by the distance of the sensor to the target.

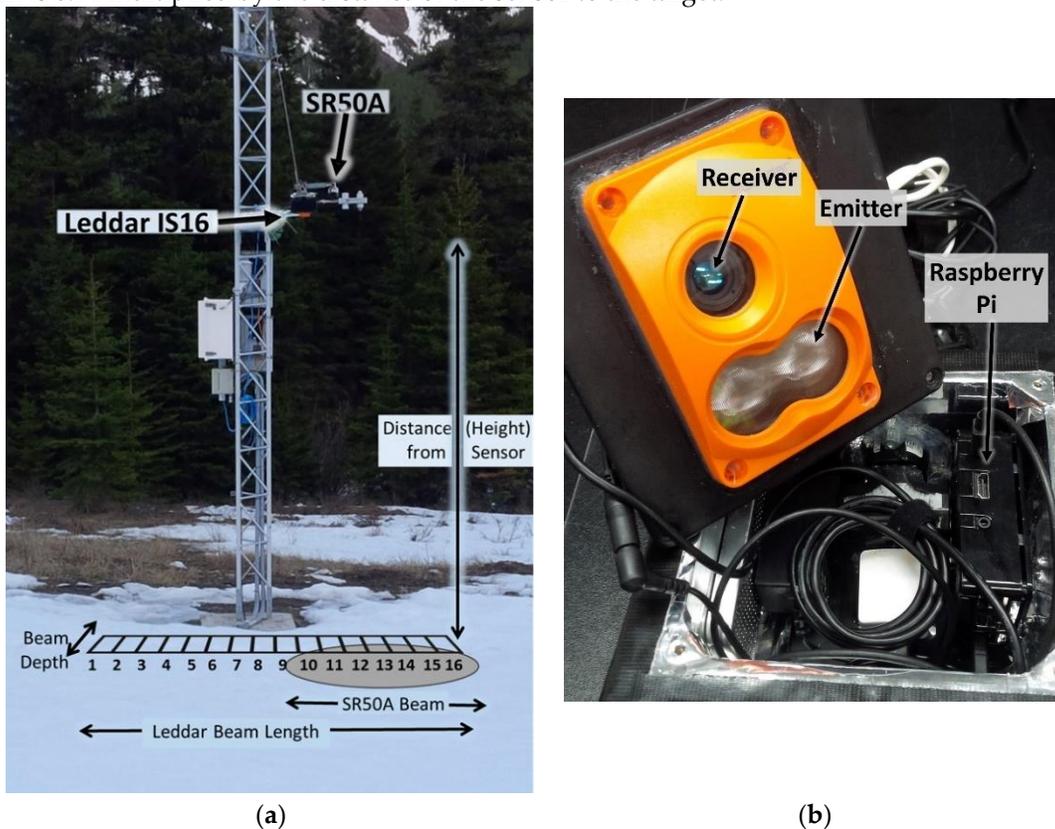


Figure 1. (a) Leddar and SR50A sonic ranging device co-located on a tower. Approximation of sensor beam footprint; not to scale. Leddar segment 1 orientation is the north side of the tower. The sensor emits defuse LED light and receives return signals for 16 segments. The ground surface footprint is dependant on the height of the sensor above the target, and the beam length and depth. (b) LeddarTech IS16 pulsed based time of flight LED LIDAR sensor logging to a Raspberry Pi 3.

LeddarTech technology is used for industrial applications, such as vehicle collision detection, in which fast response times from the sensor are required [27]. The distance measurement accuracy reported by the

manufacturer is 50 mm, the precision of the sensor (if the intensity return signal is greater than the manufacturer's specification of 15) is 6 mm, and the resolution is 10 mm. For this study, the Leddar sensor was statically mounted over the ground surface at a fixed height of 2.93 m with a calculated beam length of 2.61 m (each segment length is 0.16 m) and depth of 0.41 m. In comparison, the SR50A sensor emits ultrasonic pulses. The sensor's footprint on the ground surface is circular with an approximate radius of 0.27 multiplied by the height of the sensor above the target surface. A temperature corrected measurement for DTT is averaged over the entire footprint as a single reading. The distance range of the SR50A is 0.5 to 10 m. The SR50A DTT measurement accuracy as reported by the manufacturer is 12 mm on a 3 m high mount, and the resolution is 0.25 mm. A comparison of the technical specifications of Leddar and SR50A units is shown in Table 1.

Table 1. Leddar and SR50A sonic ranger manufacturer technical specifications [24, 29].

| Sensor | Leddar IS16 | SR50A Sonic Ranging Device |
|------------------|---|---|
| Type | Leddar LED Multichannel LIDAR sensor, built-in processing chip performing proprietary temperature adjusted Full-Waveform analysis for multi-object detection distance measurement | SR50AA Sonic Ranging Sensor with independent temperature compensation |
| Manufacturer | LeddarTech Inc. | Campbell Scientific (Canada) Corp. |
| Distance | 0 to 50 m | 0.5 to 10 m |
| Operating Temp | -40 °C to +50 °C | -45 °C to +50 °C |
| Accuracy | ±50 mm Quoted for a moving target | ±10 mm or 0.4% of DDT (greater value) |
| Precision | 6 mm (manufacturer specification if intensity > 15) | |
| Resolution | 10 mm | 0.25 mm |
| Measurement Rate | Up to 50 Hz | Less than 1.0 second |
| Emitter | Single LED diffused light source beam | Sonic Ranging ultrasonic pulses |
| Receiver | Measurement of backscatter on a 16-Channel photodetector array | Listening for return echoes |
| Beam Length | 48° (Distance from sensor * 0.8905) | 30° (Radius = 0.268 * Height) |
| Segment Length | 1/16 of the Beam Length (Beam Length / 16) | N/A |
| Beam Depth | 8° (Distance from sensor * 0.1402) | N/A |
| Wavelength | 940 nm (infrared) | 50 kHz (Ultrasonic) electrostatic transducer |

2.2. Field Deployment Setup, Configuration, and Data

Leddar and SR50A instrument testing and snowpack validation were completed at the University of Lethbridge West Castle Field Station (WFS), located in the headwaters of the Oldman River Basin, Alberta, Canada from 2017 (December 14) to 2018 (April 27). The Leddar unit was mounted on the weather station tower coincident with a SR50A such that the beams of the two sensors partially overlapped, as shown in Figure 1a. The ground surface beneath the Leddar and SR50A sensors had a downward slope of 0.07 m over a distance of 3.0 m. Meteorological data was collected at the tower site for wind speed, wind direction, temperature (temp), barometric pressure (BP), relative humidity (RH), and incoming and reflected shortwave (SW) and longwave (LW) radiation (Figure 2). A temperature sensor under the Leddar unit at the ground surface collected ground temperature. A totalizing precipitation gauge located nine meters south of the tower provided cumulative precipitation data.

To validate Leddar and SR50A sensor snow depth measurements, eleven site visits took place from December 21 to April 27. Biweekly to monthly field measurements occurred during the months of December to March. At the onset of snowmelt, sampling was done at a weekly to daily time interval. Snow depth field measurements using a graduated avalanche probe were taken by standing behind the tower out of the sensor field of view and sampled directly under the Leddar unit. Field depth measurements collocated and coincident in time with the Leddar deployment were collected to support and validate both the Leddar and SR50A measurements. Leddar, SR50A, and meteorological data were extracted from the 15-minute timestep dataset for a four-hour time period when field measurements occurred. The mean, minimum, and maximum were calculated for the SR50A. The mean for the Leddar was derived from the mean value of all segments. Minimum and maximum were extracted based on the range of values from all segments of the Leddar sensor.

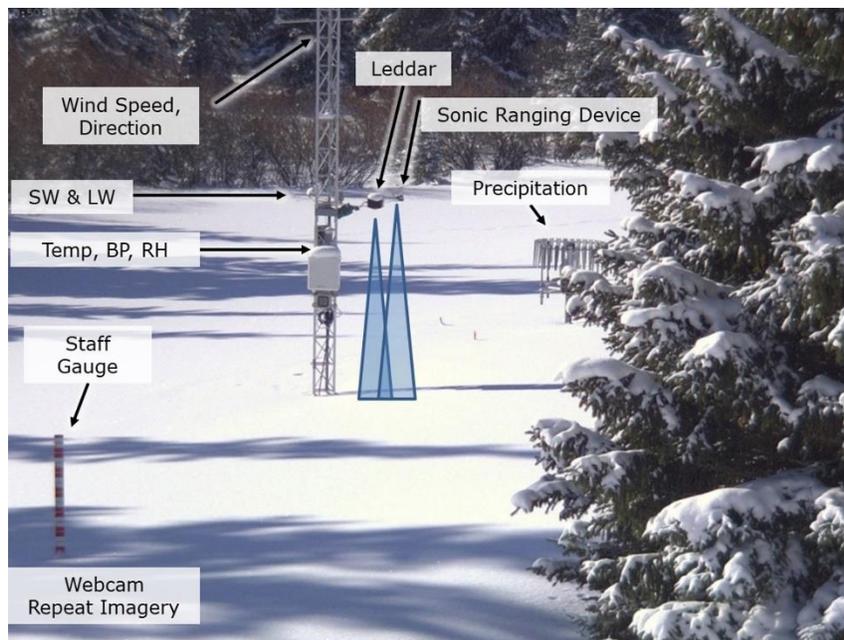


Figure 2. The West Castle Field Station tower and snow depth monitoring sensors (Leddar and SR50A), the totalizing precipitation gauge located south of the tower, and the meteorological sensors for the 2017–2018 snow season collecting wind speed, wind direction, air temperature (temp), barometric pressure (BP), relative humidity (RH), and incoming and reflected shortwave (SW) and longwave (LW) radiation. A temperature sensor was located at the ground surface to collect ground temperature.

2.2.1. LeddarTech IS16 (Leddar) Configuration and Laboratory Calibration

The Leddar configuration settings determine the accuracy, precision, and resolution of the emitted and return signals for DTT and intensity measurements (Table 2). The Leddar’s LED pulse rate is 102.4 kHz. The parameter configuration selected maximized the accuracy and precision of object detections that resulted in a maximum range of 21.3 m [24]. “Crosstalk removal” and “object demerging” were both enabled to reduce return signal degradation from objects detected in other segments.

The control and logger system for the Leddar sensor had a similar design to that of [30], wherein the Leddar unit was used for 3D digital canopy foliage sampling. The Leddar unit (Figure 1b) was connected to a Raspberry Pi 3 (RP) running the Raspbian OS through a 2.0, 12 Mbits/s USB cable (Figure 3). A computer program [31] executed Leddar SDK (software development kit) commands to obtain DTT, intensity, and status flag for a 15-minute time step. At the start of the collection interval, the Leddar unit was “turned on” to acquire continuous pulse measurements for a period of one minute. For each measurement, the 16-channel photodetector divided the return signal into separate segments where a full waveform analysis was performed by the Leddar onboard processing chip. The resulting measurement was returned to the RP. The program then went into a “wait” state for the next 14 minutes.

The RP program converted Leddar range measurements (R) into vertical distance (V_T) [32] values using Equation (1).

$$V_T = (R_i + R_c) \cos(\varphi) \cos(\alpha + \beta(i-8.5) + \gamma(i-8.5)^2) \quad (1)$$

R_i is the range measurement of the i^{th} segment, R_c is the fixed shift of the range measurement from the true range, φ is the fixed zenith boresight shift, α is the azimuth boresight shift, β is the i^{th} segment’s angle deviation from the boresight on Leddar beam plane, and γ is the 2nd order non-linear angle deviation (8.5 is the offset to the center of segments 1 to 16). Leddar calibration parameters were determined in the laboratory. The Leddar unit was mounted facing a flat level reference surface with no zenith deviation ($\varphi = 0$). Measurements were taken at varying boresight angles (α) and vertical distances (V_T) from Leddar optical emitter center to the reference surface. R_c , α , β , and γ are the calibration parameters. V_T , R_i , and φ are known values. This made it possible to infer R_c , β , and γ from Equation (1) based on the Gauss–Newton algorithm with Huber robust function [33] to get the best parameter fit in order to evaluate the accuracy of the Leddar measurement.

Table 2. Leddar measurement configuration settings used for the 2017–2018 snow season.

| Parameter | Configuration | Description |
|------------------|---------------|---|
| Distance Units | cm | Unit of measurement for distance to target |
| Accumulations | 1024 | Range: 0 to 1024. Higher values enhance the range for DTT below 10 m, reduce the measurement rate and noise |
| Measurement Rate | 1.5625 Hz | Range: 1.5625 to 50 Hz. Rate of signal measurement. Lower values give highest accuracy and precision (also known as Refresh Rate) |
| Oversampling | 8 | Range: 1–8. High values reduce measurement rate and increase accuracy |
| Point Count | 12 | The number of base sample points |

| | | |
|-------------------|------------|--|
| Threshold Offset | 0.00 | Range: -5% to 100%. Modifies intensity threshold. At 100%, no detections. Negative values increase likelihood of false measurements. |
| LED Control | Automatic | LED power level setting |
| Change Delay | 1 (640 ms) | Number of measurements before sensor changes LED power level |
| Object Demerging | Enabled | Indicates detection of multiple objects in return signal |
| Crosstalk Removal | Enabled | Degradation compensation from object detections in other segments |
| Useful Range | 21.3 m | Leddar sensor computed value based on configuration settings |

Two data files were created: the first containing the raw measurements per pulse for each of the 16 segments stored on the RP (see Figure 3 *raw.txt Data File at the top right of the diagram); the second created from the downloaded raw RP data file through a post-processing step. The raw RP data file contained a record for each segment number, timestamp, V_T , intensity, and status flag of the return signal. For each of the 16 segments, two possible status flags could be received, those being "Flag = 1" and "Flag = 35." Depending on the status flag received for the measurement cycle, the program captured between 20 to 60 observations for the 15-minute time interval. The post processing step creates a single record in the second data file for each 15-minute collection interval using the raw RP data as the input file. The record contains the timestamp; and for each segment the record is appended to include segment number, the mean distance to target and intensity for each flag, the count of the number of measurements for each flag, the total number of measurements, and a percentage of the number of "Flag = 1" returns for the given 15-minute measurement interval (see Figure 3 *Final.csv Data File at the bottom left of the diagram).

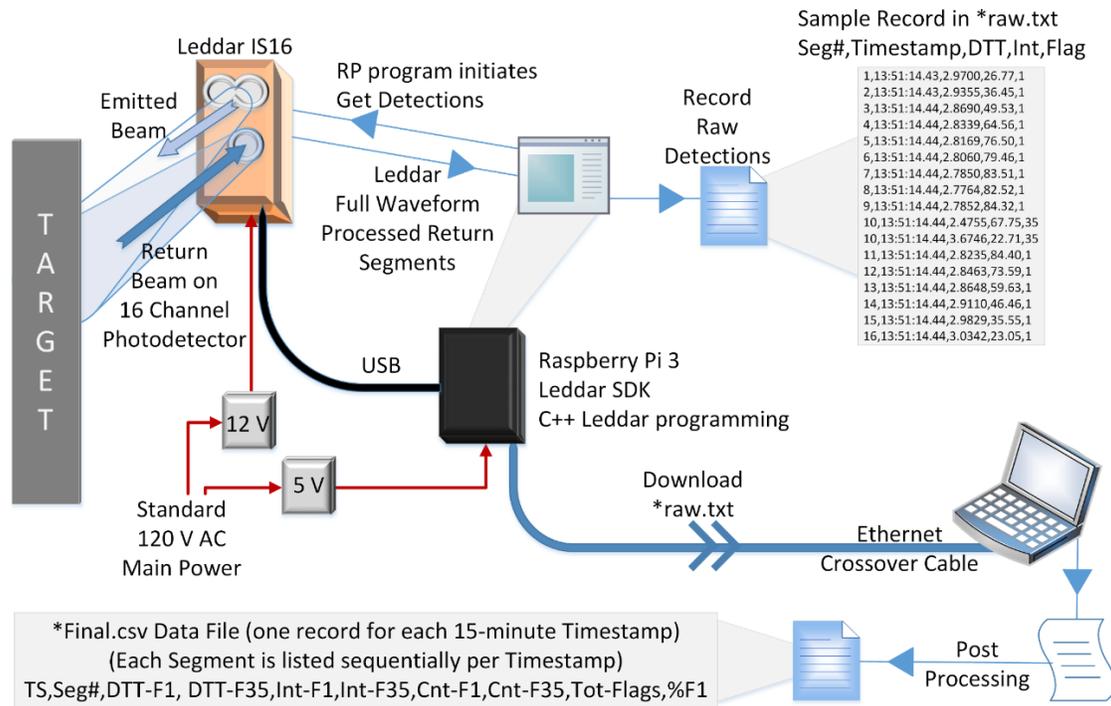


Figure 3. Data flow diagram of the Leddar-Raspberry Pi for the 2017–2018 Snow Season. TS refers to timestamp, Seg# refers to segment number, DTT-F# is the distance to target for flag number, Int-F# is the intensity for flag number, Cnt-F# is the number of returns received for the status flag number, Tot-Flags is the total count of both flag numbers, and %F1 is the percentage of Flag = 1 returns for the segment.

The Leddar’s on-board processing chip computes status “Flag = 1” as a valid return signal of a single object detected in the emitted beam. “Flag = 35” is a valid return but is interpreted by the sensor as more than one object detected within the same segment as a result of the enabled “object demerging” configuration setting. “Flag = 35” is considered noise in the data due to an implicit assumption that no objects should exist between the statically mounted Leddar sensor and the ground or snowpack surface. These measurements are retained in the final post processed data file, but they are not used for the snowpack depth calculation. The presence of “Flag = 35” does not mean all data for a given segment are invalid. If the timestamp contains data with the “Flag = 1” status, a valid return signal has been received for the specific segment. As part of the quality control process, the “noisy” (Flag = 35) V_T and intensity values are discarded from the specific segment. The count of noisy returns is retained and used in the analysis.

2.2.2. SR50A Sonic Ranging Device

The SR50A was configured using a Campbell Scientific CR1000 data logger [29] with standardized temperature-adjusted programming for a 15-minute data collection interval for the distance to target measurement. Real-time temperature compensation is performed using the temperature sensor located adjacent to the Leddar and SR50A on the instrument tower (Figure 2).

2.2.3. Meteorological Sensors

Quality assurance corrections were applied to the complete meteorological dataset. The totalizing precipitation weighing gauge accumulates rain and snow in a catchment bucket. It is susceptible to diurnal and long-term drift, evaporation, the under-catch of snow caused by wind during a precipitation event, and the over-catch from blowing snow [34, 35]. Totalizing precipitation weighing gauge data adjustments followed the methodology described by [36, 37]. Initial manual cleaning was completed to correct gauge measurements resulting from sensor maintenance tasks such as emptying liquid when the catchment bucket was full. To correct for diurnal, long term drift, and evaporation, negative and small positive changes were removed using a threshold of 0.11 mm unless a precipitation event was in progress in prior and subsequent measurements. To reduce overestimation of under catch above the sensor, the correction of the wind field was limited to speeds greater than 1.2 ms⁻¹ and less than 6.5 ms⁻¹. Precipitation measurements were calculated using the change in the catchment bucket volume from the previous measurement.

Incoming shortwave radiation (SWI) and reflected shortwave radiation (SWR) were collected from a Campbell Scientific CNR1 [38] net radiometer mounted on the same tower as the Leddar and SR50A units. CNR1 SWI and SWR data were used to calculate albedo (ratio of SWR over SWI) for both a one-hour and daily averaged interval. Measurements greater than one were discarded as SWI must always be greater than SWR. Undulations in the snow surface or other highly reflective objects in close proximity to the CNR1 can produce additional backscatter that is recorded by the sensor [39]. There were five days when the albedo calculation was greater than one. For those days, daily albedo was computed using the mean from the previous and subsequent day.

2.3. Aggregated Datasets for Analysis

The final post-processed 15-minute dataset contained 12,844 records. Snow depth (D_{snow}) was calculated for each of the Leddar segments and SR50A sensors by subtracting the baseline “no-snow” ($V_o = 2.93$ m) ground measurements from subsequent “snow” surface (V_T) measurements (Equation (2)). The mean Leddar D_{snow} , V_T , and intensity were calculated across the 16 segments, where V_T is estimated from Equation (1). Several of the Leddar segments possessed “noisy” (Flag = 35) data records, which were removed prior to calculating the means.

$$D_{\text{snow}} = V_o - V_T \quad (2)$$

Aggregated datasets were created using quality-controlled data for all sensors to reduce the dataset to a manageable size. A daily timestep was created using the 24-hour mean for all variables. To maintain a higher temporal resolution while retaining diurnal variability, a second data file contained records for the mean of the 15-minute data at a one-hour timestep. The dataset contained 3159 one-hour observations. The “proportion of clean returns” is the total number of one-hour observations containing “Flag = 1” (clean) data relative to the “total number of observations” for the timestep. The calculated percentage is the “total observations” divided by the total number of “clean” observations. The “proportion of noisy returns” is the total number of one-hour observations divided by the total number of “Flag = 35” (noise) one-hour observations.

2.4. Controls on the Leddar Intensity Signal

To investigate the drivers of the Leddar intensity signal through time, a Pearson’s correlation matrix was computed for the hourly dataset to examine relationships and potential collinearity between variables. Variables selected were wind speed (WS), daily albedo, air temperature, relative humidity (RH), ground

temperature, SR50A snow depth, hourly precipitation, the Leddar V_T , intensity, and proportion of clean returns per timestep. Univariate statistical analysis was completed for the Leddar intensity and proportion of clean return signals for air temperature, daily albedo, V_T , and relative humidity after being identified as “variables of interest” in the correlation matrix. The correlation analysis was limited to only sample periods when complete snow cover was present beneath the Leddar unit (21 December 2017 to 27 April 2018).

2.5. In-Situ Evaluation of LeddarTech IS16 Sensor’s Precision, Accuracy, and Performance

An initial test of the Leddar sensor precision was completed in the field on 14 December 2017 before the start of the 2017–2018 data collection. A 1.22 m × 2.44 m painted plywood reflective target used by [40] for LIDAR radiometric calibration was placed level on the ground to compensate for the sloped surface under the Leddar sensor to increase the intensity of the return signal. Leddar DTT and intensity data as well as SR50A DTT measurements were captured in 15-minute increments for a four-hour time period. The SR50A and the Leddar segment 16 beams did not fall fully within the target surface due to the sensor field of view being larger than the target dimension. The Leddar beam length was 2.61 m, which exceeded the painted target’s length of 2.44 m. Leddar status “Flag = 1” observations were used to calculate the mean distance to target and intensity for each of the 16 segments.

A second analysis was completed to evaluate the precision of the Leddar unit when operating over a snowpack surface under completely stable conditions; i.e., during a period when the snowpack was in a stable state such that settling, compaction, and crystallization had no significant impact on sensor observations. Several criteria were used to select a time period to test for consistent measurement in snow depth once snowpack settling had taken place. Influences from solar radiation were removed by selecting a sampling interval between 20:00 to 06:00. The air temperature was less than -5.0 °C prior to and throughout the sampling interval to avoid the melting and refreezing metamorphoses of crystalline structures [14, 16]. No precipitation event occurred for several days prior to observation to remove accumulation, compaction, and settling influences [21]. Wind speeds approaching zero were desired to remove redistribution of the snowpack surface [15]. Due to an unusually warm winter with few days between precipitation events, there was only one time period from 26 December 2017 to 27 December 2017 meeting these criteria.

3. Results and Discussion

The Leddar daily data were used to plot both snowpack depth and intensity for the winter 2017–2018 snow season. Figure 4 shows snowpack depth variability across the 16 segments of the sensor footprint, while increases in depth through time represent the snow fall accumulation events. Settling, compaction, or a mid winter melt are seen as decreases in depth over time. Snowpack surface morphology and texture within the sensor field of view are illustrated orthogonal to the time axis. The rapid reduction in snow depth at the end of the time series was due to spring melting as air temperatures and day length increased in mid to late April.

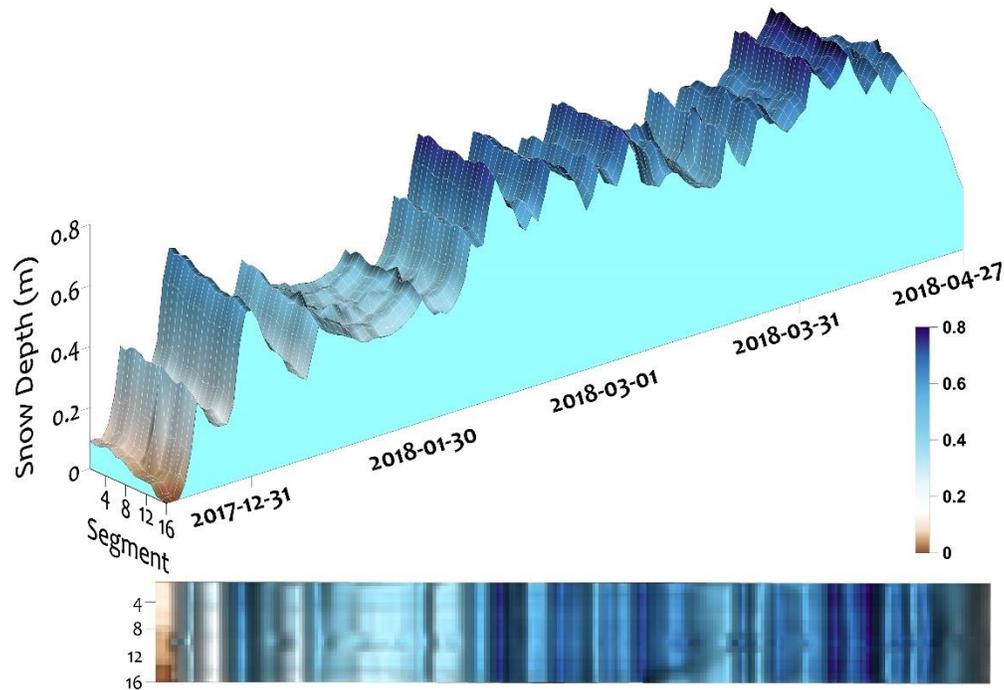


Figure 4. Leddar snow depth 15 December 2017 to 27 April 2018.

The Leddar intensity for the 16 segments is shown in Figure 5. At the beginning and end of the series, low intensities are associated with patchy snow combined with bare soil and vegetation at the ground surface level. Higher intensity values occur, when snow cover completely fills the sensor field of view. During the period of 100% snow covered area (SCA), intensity increases immediately following snow accumulation events and then gradually decreases as the snowpack settles and metamorphoses. These patterns of increasing and decreasing intensity are synchronous with increases and decreases in depth, but the magnitudes of depth and intensity change are not visually correlated. Intensity responses are strongest at nadir segments with decreasing values toward the outer edge of the field of view. The laser radar equation shows there is an inverse relationship with DDT and the intensity response [32, 41]. Segments at nadir are closest to the ground surface and have the highest intensity. For each segment starting at nadir going to the edge segments of the sensor, distance to the ground increases and there is a corresponding decrease in intensity.

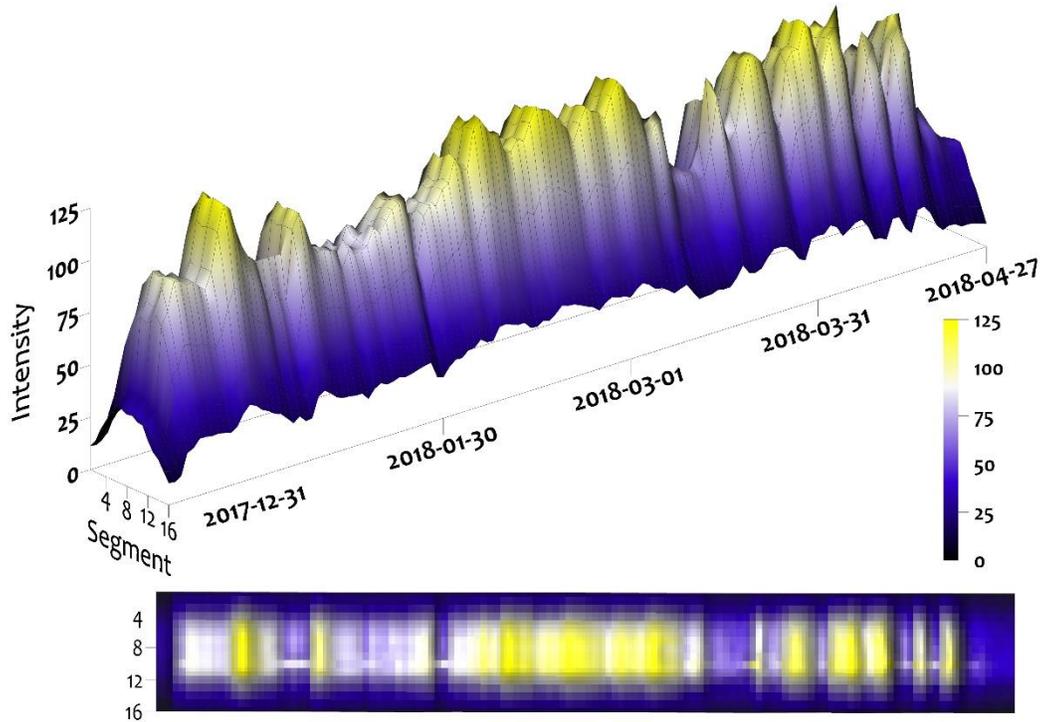


Figure 5. Leddar intensity return signal from 15 December 2017 to 27 April 2018. Before the implementation of all of the quality control steps.

3.1. LeddarTech IS16 Sensor Performance

3.1.1. Signal data noise

The Leddar DTT measurement is dependent on the “clean” (“Status Flag = 1”) return signals and intensity amplitude. Due to the Leddar demerging configuration setting being “enabled,” the sensor occasionally recorded signals with a “Flag = 35” status (noisy returns). Analysis was completed for the entire hourly time series (14 December 2017 to 27 April 2018) dataset to evaluate the amounts of “clean” (proportion of clean returns) and “noisy” (proportion of noisy returns) return signals for all segments (Table 3a). The proportion of timestep observations with no clean returns (Table 3b) is the amount of missing clean data for the individual 15-minute data sampling measurements. Intensity (Table 3c) for segments near nadir received the highest values, while edge segments received the lowest intensity measurements. Of note, the two edge segments on both sides of nadir had minimum intensity values below the manufacturer’s specifications for DTT accuracy threshold which occurred during the time period of no-snow. Segment 10 had an abnormally high noise fraction throughout most of the observation period. This segment does not appear to be representative of the sensor. The noise was potentially caused by internal damage or contamination within the optical receiver.

Table 3. From 14 December 2017 to 27 April 2018, n = 3211 (1-hour timestep) per segment (SEG) for the entire data collection period. (a) Leddar data series proportion of clean and noise returns. (b) The proportion of timestep observations with no clean returns refers to the amount of data for all 15-minute measurement cycles where the sensor did not detect any clean readings. (c) Leddar intensity signal minimum, maximum, mean, and range.

| SEG | Proportion of Clean returns | Proportion of Noisy returns | Proportion of timestep observations with no Clean returns | Intensity | | | |
|-----|-----------------------------|-----------------------------|---|-----------|-------|-------|-------|
| | | | | Min | Max | Mean | Range |
| 1 | 99.2% | 0.8% | 0.0% | 10.0 | 45.2 | 29.3 | 35.2 |
| 2 | 99.7% | 0.3% | 0.0% | 13.7 | 61.6 | 40.8 | 47.9 |
| 3 | 97.5% | 2.5% | 0.0% | 18.6 | 82.0 | 55.0 | 63.4 |
| 4 | 96.4% | 3.6% | 0.1% | 24.8 | 105.1 | 70.9 | 80.3 |
| 5 | 94.6% | 5.4% | 0.1% | 26.0 | 123.2 | 83.4 | 97.2 |
| 6 | 91.1% | 8.9% | 2.1% | 30.3 | 128.0 | 87.8 | 97.6 |
| 7 | 88.9% | 11.1% | 2.2% | 32.1 | 134.0 | 92.6 | 101.9 |
| 8 | 84.9% | 15.1% | 5.6% | 31.6 | 131.7 | 92.7 | 100.1 |
| 9 | 81.7% | 18.3% | 5.5% | 31.3 | 134.5 | 94.8 | 103.1 |
| 10 | 33.4% | 66.6% | 38.5% | 29.2 | 135.3 | 105.1 | 106.1 |
| 11 | 87.5% | 12.5% | 1.2% | 29.3 | 136.1 | 94.3 | 106.8 |
| 12 | 84.8% | 15.2% | 5.9% | 24.6 | 118.8 | 83.9 | 94.2 |
| 13 | 88.4% | 11.6% | 0.3% | 20.4 | 96.1 | 66.3 | 75.7 |
| 14 | 97.0% | 3.0% | 0.0% | 17.6 | 75.6 | 51.9 | 58.0 |
| 15 | 98.5% | 1.5% | 0.4% | 12.9 | 57.8 | 39.7 | 44.8 |
| 16 | 96.9% | 3.1% | 1.4% | 7.7 | 38.3 | 26.0 | 30.6 |

The noise was plotted over time to find potential relationships with snowpack conditions and meteorological influences. Figure 6 shows the proportion of the data per time step that received noisy returns in relation to the total number of returns. The least amount of noise was observed in the middle of the series once the SCA was fully present at the end of December and before melt conditions initially occurred in mid-March. Segments near nadir experienced more noise during episodic snowmelt periods late in the winter season, when there was higher water content in the snowpack. Some noise may have been caused by solar contamination, precipitation, wind redistribution of snowpack surface grains, or higher moisture content when the pack entered a freeze/thaw stage or became isothermal and melted out.

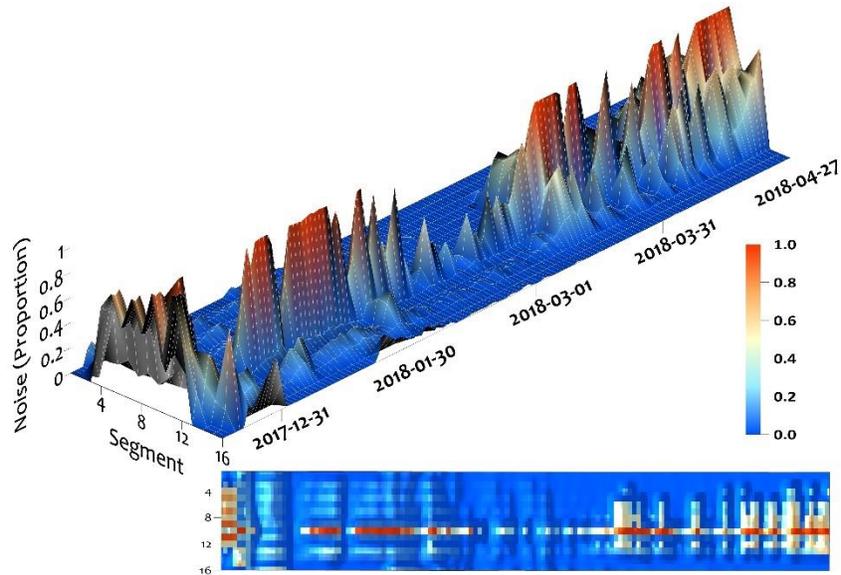


Figure 6. Leddar “noise” in the data from 14 December 2017 to 27 April 2018. The noise is the proportion of the data per time step that received noisy returns in relation to the total number of returns. Time series of readings culminated in Table 3a—proportion of noisy returns.

3.1.2. Temperature Sensitivity

The high-resolution time series data demonstrated the intensity amplitude shifted by 10 or more units immediately before some snowfall events (Figure 7). This indicated at least some of the shift in signal intensity associated with snowpack accumulation events was not a function of the snowpack surface condition but instead was impacted by the inverse relationship between LED light output power and ambient temperature [42]. Figure 7 show five events (precipitation: blue bars; air temperature: black; intensity: red) with increases and decreases in the intensity response prior to and during the events. These data show that signal intensity tends to vary inversely with air temperature leading into the precipitation event, and is likely a hardware response demonstrating some thermal sensitivity within the Leddar unit. Consultation with the manufacturer confirmed both internal hardware components and environmental conditions influence the sensitivity of the Leddar intensity signal [28]. Further analysis is required to separate out the ambient temperature influence on the hardware vs. the surface reflectance response; however, this preliminary illustration of an inverse relationship with temperature suggests that temperature-based correction of the intensity response, while not implemented here, could be achievable.

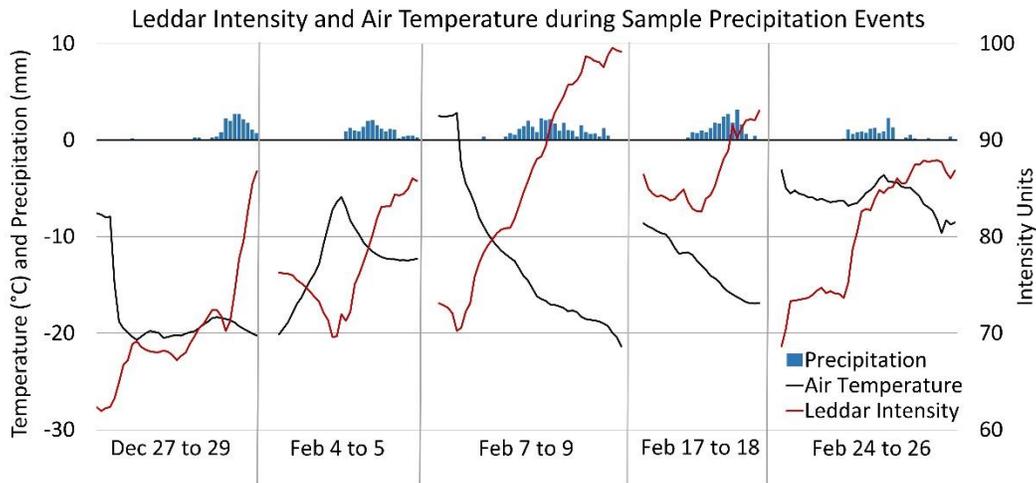


Figure 7. Leddar intensity (mean all segments), air temperature, sample precipitation events. Initial examination of the Leddar intensity signal suggests an LED light output sensitivity to air temperature.

3.1.3. Controls on the Leddar Intensity Signal

The full winter season 1-hour dataset with Leddar intensity signal, proportion of clean returns, air temperature, daily albedo, and Leddar V_T data show moderate to strong correlations (Table 4).

Table 4. Pearson’s Correlation coefficient (r) for the Leddar (mean of all segments) sensor and other meteorological variables for the given timestep. “% Clean” is the proportion of clean returns.

| | Daily Albedo | Air Temperature | Leddar V_T | Leddar Intensity |
|-------------------------|--------------|-----------------|--------------|------------------|
| | r | r | r | r |
| Leddar Intensity | 0.77 | -0.77 | -0.43 | - |
| Leddar % Clean | 0.59 | -0.57 | -0.13 | 0.74 |

The daily albedo and Leddar intensity return signal show a correlation of 0.77. There should be a relationship between these two variables as they are both a measure of reflectance. Higher reflectance values are received from fresh dry snow when smaller snow grain sizes are present [43]. Lower reflectance values occur as the snowpack metamorphoses as a result of changes in crystal structure and increased snow grain size [22, 44]. Reflectance of aged, wet, and melting snow is lower than that of fresh, dry snow [21, 45]. Ice layers formed from melt/freeze cycles have similar reflectance properties to that of wet snow [12, 19]. Albedo and the proportion of clean Leddar returns showed a moderate correlation of 0.59, which indicates that noise tends to diminish as snowpack reflectance increases.

There is a correlation of -0.77 between the air temperature and the Leddar intensity signal (Table 4). In addition to potential hardware influences (discussed above), air temperature also influences the snowpack properties of grain size and structure [16]. As air temperatures increase and approach $0\text{ }^\circ\text{C}$, the snowpack begins to ripen (higher water content) and surface reflectance reduces [46].

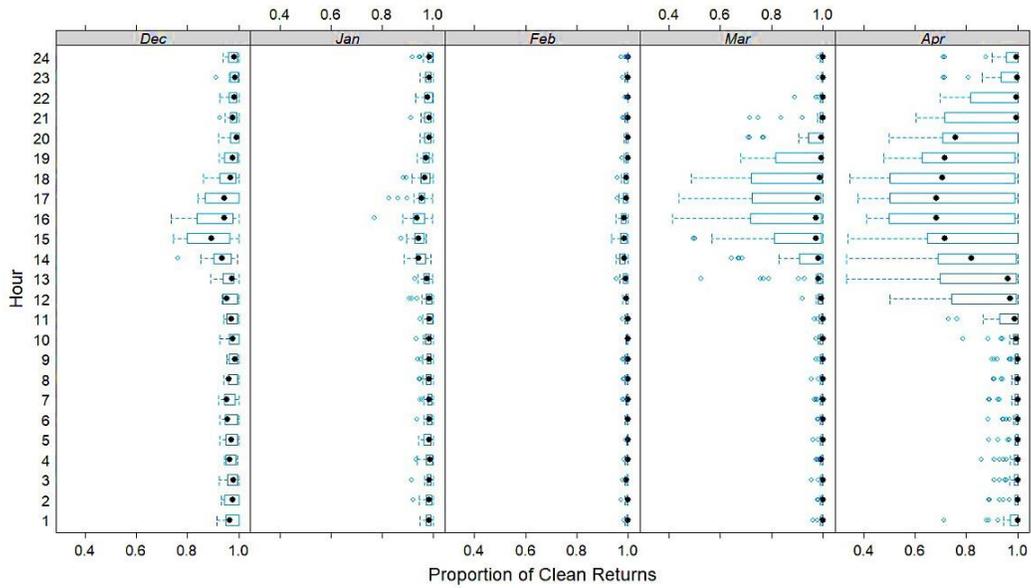
There is negative correlation between air temperature and the proportion of clean Leddar returns at -0.57 (Table 4). The proportion of clean returns was elevated during cooler temperatures. As air temperature approached $0\text{ }^\circ\text{C}$ and warmer, more noise was present in the data. This could be influenced

by another driver that auto-correlates with seasonal variations in temperature, such as freeze/thaw conditions at the surface of the pack during the melt phase, suggesting noise levels increase as the snow surface melts.

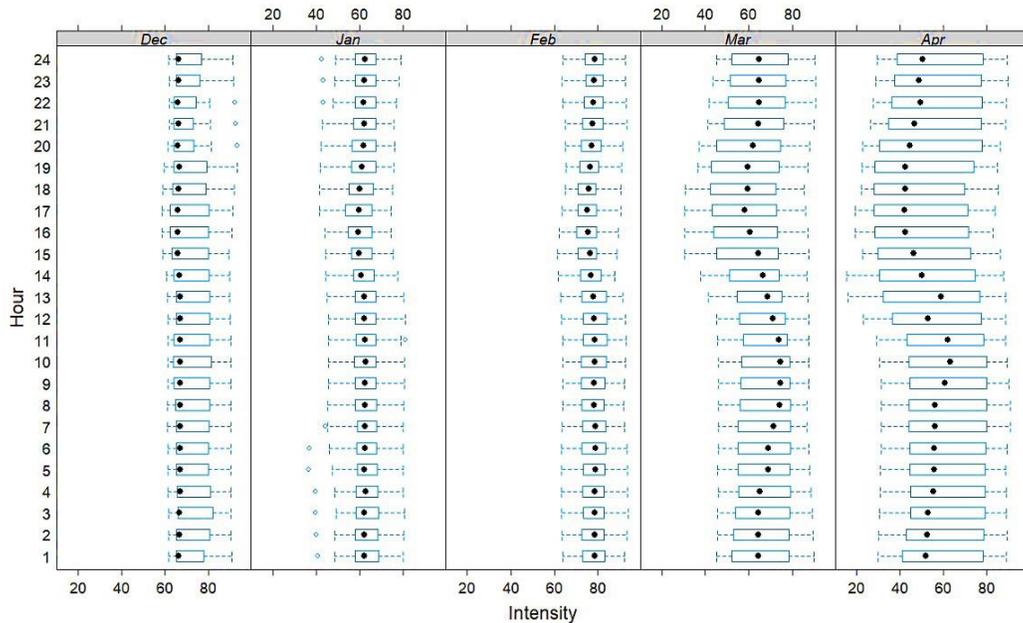
No relationship exists for V_T and the clean return signal ($r = -0.13$), but V_T and intensity demonstrate a correlation of -0.43 . The further the target is from the sensor, the less backscatter received [47]. Since the Leddar unit was statically mounted on a tower facing the ground surface, the largest DDT values occur before the snowpack is present. At this point, the sensor is observing the vegetation and soils which have a lower spectral reflectance than snow in the NIR 940 nm band. As the snowpack builds, DDT is reduced, and Leddar intensity increases due to the higher albedo of snow, as well as the shorter range. As the pack starts to melt out, spectral reflectance values decrease from the higher water content in the snowpack as well as the increased distance to the ground surface.

Intensity and proportion of clean returns are positively correlated ($r = 0.74$, Table 4). The intensity signal and proportion of clean returns were plotted through time once the snowpack emerged. Box plots in Figure 8 are broken down by hour-of-the-day and further by month. Both the proportion of clean returns (Figure 8a) and the intensity signal (Figure 8b) show diurnal patterns. Signal “noise” and intensity ranges are elevated during daylight hours from midday onwards, which appear to correspond with temperature and daylight variations. Maximum noise and lowest intensity are associated with warmer afternoon temperatures, while minimal noise and slightly elevated Leddar reflectance occur in the morning after sunrise but during local diffuse sky radiation. The pattern was more dominant in March and April, when most daytime air temperatures were above 0 °C. The least amount of noise occurred in February when air temperatures were between -20 to -35 °C with only a few days approaching or above 0 °C. There are two potential contributing factors: (i) occasional melt, ripening processes, and higher water content at the surface of the snowpack; and/or (ii) solar contamination of the signal during the afternoon. With the diurnal variation in sensible heat flux and net surface radiation balance over the snowpack, it is possible the Leddar signal is sensitive to the changing snowpack surface structure and increasing snow grain size. There is a corresponding increase in noise levels in March and April when diurnal energy inputs to the pack can be most extreme. Any sensitivity to changes in surface structure, would be expected to be observed during the onset of melt conditions late in the season.

Solar contamination cannot be ruled out, however, as the skyview surrounding the instrument tower is most open to the west with mountain ridges dominating the south and north skyline. Consequently, if solar contamination occurs, it would be expected in the afternoon, with the level of contamination increasing as the solar zenith and range in azimuth increase later in the season. It is believed that both surface freeze/thaw processes and solar contamination play a role in influencing diurnal intensity and noise patterns, but further investigation is required to quantify the relative influence and identify which is dominant.



(a)



(b)

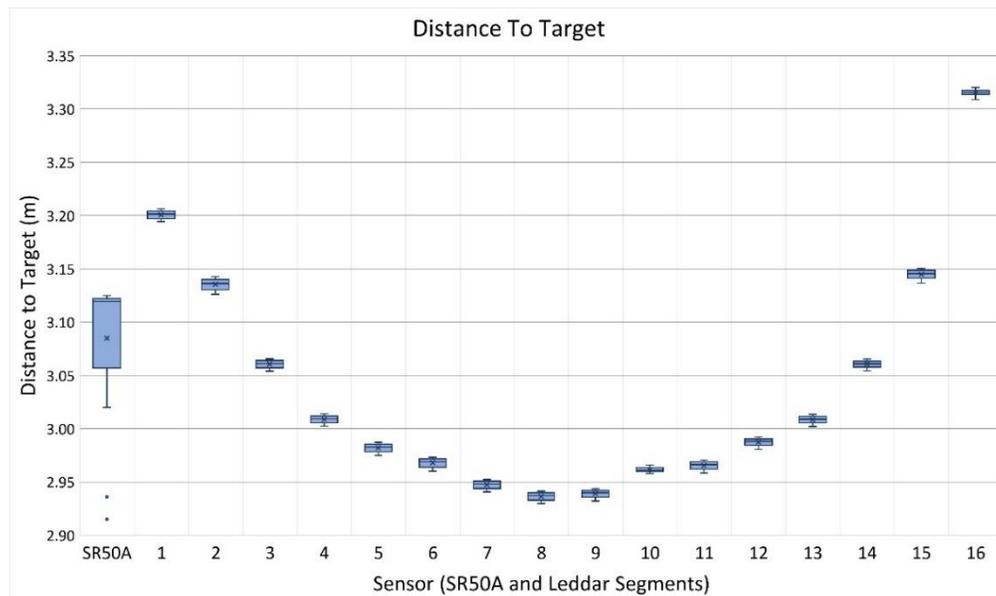
Figure 8. Proportion of clean signals (a) and mean hourly Leddar signal intensity (b) by month from December to April. The box delineates the lower 25th and upper 75th percentiles. The dashed T-lines are the minimum and maximum values and open circles to the left and right of the dashed T-line are outliers. The black dot inside the box is the mean.

3.2. Range and Depth Observations

3.2.1. Leddar Calibration

The root-mean-square-error (RMSE) of range-to- V_T calibration conducted in the laboratory using Equation (1) was 13.0 mm. The parameter values (and standard error): β , 3.134° (0.010°); γ was -0.044° (0.003°); and R_c was -0.307 m (< 0.001 m). The calibrated β was just over the LeddarTech product specification of 3.0° [24].

Results of the Leddar and SR50A reflective target testing from December 14, 2017 (as described in Section 2.5), are shown in Figure 9. The Leddar DTT follows a systematic arc from the optical receiver to the target surface [24]. Segment DTT is shortest at nadir and increases toward the edge segments. The range of DTT values show consistent measurements for all segments within 20 mm and a maximum standard deviation for all segments of 5 mm (Figure 9a) over a four-hour duration. The SR50A measurement varied to a higher degree with a DTT range of 210 mm and a standard deviation of 70 mm. Under controlled conditions, variation in the Leddar measurements was very low and more precise than the SR50A. The intensity signal (Figure 9b) follows an expected systematic inverse pattern with highest values at nadir and decreasing towards the outer segments. The footprint for segment 16 was at the edge of the target and was contaminated with bare soil and grass returns, resulting in the lowest intensity relative to other segments. The intensity measurement over the four-hour period for segment 16 was 11.1 which was below the Leddar manufacturer’s threshold of 15. However, the DTT standard deviation for segment 16 was 3 mm. The temporal stability of measurements from this Leddar sensor testing are consistent with the findings of [48] in their preliminary investigation of the Leddar technology.



(a)

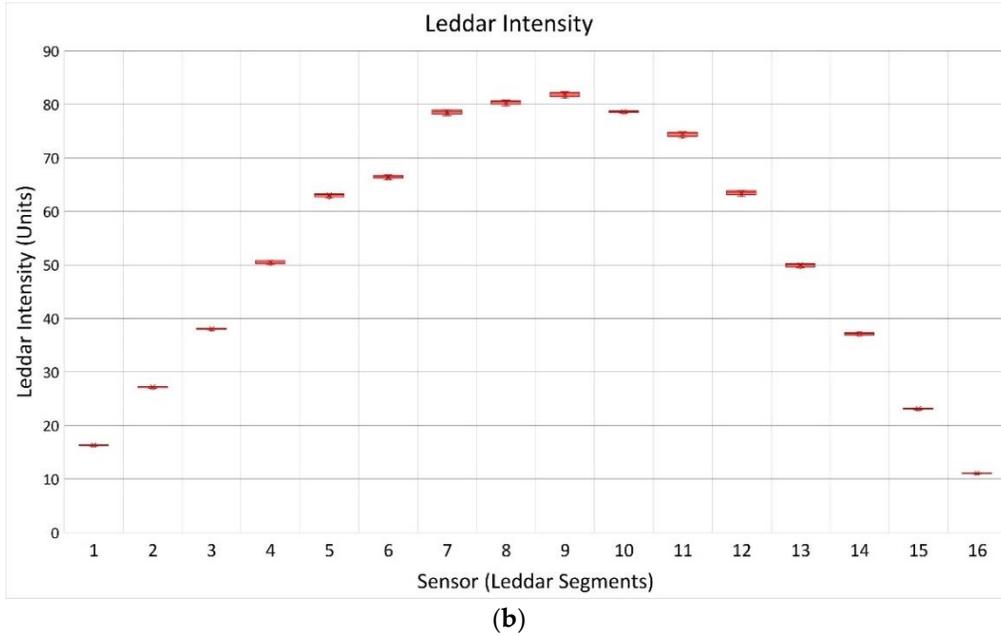


Figure 9. Reflective target placed on the ground under the Leddar unit. SR50A footprint and Leddar Segment 16 were not contained within the target surface. (a) Boxplot of Leddar and SR50A DTT, (b) boxplot of Leddar Intensity. (Note: SR50A does not return intensity signal measurements).

3.2.2. Snow Depth Validation

The full winter season 1-hour Leddar and SR50A depth dataset contained 3211 records. Table 5 shows snowpack depth for the SR50A and each Leddar segment when the snowpack reached maximum depth on 8 April 2018 at 10:00. Snow depth varies across the Leddar field of view with lowest measurements of 0.75 m in segments 5 through 7 and largest observations of 0.85 m for segments 14–16. This difference is easily within the expected variance associated with the slightly sloping ground surface and undulations in the snowpack surface. The snow depth recorded for the SR50A was 0.83 m. The SR50A measurement falls within the range of the Leddar snow depth for segments 13 and 14. SR50A observations are not necessary expected to be the same as any individual Leddar segment due to the size and shape of the different sampling footprints (Figure 1).

Table 5. Maximum snowpack depth occurred on 8 April 2018 10:00. Leddar segments and SR50A snow depth. From 14 December 2017 to 27 April 2018 winter season 1-hour dataset, n = 3211.

| SR50A | Leddar Segment (m) | | | | | | | | | | | | | | | |
|-------|--------------------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| (m) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| 0.83 | 0.81 | 0.79 | 0.80 | 0.77 | 0.75 | 0.75 | 0.75 | 0.76 | 0.77 | 0.75 | 0.78 | 0.78 | 0.81 | 0.85 | 0.85 | 0.85 |

Comparable Leddar, SR50A, and field validation data are presented in Table 6. The Leddar “Proportion clean returns” reduces as air temperatures increase during melt conditions. The Leddar footprint shows a range in snow depth from a low of 0.08 m to a high of 0.17 m across all segments. The

SR50A snow depth measurement was between zero to 0.04 m for all site visits except the last when significant melt occurred in a short period of time.

Table 6. Manual field measurement, SR50A, Leddar (all 16 segments), mean air temperature, and daily albedo. Manual snow depth measurements taken from behind the tower out of the sensor footprints under the Leddar unit, but exact location varied with each visit. All data collected on calm days with no precipitation.

| Site Visit Date | 21- Dec | 05- Jan | 20- Jan | 21- Feb | 04- Mar | 09- Apr | 15- Apr | 19- Apr | 26- Apr | 27- Apr | 27- Apr |
|--------------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| Start Time | 18:30 | 11:00 | 15:00 | 10:45 | 14:45 | 14:00 | 14:00 | 18:15 | 06:00 | 07:00 | 11:15 |
| End Time | 22:30 | 15:00 | 19:00 | 14:45 | 18:45 | 18:00 | 18:00 | 21:45 | 10:00 | 11:00 | 12:30 |
| Snow Depth (m) | | | | | | | | | | | |
| Manual Field Sample | 0.35 | 0.40 | 0.36 | 0.70 | 0.74 | 0.77 | 0.53 | 0.58 | 0.28 | 0.20 | 0.17 |
| SR50A Mean | 0.31 | 0.44 | 0.36 | 0.66 | 0.72 | 0.71 | 0.59 | 0.57 | 0.29 | 0.21 | 0.17 |
| SR50A Min | 0.29 | 0.42 | 0.35 | 0.65 | 0.72 | 0.70 | 0.59 | 0.56 | 0.28 | 0.20 | 0.06 |
| SR50A Max | 0.31 | 0.45 | 0.36 | 0.69 | 0.72 | 0.73 | 0.60 | 0.58 | 0.29 | 0.22 | 0.20 |
| Leddar Mean | 0.28 | 0.36 | 0.31 | 0.61 | 0.66 | 0.65 | 0.54 | 0.53 | 0.23 | 0.14 | 0.12 |
| Leddar Min | 0.24 | 0.32 | 0.27 | 0.54 | 0.61 | 0.59 | 0.48 | 0.46 | 0.16 | 0.08 | 0.06 |
| Leddar Max | 0.38 | 0.40 | 0.37 | 0.70 | 0.74 | 0.75 | 0.62 | 0.61 | 0.32 | 0.24 | 0.21 |
| Proportion clean returns | 92% | 83% | 81% | 93% | 93% | 59% | 47% | 49% | 79% | 63% | 28% |
| Air Temperature (°C) | -6.2 | 3.2 | -0.3 | -13.4 | -7.4 | 7.4 | 7.5 | 5.0 | 4.8 | 5.1 | 16.8 |
| Daily Albedo | 0.84 | 0.83 | 0.80 | 0.81 | 0.85 | 0.77 | 0.69 | 0.73 | 0.59 | 0.58 | 0.58 |

A regression plot of Leddar and SR50A snow depth (Figure 10a) shows strong agreement between the two sensors ($r^2 = 0.98$). Figure 10b shows the mean for the Leddar snowpack depth values across all segments plotted against the single field snow depth measurement below the Leddar unit. There is low confidence that the manually collected field data represent the whole Leddar or SR50A sampling footprint. Single manual sampling depth observations were obtained from the back of the tower outside the sensor footprints reaching approximately 1 m to the location under the Leddar unit attempting to minimize snow surface disturbance. The manual measurements fall within the Leddar segment array distribution range as seen in Table 6 but were not co-located to any individual segment. The Leddar segment means tend to fall below the field measurement resulting in a bias on the order of ~ 0.05 m due to variability in snow depth among all segments or manual field measurement errors such as the probe not penetrating to the ground surface from ice layers at the base of the pack. However, despite this small bias Leddar and field depth data show a strong regression ($r^2 = 0.98$).

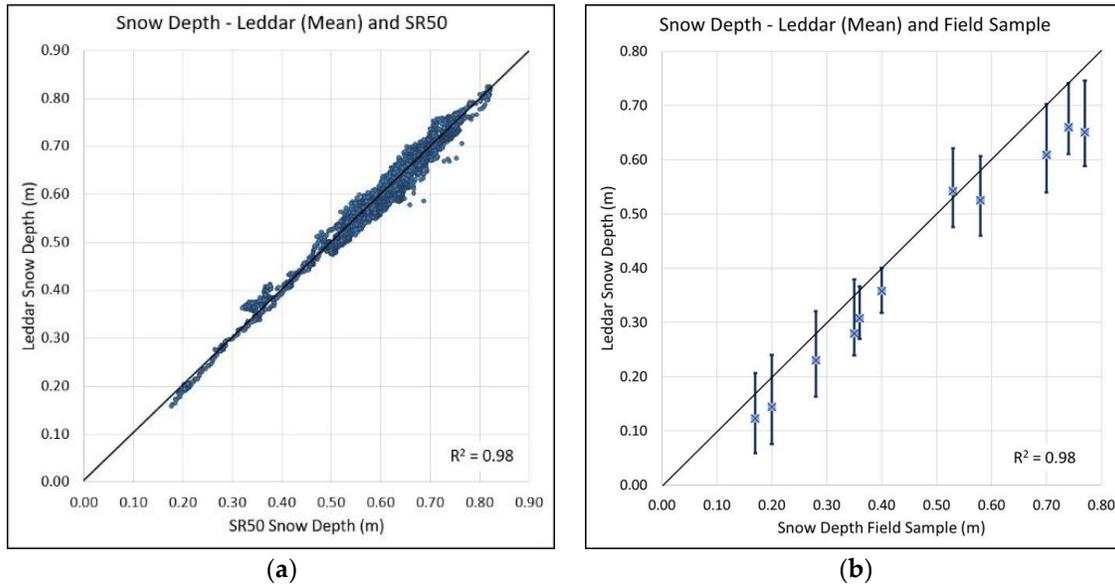


Figure 10. (a) Plot of Leddar segment mean against SR50A mean snowpack depth. (b) Plot of Leddar segment depth range (bars) and mean (blue square) against manual field measurements.

3.2.3. Leddar Snow Depth Stability

Table 7. Leddar and SR50A sensor snow depth data from 20:00 26 December 2017 to 06:00 27 December 2017 nighttime sampling period meeting the minimum specifications.

| Leddar Segment | Snow Depth (m) | | | | |
|------------------|----------------|-------|-------|-------|-------|
| | Mean | STDev | Min | Max | Range |
| 1 | 0.191 | 0.001 | 0.190 | 0.194 | 0.004 |
| 2 | 0.181 | 0.001 | 0.180 | 0.183 | 0.003 |
| 3 | 0.189 | 0.001 | 0.188 | 0.190 | 0.002 |
| 4 | 0.179 | 0.001 | 0.178 | 0.181 | 0.003 |
| 5 | 0.157 | 0.001 | 0.156 | 0.158 | 0.002 |
| 6 | 0.162 | 0.001 | 0.161 | 0.164 | 0.003 |
| 7 | 0.161 | 0.001 | 0.160 | 0.163 | 0.003 |
| 8 | 0.174 | 0.001 | 0.173 | 0.176 | 0.003 |
| 9 | 0.179 | 0.001 | 0.178 | 0.182 | 0.003 |
| 10 | 0.169 | 0.001 | 0.168 | 0.172 | 0.003 |
| 11 | 0.173 | 0.001 | 0.172 | 0.175 | 0.003 |
| 12 | 0.176 | 0.001 | 0.175 | 0.178 | 0.003 |
| 13 | 0.191 | 0.001 | 0.189 | 0.193 | 0.003 |
| 14 | 0.212 | 0.001 | 0.211 | 0.215 | 0.004 |
| 15 | 0.197 | 0.002 | 0.195 | 0.200 | 0.005 |
| 16 | 0.211 | 0.001 | 0.210 | 0.214 | 0.004 |
| SR50A (m) | 0.174 | 0.008 | 0.164 | 0.185 | 0.021 |

The night-time sampling interval that met the stable sampling constraints of air temperatures consistently below -5°C with no precipitation or wind, was a 10-hour period between 20:00 on 26 December 2017 to 06:00 27 December 2017. Table 7 shows snow depth for the Leddar segments and SR50A. The range in snow depth measurement of each individual Leddar segments was <5 mm, while the range

in SR50A snow depth over the single footprint was >20 mm. The precision observed for the Leddar sensor exceeds the manufacturer quoted 6 mm when intensity is greater than 15 [24].

4. Conclusions

This proof of concept study examined the feasibility of using a LeddarTech IS16 LED LIDAR-based sensor for snowpack surface morphology, depth, and reflectance measurements. The sensor was co-located and tested against a Campbell Scientific SR50A sonic ranging snow depth sensor during the 2017–2018 winter season. Controls on the intensity signal, comparison to SR50A snow depths, Leddar distance-to-target (DTT) precision, and accuracy were evaluated.

The Leddar sensor data showed variability in magnitude of the intensity signal during the formation, settling, ripening, and ablation of the snowpack as well as when precipitation events took place. Prior to snowpack formation, vegetation and soils produced more noise in the dataset as well as lower intensity responses. Lower intensity and increased noise also occurred during surface melt conditions. More noise was present in segments closer to nadir than at the edges of the Leddar footprint. Diurnal variations in the intensity and noise levels increased later in the season, suggesting that solar contamination of the Leddar signal could be present and that mitigation measures need to be explored. There was a high correlation between albedo and the intensity signal (0.77) as well as noise (0.59). Air temperature demonstrated a negative correlation (−0.77) with intensity, which was likely at least partly a function of hardware sensitivity, though it could also be influenced by surface melt lowering the intensity response. Further analysis is required to quantify and better understand the impact of air temperature on the intensity signal.

The Leddar sensor data showed variability in snow depth measurements across the 16-segment footprint during snowpack accumulation, compaction, and ablation events. Examining sensor performance when the snowpack was in a stable state showed Leddar snow depth observations varied <5 mm in comparison to >20 mm for the SR50A measurements. These Leddar observations exceed the manufacturer’s specification for accuracy at 50 mm and precision of 6 mm for intensity values greater than 15. Despite the environmental sensitivities experienced by the Leddar unit, the sensor demonstrated accurate and precise depth measurements throughout the winter season. The Leddar and SR50A snow depth measurements displayed a strong linear relationship ($r^2 = 0.98$), with virtually no bias and a slope of 0.99. Leddar and manual field measurements also showed strong agreement ($r^2 = 0.98$). Leddar measurements during initial testing using a reflective target were within 20 mm for DTT, compared to 210 mm for the SR50A sensor.

LED LIDAR has great potential for low-cost monitoring of a range of snowpack conditions, including depth measurements for 16 independent beam segments, surface morphology, and possibly surface reflectance if temperature influences on the internal optics can be mitigated. The base components used in this study were obtained for a combined approximate cost of <CA\$1000, which is similar to a new sonic ranging sensor. Meanwhile, *in-situ* Leddar snow depth observations can be made at higher precision and accuracy, over multiple segments along a transect, while the intensity signal and noise response show promise for either characterising snow surface properties and/or perhaps precipitation. Compared to traditional laser-based ranging sensors, LED LIDAR has a low intermittent operational power requirement, making it a suitable candidate for remote deployment where power options are limited.

Author Contributions: Conceptualization, C.H. and C.B.; methodology, C.B.; software, Z.X.; validation, C.H.; formal analysis, C.B.; investigation, C.B., C.H., T.P., and Z.X.; resources, C.H.; data curation, C.B., T.P. and Z.X.; writing—original draft preparation, C.B.; writing—review and editing, C.B. and C.H.; supervision, C.H.; project administration, C.H.; funding acquisition, C.H.

Funding: This research was funded through grants awarded to Hopkinson: (i) NSERC ENGAGE program; (ii) Alberta Innovates Energy and Environmental Solutions (AIEES)—Water Innovation Program; and (iii) the Government of Alberta Environment and Parks. Laboratory equipment funding was provided through the Canada Foundation for Innovation and Western Economic Diversification Canada.

Acknowledgments: Celeste Barnes would like to thank Dan Johnson from the University of Lethbridge for comments; Reed Parsons for his preliminary work with the LeddarTech IS16 sensor; the staff at LeddarTech are acknowledged for providing early versions of the IS16 hardware at no cost for preliminary testing purposes, for answering hardware related questions and occasional troubleshooting support.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Natural Resources Canada. *Atlas of Canada, Average Maximum Snow Depth*, 6th ed; N.R.C. Government of Canada, Earth Sciences Sector, Atlas of Canada, Ed.; Government of Canada, Natural Resources Canada, Earth Sciences Sector, Atlas of Canada: Ottawa, ON, Canada, 2010.
2. Barrows, H.K.; Horton, R.E. *Determination of Stream Flow During the Frozen Season*; US Government Printing Office: Washington, DC, USA, 1907.
3. Paznekas, A.; Hayashi, M. Groundwater Contribution to Winter Streamflow in the Canadian Rockies. *Can. Water Resour. J. Rev. Can. Des Ressources. Hydr.* **2016**, *41*, 484–499.
4. Grünewald, T.; Schirmer, M.; Mott, R.; Lehning, M. Spatial and temporal variability of snow depth and SWE in a small mountain catchment. *Cryosphere* **2010**, *4*, 215–225.
5. Government of Alberta. *Snow Does Not Equal Flood: Alberta's Snowmelt and its Impact on Alberta Rivers*; alberta.ca: Edmonton, AB, Canada, 2018.
6. Kinar, N.J.; Pomeroy, J.W. Measurement of the physical properties of the snowpack. *Rev. Geophys.* **2015**, *53*, 481–544.
7. López-Moreno, J.I.; Fassnacht, S.R.; Heath, J.T.; Musselman, K.N.; Revuelto, J.; Latron, J.; Morán-Tejeda, E.; Jonas, T. Small scale spatial variability of snow density and depth over complex alpine terrain: Implications for estimating snow water equivalent. *Adv. Water Resour.* **2013**, *55*, 40–52.
8. Pickering, R. *Snow Surveying Field Sampling Protocols*; Enrollment Management and Services Division, Ed.; Alberta Environment and Parks: Edmonton, AB, Canada, 2018; p. 10.
9. Campbell, S. *Acoustic Snow Depth Sensor Operation*; Enrollment Management and Services Division, Ed.; Alberta Environment and Parks: Edmonton, AB, Canada, 2018; p. 7.
10. Gubler, H. An inexpensive remote snow-depth gauge based on ultrasonic wave reflection from the snow surface. *J. Glaciol.* **1981**, *27*, 157–163.
11. Ryan, W.A.; Doesken, N.J.; Fassnacht, S.R. Evaluation of ultrasonic snow depth sensors for u.s. snow measurements. *J. Atmos. Ocean. Technol.* **2008**, *25*, 667–684.
12. Warren, S.G. Optical properties of ice and snow. *Philos. Trans. R. Soc. A* **2019**, *377*, 20180161.
13. Fischer, A.P. The measurement factors in estimating snowfall derived from snow cover surfaces using acoustic snow depth sensors. *J. Appl. Meteorol. Clim.* **2011**, *50*, 681–699.
14. Pomeroy, J.W.; Gray, D.M.; Shook, K.R.; Toth, B.; Essery, R.L.H.; Pietroniro, A.; Hedstrom, N. An evaluation of snow accumulation and ablation processes for land surface modelling. *Hydrol. Process.* **1998**, *12*, 2339–2367.
15. Essery, R.; Li, L.; Pomeroy, J. A distributed model of blowing snow over complex terrain. *Hydrol. Process.* **1999**, *13*, 2423–2438.

16. Warren, S.G. Optical properties of snow. *Rev. Geophys.* **1982**, *20*, 67–89.
17. Deems, J.S.; Painter, T.H.; Finnegan, D.C. Lidar measurement of snow depth: A review. *J. Glaciol.* **2013**, *59*, 467–479.
18. Seidel, F.C.; Rittger, K.; Skiles, S.M.; Molotch, N.P.; Painter, T.H. 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. *Cryosphere* **2016**, *10*, 1229–1244.
19. Hannula, H.-R.; Pulliainen, J. Spectral reflectance behavior of different boreal snow types. *J. Glaciol.* **2019**, *65*, 926–939.
20. Eppanapelli, L.K.; Lintzén, N.; Casselgren, J.; Wåhlin, J. Estimation of liquid water content of snow surface by spectral reflectance. *J. Cold Reg. Eng.* **2018**, *32*, 05018001.
21. Peltoniemi, J.I.; Kaasalainen, S.; Naranen, J.; Matikainen, L.; Piironen, J. Measurement of directional and spectral signatures of light reflectance by snow. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 2294–2304.
22. Negi, H.S.; Singh, S.K.; Kulkarni, A.V.; Semwal, B.S. Field-based spectral reflectance measurements of seasonal snow cover in the Indian Himalaya. *Int. J. Remote Sens.* **2010**, *31*, 2393–2417.
23. Kulkarni, A.V.; Srinivasulu, J.; Manjul, S.S.; Mathur, P. Field based spectral reflectance studies to develop NDSI method for snow cover monitoring. *J. Indian Soc. Remote Sens.* **2002**, *30*, 73–80.
24. LeddarTech Inc.. *Leddar™ Sensor Module User Guide*; 2018, LeddarTech Inc.: Quebec, QC, Canada, 2018; p. 13.
25. Kipp, S.; Mistele, B.; Schmidhalter, U. The performance of active spectral reflectance sensors as influenced by measuring distance, device temperature and light intensity. *Comput. Electron. Agric.* **2014**, *100*, 24–33.
26. Bullough, J. *LED Lighting Systems*, in *National Lighting Product Information Program's (NLPIP's)*; Lighting Research Center: Troy, NY, USA, 2003; p. 47.
27. Olivier, P. *Leddar Optical Time-of-Flight Sensing Technology: A New Approach to Detection and Ranging*; LeddarTech Inc.: Quebec, QC, Canada, 2016; p. 13.
28. Messiaen, P. *LeddarTech IS16 [case #1499]*; Barnes, C., Ed.; LeddarTech Inc.: Quebec, QC, Canada, 2019.
29. Campbell Scientific (Canada) Corp. *SR50A and SR50AT Sonic Ranging Sensors*; Campbell Scientific (Canada) Corp.: Edmonton, AB, Canada, 2016; p. 64.
30. Xi, Z.; Hopkinson, C.; Rood, S.B.; Barnes, C.; Xu, F.; Pearce, D.; Jones, E. A lightweight leddar optical fusion scanning system (FSS) for canopy foliage monitoring. *Sensors* **2019**, *19*, 3943.
31. Xi, Z. *Rasbian Leddar Detection Sampling Computer Program*; University of Lethbridge: Lethbridge, AB, Canada, 2017.
32. Baltsavias, E.P. Airborne laser scanning: Basic relations and formulas. *ISPRS J. Photogramm. Remote Sens.* **1999**, *54*, 199–214.
33. Binder, T.; Kostina, E. Gauss–newton methods for robust parameter estimation. In *Model Based Parameter Estimation*; Springer: Berlin, Germany, 2013; pp. 55–87.
34. Earle, M.; Reverdin, A.; Wolff, M.; Smith, C.; Morin, S.; Rodica, N.; Data processing and quality control methodology for the derivation of reference datasets. In *Project Team and (Reduced) International Organizing Committee for the WMO Solid Precipitation Intercomparison Experiment Final Report*; World Meteorological Organization: Sodankylä, Finland, 2014; p. 77.
35. Kochendorfer, J.; Nitu, R.; Wolff, M.; Mekis, E.; Rasmussen, R.; Baker, B.; Earle, M.E.; Reverdin, A.; Wong, K.; Smith, C.D.; et al. Testing and development of transfer functions for weighing precipitation gauges in WMO-SPICE. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 1437–1452.

36. Leeper, R.D.; Palecki, M.A.; Davis, E. Methods to calculate precipitation from weighing-bucket gauges with redundant depth measurements. *J. Atmos. Ocean. Technol.* **2015**, *32*, 1179–1190.
37. Pan, X.; Yang, D.; Li, Y.; Barr, A.; Helgason, W.; Hayashi, M.; Marsh, P.; Pomeroy, J.; Janowicz, R.J. Bias corrections of precipitation measurements across experimental sites in different ecoclimatic regions of western Canada. *Cryosphere* **2016**, *10*, 2347–2360.
38. Campbell Scientific (Canada) Corp. *CNR1 Net Radiometer*; Campbell Scientific (Canada) Corp.: Edmonton, AB, Canada, 2015; p. 32.
39. Mannstein, H. The interpretation of albedo measurements on a snowcovered slope. *Arch. Meteorol. Geophys. Bioclimatol. Ser. B* **1985**, *36*, 73–81.
40. Okhrimenko, M.; Coburn, C.; Hopkinson, C. Multi-spectral lidar: radiometric calibration, canopy spectral reflectance, and vegetation vertical SVI profiles. *Remote Sens.* **2019**, *11*, 1556.
41. Grams, G.W.; Blifford Jr., I.H.; Schuster, B.G.; DeLuisi, J.S. Complex index of refraction of airborne fly ash determined by laser radar and collection of particles at 13 km. *J. Atmos. Sci.* **1972**, *29*, 900–905.
42. Huang, B.-J.; Tang, C.-W.; Wu, M.-S. System dynamics model of high-power LED luminaire. *Appl. Therm. Eng.* **2009**, *29*, 609–616.
43. Dozier, J.; Green, R.O.; Nolin, A.W.; Painter, T.H. Interpretation of snow properties from imaging spectrometry. *Remote Sens. Environ.* **2009**, *113*, S25–S37.
44. Dozier, J.; Painter, T.H. Multispectral and hyperspectral remote sensing of alpine snow properties. *Annu. Rev. Earth Planet. Sci.* **2004**, *32*, 465–494.
45. Dozier, J.; Davis, R.E.; Chang, A.T.C.; Brown, K. The spectral bidirectional reflectance of snow. *Spectr. Signat. Objects Remote Sens.* **1988**, *287*, 87.
46. Kaasalainen, S.; Kaasalainen, M.; Mielonen, T.; Suomalainen, J.; Peltoniemi, J.I.; Näränen, J. Optical properties of snow in backscatter. *J. Glaciol.* **2006**, *52*, 574–584.
47. Kashani, A.; Olsen, M.J.; Parrish, C.E.; Wilson, N. A review of LIDAR radiometric processing: From Ad Hoc intensity correction to rigorous radiometric calibration. *Sensors* **2015**, *15*, 28099.
48. Parsons, N.R.; Hopkinson, C. The Development of an In-situ Light Emitting Diode Detection And Ranging Technology for Monitoring Snow Depth and Snow Surface Topography; *Proceedings of the 73rd EASTERN SNOW CONFERENCE, Columbus, Ohio, 14-16 June.* **2016**, 66-76.

Appendix 7: Airborne LiDAR Flight Path Sampling Optimization for Watershed-wide Snowpack Depth Monitoring

Barnes and Hopkinson. AIRBORNE LIDAR FLIGHT PATH SAMPLING OPTIMIZATION FOR WATERSHED-WIDE SNOWPACK DEPTH MONITORING. To be submitted to *Hydrological Processes* (Abstract only)

Airborne LiDAR Flight Path Sampling Optimization for Watershed-wide Snowpack Depth Monitoring

Celeste Barnes and Christopher Hopkinson

Department of Geography, University of Lethbridge, Lethbridge, AB T1K 3M4, Canada

ABSTRACT:

Due to the importance of mountain snowpack to Alberta's water supply and spring flood risk, the Government of Alberta collects field snowpack sampling data across the Rocky Mountain front ranges every month during winter. There is high spatial variability of snow depth in mountain environments so manual sampling alone cannot characterize snowpack conditions for the watersheds where water budget estimates are needed. Elevation, aspect, terrain relief, and vegetative cover influence snow accumulation and melt volumes and are highly variable in mountain regions. Our objective is to devise an optimal remote sensing snow depth monitoring framework that complements the existing monitoring program and delivers a more accurate and spatially explicit data product at low cost.

A high-resolution snow depth model (SDM) is created from airborne lidar by subtracting a pre-snow digital elevation model (DEM) from a LiDAR snowpack surface model (SSM). LiDAR coverage within a single flight over mountains in winter is limited to a few 100 square kilometers. Surplus aircraft and crew availability exist in the winter months, however survey resources may be on "stand-by" waiting for optimal conditions to fly due to limited daylight and favourable weather conditions. There is a need for optimal flight line sampling strategies that do not rely on 'wall-to-wall' coverage of watersheds such that sampling the smallest representative portion of the watershed will enable accurate spatial prediction of snow depth across the full watershed.

Several airborne LiDAR missions were flown over the West Castle watershed, a headwater of the Oldman River, Alberta, since 2014 with four bare-earth and eight winter snow depth surveys. Four dominant snow depth drivers were mapped across the West Castle watershed (elevation, aspect, topographic position index (TPI) and canopy coverage) and combined to develop spatially distributed SDM driver signatures. This produced 163 of a possible 180 signatures. With the goal of sampling as many signatures within the shortest flight line distance possible, two different configurations were tested. The first, shorter, flight line was created using an optimal path analysis capturing 120 signatures. The second path was generated using 'expert' selection criteria to sample the length, width and elevation range of the watershed using three transects perpendicular to the watershed centre line containing 131 signatures. Imputation was done to fill in data for the missing snow depth signatures for each of the optimal flight lines using sampled SDM data from March 2020. This dataset was then used to train a random forest prediction model across the watershed and compared with independently collected LiDAR SDM data. The two optimal sampling flight line swath extents from the March 2020 LiDAR mission were used to extract training data for the balance of the 2014 to 2020 winter LiDAR SDMs and compared to their respective full coverage datasets.

Initial results show that flight line sampling, spatial imputation, and prediction are effective at reconstructing snow depth at watershed-scale. The time- and cost-effectiveness, as well as accuracy of the two optimization strategies will be compared and presented. Overall, utilizing a sampling vs 'wall-to-wall' mapping approach reduces flight time by up to 80%, which, if implemented, should make LiDAR SDM sampling viable for operational monitoring.

Appendix 8: Watershed-wide Snow Density and SWE Estimation using Spatially Distributed Time Series Meteorological Observations, Airborne LiDAR, and Field Measurements

Barnes and Hopkinson. WATERSHED-WIDE SNOW DENSITY AND SWE ESTIMATION USING SPATIALLY DISTRIBUTED TIME SERIES METEOROLOGICAL OBSERVATIONS, AIRBORNE LIDAR, AND FIELD MEASUREMENTS. To be submitted to *Hydrological Processes*. (Abstract only)

Watershed-wide Snow Density and SWE Estimation using Spatially Distributed Time Series Meteorological Observations, Airborne LiDAR, and Field Measurements

Celeste Barnes and Christopher Hopkinson

Department of Geography, University of Lethbridge, Lethbridge, AB T1K 3M4, Canada

Extended abstract:

The Alberta Rocky Mountain snowpack provides a significant quantity of headwater resources for downstream use [1, 2] in the form of snow water equivalent (SWE) derived from snow depth and density measurements. Snow depth and density changes over time from when the pack first emerges until it completely melts out. Spatial variability of snow density is influenced by landform, terrain, and antecedent meteorological conditions. Within a forest stand, density varies based on amount of canopy coverage [3] where the interior is more sheltered, and less solar radiation penetrates to the ground surface [4]. Snow is persistent in higher elevations with the exception of wind scoured locations [5] as well as on north facing aspects. It melts quicker on south facing aspects, which in turn affects snow depth and density [6, 7]. As the snowpack metamorphoses and melts out, depth decreases and density increases [8-10]. At the watershed scale, a better understanding of spatial and temporal meteorological and land surface relationships related to snowpack properties is required to improve snow density modelling.

Well established airborne LiDAR methodologies have been developed to create snow depth models (SDM) at the watershed scale [11, 12]. To model and estimate snow density and SWE, various techniques exist ranging from data collection in highly instrumented *in-situ* installations [13-16] to larger regional sampling methods [17-19]. However, there still is a requirement to identify the drivers controlling variability in snowpack density across mountain watersheds. The objective of this study is to develop a computer machine learning imputation and prediction model based on the drivers of spatial snow density variability at the watershed scale. Model inputs use spatially distributed meteorological data such as solar radiation, temperature, precipitation, relative humidity, and wind as well as landform variables derived from high resolution airborne LiDAR bare-earth and SDMs.

High-resolution spatial and temporal meteorological data is very sparse in the headwaters of the Alberta Rocky Mountains. Public available datasets contain temperature and precipitation from sites primarily located in valley bottoms and lower elevations in the mountain regions [20, 21]. Snowpack monitoring has been operational for several decades capturing snow depth and snow water equivalent, with some stations collecting more extensive meteorological variables such as temperature, precipitation, wind, and relative humidity. Snow course data is available but again, there are few sites providing coverage for the Alberta mountain range.

The Westcastle watershed, located on the eastern side of the continental divide in the Alberta Rocky Mountains, contains the headwaters for southern Alberta and the Canadian Prairie Provinces [22]. The University of Lethbridge ARTeMiS Lab installed weather stations in the watershed and were fully operational by 2017 at different elevational and directional aspects in order to gain a better understanding of spatiotemporal meteorological relationships driving

snowpack variability. Three sites located at the valley, mid-mountain treeline, and ridge elevations record temperature, barometric pressure, relative humidity, wind, solar radiation, and snow depth. A new LED LiDAR sensor [23] was added to the University of Lethbridge Westcastle Field Station (WFS) weather station to collect snow depth data along 16 separate segments in the sensor's field of view. Precipitation is recorded at the WFS and ridge sites. Three additional treeline sites based on directional aspect collect air temperature, incoming shortwave solar radiation, and net radiation data. A near real-time website displays weather and environmental conditions for the Westcastle valley and ridge locations (<http://artemis.uleth.ca:8080/castlemet/index.html>).

Twelve high resolution airborne LiDAR missions were flown over the Westcastle watershed from 2014 to 2020. Four "bare-earth" surveys were completed during summertime months. Digital elevation and surface models (DEM and DSM respectively) were created to characterize landform and terrain attributes required to identify significant drivers of watershed snow depth and density variability. Eight winter airborne surveys sampled watershed snow covered area (SCA). Watershed wide SDMs were created by differencing SCA datasets from the bare-earth DEM. The government managed Gardiner Headwaters snow monitoring site was added to the mission plan in the later part of the Westcastle airborne LiDAR winter data collection series for SDM validation. Snow depth and density field measurements acquired by the ARTeMiS Lab, government, and industry partners took place during SCA winter months. Field sampling also occurred during some of the winter airborne LiDAR missions.

Snow monitoring and meteorological data collection sites located in the southern Alberta mountain region will be used to derive watershed wide snow density and SWE models. Datasets were acquired and compiled from publicly available open access collections, government and industry partners, and the ARTeMiS Lab. Meteorological and snow monitoring sites are spatially distributed at different elevational and directional aspects with installations in the valley bottom, tree line, and mountain ridge. Initial analysis focuses on the meteorological elevation and directional aspect spatiotemporal patterns for air temperature and solar radiation. Spatial relationships for precipitation, SWE, and manual field snow depth and density measurements will be examined and compared for all acquired datasets. The desired outcome of this study is to create a snow density model using computed landform, terrain, and antecedent meteorological drivers representing the snowpack spatial variability at the watershed scale. The basis of the model will rely on empirical formulas from the existing literature as well as manual field measurements. The final estimated watershed SWE will be calculated from the SDM and snow density model.

References:

1. Kienzle, S.W., *Water yield and streamflow trend analysis for Alberta watersheds*. 2012, Alberta Innovates, Energy and Environment Solutions: Edmonton, Alta.
2. Statistics Canada, *Average annual runoff in Canada, 1971 to 2013*, E. Statistics Canada, Energy and Transportation Statistics Division, Editor. 2017, Statistics Canada, Environment, Energy and Transportation Statistics Division: Ottawa, ON.

3. Veatch, W., P.D. Brooks, J.R. Gustafson, and N.P. Molotch, 'Quantifying the effects of forest canopy cover on net snow accumulation at a continental, mid-latitude site'. *Ecohydrology*, 2009. **2**(2): p. 115-128..
4. Garvelmann, J., S. Pohl, and M. Weiler, *Variability of Observed Energy Fluxes during Rain-on-Snow and Clear Sky Snowmelt in a Midlatitude Mountain Environment*. *Journal of Hydrometeorology*, 2014. **15**(3): p. 1220-1237.
5. Löffler, J., *Snow cover dynamics, soil moisture variability and vegetation ecology in high mountain catchments of central Norway*. *Hydrological Processes*, 2005. **19**(12): p. 2385-2405.
6. Anderson, B.T., J.P. McNamara, H.P. Marshall, and A.N. Flores, *Insights into the physical processes controlling correlations between snow distribution and terrain properties*. *Water Resources Research*, 2014. **50**(6): p. 4545-4563.
7. Dornes, P.F., J.W. Pomeroy, A. Pietroniro, S.K. Carey, and W.L. Quanton, *Influence of landscape aggregation in modelling snow-cover ablation and snowmelt runoff in a sub-arctic mountainous environment*. *Hydrological Sciences Journal*, 2008. **53**(4): p. 725-740.
8. Bartlett, P.A., M.D. MacKay, and D.L. Verseghy, *Modified snow algorithms in the Canadian land surface scheme: Model runs and sensitivity analysis at three boreal forest stands*. *Atmosphere-Ocean*, 2006. **44**(3): p. 207-222.
9. Verseghy, D., *CLASS—the Canadian Land Surface Scheme (Version 3.4)*. Environment Canada, Toronto, Ont, 2009.
10. Hedstrom, N.R. and J.W. Pomeroy, *Measurements and modelling of snow interception in the boreal forest*. *Hydrological Processes*, 1998. **12**(10-11): p. 1611-1625.
11. Hopkinson, C., T. Collins, A. Anderson, J. Pomeroy, and I. Spooner, *Spatial Snow Depth Assessment Using LiDAR Transect Samples and Public GIS Data Layers in the Elbow River Watershed, Alberta*. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques*, 2012. **37**(2): p. 69-87.
12. Deems, J.S., T.H. Painter, and D.C. Finnegan, *Lidar measurement of snow depth: a review*. *Journal of Glaciology*, 2013. **59**(215): p. 467-479.
13. Proksch, M., H. Löwe, and M. Schneebeli, *Density, specific surface area, and correlation length of snow measured by high-resolution penetrometry*. *Journal of Geophysical Research: Earth Surface*, 2015. **120**(2): p. 346-362.
14. Vivian, M., L. Bernt, H.K. V., and L.K. Robert, *Predicting snow density using meteorological data*. *Meteorological Applications*, 2007. **14**(4): p. 413-423.
15. Kang, D.H. and A.P. Barros, *Full-System Testing in Laboratory Conditions of an L-Band Snow Sensor System for In Situ Monitoring of Snow-Water Content*. *IEEE Transactions on Geoscience and Remote Sensing*, 2011. **49**(3): p. 908-919.
16. Naderpour, R., M. Schwank, and C. Mätzler, *Davos-Laret Remote Sensing Field Laboratory: 2016/2017 Winter Season L-Band Measurements Data-Processing and Analysis*. *Remote Sensing*, 2017. **9**(11): p. 1185.
17. Cornwell, E., N. Molotch, and J. McPhee, *Spatio-temporal variability of snow water equivalent in the extra-tropical Andes Cordillera from distributed energy balance modeling and remotely sensed snow cover*. *Hydrology and Earth System Sciences*, 2016. **20**(1): p. 411.
18. Painter, T.H., D.F. Berisford, J.W. Boardman, K.J. Bormann, J.S. Deems, F. Gehrke, A. Hedrick, M. Joyce, R. Laidlaw, D. Marks, C. Mattmann, B. McGurk, P. Ramirez, M. Richardson, S.M. Skiles, F.C. Seidel, and A. Winstral, *The Airborne Snow Observatory*:

- Fusion of scanning lidar, imaging spectrometer, and physically-based modeling for mapping snow water equivalent and snow albedo. Remote Sensing of Environment, 2016. 184: p. 139-152.*
19. Mizukami, N. and S. Perica, *Spatiotemporal Characteristics of Snowpack Density in the Mountainous Regions of the Western United States. Journal of Hydrometeorology, 2008. 9(6): p. 1416-1426.*
 20. Wright, R., *Current and Historical Alberta Weather Station Data A.C.I.S. (ACIS), Editor. 2017, Alberta Climate Information Service (ACIS): Edmonton, AB.*
 21. EnvironmentCanada, *Digital Archive of Canadian Climatological Data E. Canada, Editor. 2014, Environment Canada: Fredericton, New Brunswick.*
 22. CMP, *Crown Managers Partnership Strategic Conservation Framework 2016 - 2020. 2016, Crown Managers Partnership. p. 16.*
 23. Barnes, C., C. Hopkinson, T. Porter, and Z. Xi, *In-Situ LED-Based Observation of Snow Surface and Depth Transects. Sensors, 2020. 20(8): p. 2292.*

Appendix 9: A Lightweight Leddar Optical Fusion Scanning System (FSS) for Canopy Foliage Monitoring.

Xi , Hopkinson, Rood, Barnes, Xu, Pearce, Jones. 2019. A LIGHTWEIGHT LEDDAR OPTICAL FUSION SCANNING SYSTEM (FSS) FOR CANOPY FOLIAGE MONITORING. *Sensors* 19(18), 3943; doi:10.3390/s19183943

A Lightweight Leddar Optical Fusion Scanning System (FSS) for Canopy Foliage Monitoring

Zhouxin Xi, Christopher Hopkinson*, Stewart B. Rood, Celeste Barnes, Fang Xu, David Pearce, and Emily Jones

Department of Geography, University of Lethbridge, Lethbridge, AB T1K 3M4, Canada; zhouxin.xi@uleth.ca

* Correspondence: c.hopkinson@uleth.ca; Tel.: +01-403-332-4586 (C.H.)

Received: date; Accepted: date; Published: date

Abstract: A growing need for sampling environmental spaces in high detail is driving the rapid development of non-destructive three-dimensional (3D) sensing technologies. LiDAR sensors, capable of precise 3D measurement at various scales from indoor to landscape, still lack affordable and portable products for a broad-scale and multi-temporal monitoring purpose. This study aims to configure a compact and low-cost 3D Fusion Scanning System (FSS) with a multi-beam Leddar (light emitting diode detection and ranging, LeddarTech), a monocular camera and rotational robotics to recover hemispherical colored point clouds. This includes an entire framework of calibration and fusion algorithms utilizing Leddar depth measurements and image parallax information. The FSS was applied to scan a cottonwood (*Populus spp.*) stand repeatedly during autumnal leaf drop. Results show that the calibration error based on bundle adjustment is between 1-3 pixels. The FSS scans consistently exhibit a similar canopy volume profile to the benchmarking terrestrial laser scans, with an r^2 between 0.5-0.7 in varying stages of leaf cover. The 3D point distribution information from FSS also provides a valuable correction factor for the leaf area index (LAI) estimation. We expect the FSS or similar low-cost systems will increase in accuracy and become sufficiently robust for deployment within canopy monitoring sensor networks.

Keywords: LiDAR; Leddar; Multi-beam; Sensor Fusion; Structure from Motion; Monocular camera; Terrestrial Laser Scanning; Leaf Area Index; Canopy Monitoring;

1. Introduction

Monitoring *in-situ* canopy variables such as leaf area index (LAI) from ground are valuable for developing canopy light interception and biomass growth models for forests [1,2]. Passive optical sensors such as digital cameras are cost-effective ground-based tools and have been applied to monitor canopy variables such as LAI and the fraction of absorbed photosynthetically active radiation (fPAR) [3]. However, the passive optical approach provides limited accuracy due to the 3D heterogeneity of foliage distribution. For example, the LAI estimated from the Beer Lambert geometric-optical model [4,5], also termed effective LAI, is usually 55-65% of the true LAI [6]. The constraint of sensing tools inevitably leads to sophisticated tuning efforts for the geometric-optical model, such as the introduction of gap size distribution, clumping factor, and needle-to-shoot area ratio [7]. The emerging terrestrial laser scanning (TLS) technology greatly mitigates the LAI characterization problems. The 3D datasets from TLS not only enable straightforward estimation of canopy model variables including leaf angle distribution (LAD) [8,9], clumping index [10], canopy foliage profile [11], gap fraction [12] and plant area volume density (PAVD) [13], but have also led to the development of more accurate canopy models. For example, the non-randomness of leaf distribution,

conventionally described using the clumping index or gap size distribution, can be explicitly modeled in path length distribution equations [14,15]. Basically, the path length distribution (PATH) model relates the light extinction degree with the detailed optical paths extractable from 3D point clouds. The LAI variable can be estimated from the PATH with high accuracy and stability benchmarked by the true LAI [7,14]. Therefore, capturing 3D information has great potential for modeling canopy variables.

Conventional TLS sensors have shown impressive measurement precision and reliability of capturing 3D canopy information, yet lack portability and affordability for widespread use at landscape level. An alternative solution is to integrate low-cost LiDAR sensors into an existing ground sensor network for broad-scale monitoring purposes. Manufacturers such as Faro and Leica produce portable terrestrial or mobile LiDAR at a price range of 4,000 USD and 20,000 USD with moderate frequency (20k-300k) and centimeter-level resolution. Those sensors are not cost-effective, power-saving, compact and flexible enough for widespread 3D biomass monitoring. In recent years, tiny LiDAR scanners such as from Velodyne, Ouster, Hokuyo, SICK, Ibeo and Scanse have entered the market with a price level between 100 - 4,000 USD. Most tiny scanners have limited field of view (FOV) and low point detection frequency. Many low-cost tiny scanners scan in only 2D dimensions: recording laser distance returns with a spinning mirror or motor. Multiple scanlines are thus required to produce sufficient vertical points, while low frequencies can cause serious 3D distortions from fast-moving platforms or targets [16].

A multi-beam laser scanner may provide a balanced choice of budget and frequency. Instead of repetitive scanning in the vertical direction, a multi-beam scanner relies on detection arrays to record multiple distances instantly. Each beam from a detection array usually has high detection frequency over 50 Hz. Among the multi-beam scanners, a LED-based LiDAR from LeddarTech, or termed as Leddar, stands out as an economical choice with 16 beams, 100 Hz frequency and a thousand-dollar cost. A Leddar sensor implements patented algorithms to estimate traveling distances of each pulse emitted from a LED light source and detected by an array of 16 PIN photodiodes [17]. Each beam corresponds to a solid angle of around $2.8^\circ \times 7.5^\circ$ and the field of view for 16 beams is customizable to be between 9° and 95° . The sensor can record multiple distance returns from multiple objects at different distances. Leddar's capability of rapid data acquisition and multiple object detection has enabled growing applications in canopy detection [18], autonomous driving [19,20], traffic analysis [21,22], parking assistance [23], and drone altitude estimation [24,25]. However, due to limited FOV and sparse beam, the Leddar sensor alone cannot compete with conventional static TLS for detailed 3D canopy modeling or geometric mensuration.

Adding a twin-axis rotational robot to the Leddar sensor can be a cost-effective solution to expand its FOV and enhance point clouds. The boresight of the integrated system, however, needs to be calibrated in order to deliver precise and consistent 3D data or point clouds. One type of LiDAR calibration method can be statistical correction based on a known target. For example, Bohren, *et al.* [26] corrects 3D ground point clouds from SICK and Velodyne scanners based on the planarity constraint of ground. A more rigorous calibration method is to model the physical relationship between the LiDAR system and calibration targets. For example, Muhammad and Lacroix [27] use a planar target to calibrate five intrinsic parameters of Velodyne HDL-64E S2 system mounted on a static rotator, including two beam angle and three origin parameters. Atanacio-Jiménez, *et al.* [28] expands the calibration target to be a room of five planes and calibrates intrinsic and extrinsic parameters of Velodyne HDL-64E. Zhu and Liu [16] estimate origin and orientation of the HDL64E S2 sensor by aligning point clouds based on pole-shaped features. Several other studies provide convenient physical calibration approaches without a requirement for measuring reference target coordinates. For example, Levinson and Thrun [29] proposes a global calibration method for 192 orientation or distance parameters generated from a moving trajectory of Velodyne HD-64E S2 sensor. Point clouds are self-calibrated by maximizing local planarity without a need for a deliberate reference

target. Similarly, Sheehan, *et al.* [30] proposes an entropy-based self-calibration method maximizing the sharpness of point clouds for three SICK LMS-151 laser scanning units.

The above calibration methods require accurate position and orientation data from external sources such as wheel encoder, GPS or IMU, and also dense points to support point cloud alignment and registration. These requirements are usually not easily met in a low-cost scanning system with Leddar and rotational motor. An alternative solution is to integrate a camera sensor. The extrinsic parameters such as pose and origin can be estimated using photogrammetric methods. The predominant approach from existing studies is to re-project LiDAR point clouds to a 2D plane and co-register with an image based on corresponding features [31-36]. The calibration accuracy from these studies greatly depends on highly dense point clouds which a low-cost Leddar sensor is unable to produce. Among the few studies on calibrating sparse LiDAR data, Debattisti, *et al.* [37] places focus on edge points of artificial targets visible from both an image and a SICK LMS221 sensor. Debattisti, Mazzei and Panciroli [37] also points out that point clouds from low-cost LiDAR usually leads to laborious scanning in pursuit of sufficient corresponding points from both image and point cloud. Considering the characteristics of a low-cost scanning system, this study will utilize a planar calibration target to avoid the prerequisite for dense point clouds and also integrate a camera to provide extrinsic pose estimation.

Adding a camera sensor to the scanning system not only satisfies the calibration and alignment need but also provides useful texture details. A question of interest is how to integrate the texture information from camera with the sparse point clouds from a Leddar scanning system, to produce dense colored point clouds. Indeed, a variety of existing literature, Hartley and Zisserman [38] in particular, already illustrate the feasibility of reconstructing dense 3D point clouds from stereo or multiple images directly without LiDAR distance. However, with regard to our small scanning system with one rotational camera, the short movement baseline will lead to poor 3D reconstruction quality. Assuming the target is far from the camera, the error of depth estimate (ΔD) is related to the error of stereo matching between two images (Δx) in Equation 1:

$$\Delta x = \frac{fd\Delta D}{\mu D^2} \quad (1)$$

Assuming the baseline d of the small scanning system is about 3 cm, focal length f 3.6 mm, pixel size μ 2.8 μ m and target distance D is 20 m, then a small Δx of 0.1 pixel will lead to a ΔD of 1 m. We can conclude that using the monocular camera alone without Leddar point clouds in this case cannot produce an accurate point cloud. Therefore, a camera cannot replace Leddar in recovering depth (Z) information, but can be useful to fill 2D planar (XY) gaps between the beams from the interpolation point of view. The methods for interpolating or filling point clouds are many, yet few studies consider this level of fusion. De Silva, *et al.* [39] is among the few studies which matches resolution between image and point clouds using a Gaussian process model. This study will combine both the classic structure from motion (SFM) method and posterior interpolation method for dense 3D recovery.

The objective of this chapter is to (1) configure a compact and low-cost fusion scanning system (FSS) including multi-beam Leddar, monocular camera and rotational robotics, and (2) propose an entire framework of calibration and fusion algorithms that produce dense colored point clouds covering a hemispherical view for 3D canopy monitoring. The dense point clouds can provide rapid canopy information such as gap fraction and LAI, useful for monitoring foliage biomass status and changes. Most parts of the framework are automatic and aimed to reduce potential manual cost. For example, only a few ground truth measurements and one horizontal scan is required to complete calibration, and only ten minutes of field scanning is necessary to reconstruct dense point clouds over a hemispherical view. We

expect our work could promote more research into lightweight scanning systems for cost-effective 3D environmental mapping and monitoring.

2. Materials and Methods

2.1. Hardware Customization and Data Processing Framework

The Leddar optical FSS mainly consists of two sensors: a 16-beam Leddar sensor (Leddar M16) and a web camera sensor in a 3D printed enclosure (13.3 x 9.1 x 4.1 cm). The two lightweight sensors (< 300 gram) sit on top of a tilting arm (DDT560 Direct Drive Tilt) and a panning base shown in Figure 1. Beside and beneath the pan-tilt arms are rotational servos which drive the pan-tilt movement and determine the angular resolution and span of the rotation. Specifically, the tilting servo (Hitec HS-5485HB) is a standard digital servo that rotates between 0-118° with highest resolution of 0.6°. The pan servo (Dynamixel MX-12W) can have 360° rotation with 0.08° resolution and can feedback rotation angles in real time. Specifications of Leddar, camera, and servos are provided in Table 1. Camera video and Leddar distance data are collected and stored by Raspberry Pi and servo rotations are manipulated by an Arduino Mega 2560 board. The detailed connections between sensors, pan-tilt robotics and electronic controllers are shown in Figure 2.

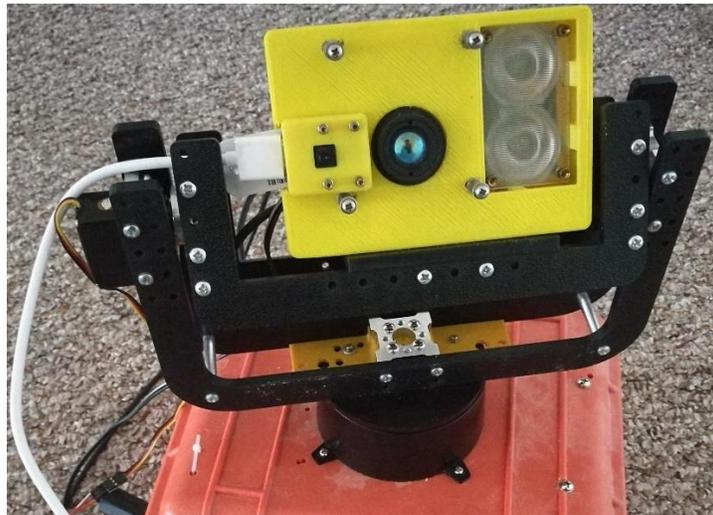


Figure 1. A Fusion Scanning System (FSS) with Leddar and monocular camera sensors.

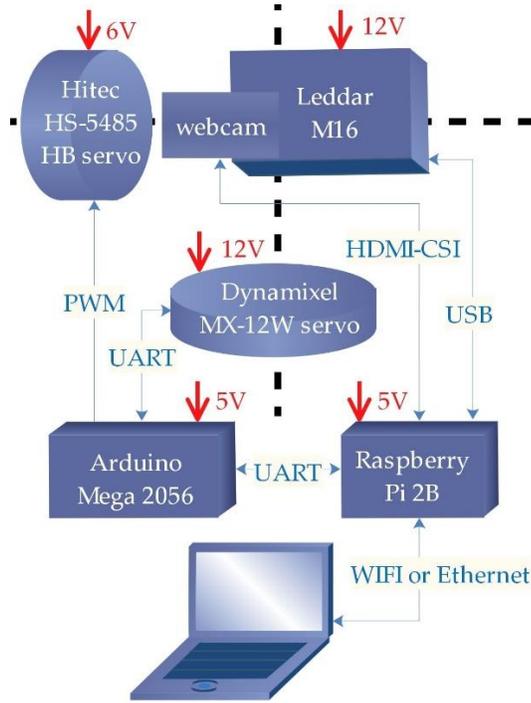


Figure 2. Hardware components and connections of the FSS. PWM: Pulse Width Modulation. UART: Universal Asynchronous Receiver/Transmitter. CSI: Camera Serial Interface.

Table 1. Major hardware specifications.

| Camera | Leddar | Tilt servo | Pan servo |
|---------------------------|------------------------|---------------------|-----------------------|
| OmniVision OV5647 | M16 module | Hitec HS- 5485HB | Dynamixel MX- 12W |
| FOV: 54° x 41° | Distance: 0 to 50 m | Max angle: 118° | Max angle: 360° |
| Lens: f=3.6 mm, f/2.9 | Frequency: ≤100 Hz | PWM: 750-2250 μs | Steps: 4096 |
| Calibration: no IR | Wavelength: 940 nm | Deadband: 8 μs | Resolution: 0.088° |
| Application: IR filter | Power: 12/24 V, 4 W | Power: 4.8-6.0 V | Voltage: 12 V |

To scan a wide field of view (e.g. hemispherical view), the pan-tilt system rotation follows a common raster scanning scheme: setting a fixed tilt angle for one horizontal scanline rotation and changing tilt angle for the next. For each horizontal scanline, recordings from Leddar, camera and servo are stored asynchronously in separate files by Raspberry Pi. The Leddar sensor outputs timestamp, beam distance, echo amplitude and echo quality index with an updating frequency of 100 Hz. Raspberry Pi camera captures 720p video with a rate of 25 frames per second (FPS). The rotation angle from pan servo is saved

every 4°. The Leddar and pan angle readings are synchronized based on their millisecond-level timestamps tagged by Raspberry Pi. The camera videos do not have timestamps and their timing is inferred from motion detection.

Constructing the hardware of the FSS is technically straightforward. The system components are inexpensive and uncomplicated compared to most commercial LiDAR scanning systems, yet fusion of different low-resolution data sources to generate dense point cloud is the primary challenge. The framework of our multi-source data fusion is illustrated in Figure 3, including: (1) mapping discrete Leddar distances onto individual video frames to create “3D pixels”, (2) aligning video frames globally to cover a panorama field of view, and (3) adjusting frame alignment and extrapolating 3D point clouds based on the “3D pixels”.

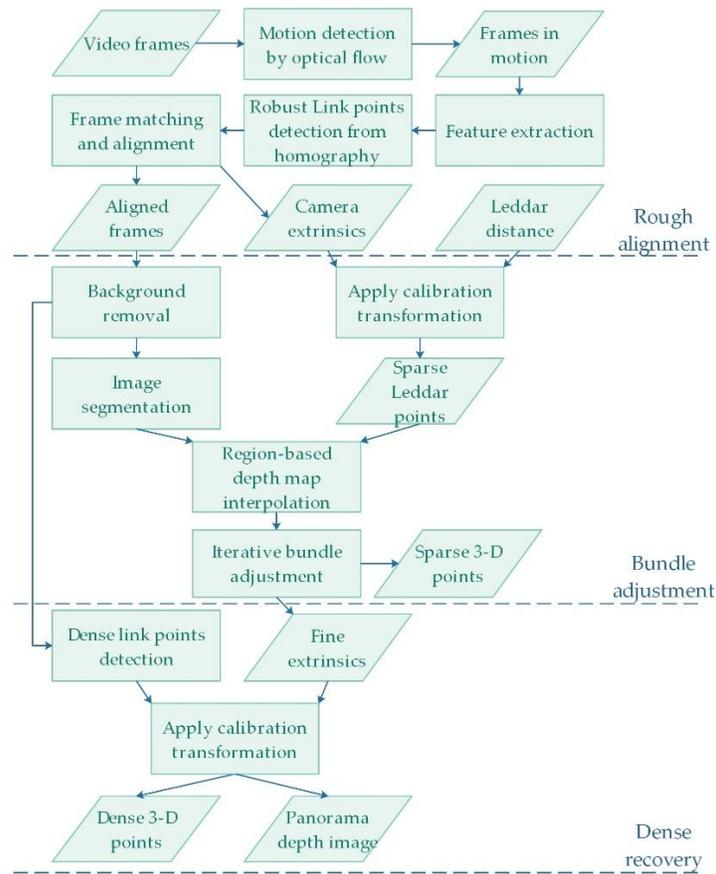


Figure 3. Framework of point cloud recovery from monocular camera and sparse Leddar beams.

2.2. Coordinate System Conversion for Calibration

Image frames decomposed from camera video do not have timestamps directly linkable to Leddar timestamps. It is necessary to match the camera motion with the Leddar motion and assign timestamps to camera frames. The start and stop of camera motion are detected by optical flows from neighboring frames: frames with average pixel velocity above a threshold of 0.6 pixel are considered moving. With start and

stop timestamps, the entire timestamps of the moving frames can be recovered based on linear interpolation.

The synchronization of camera and Leddar enables one-to-one mapping between Leddar distances and camera frames. We need to further locate Leddar footprints on the pixels of each video frame. This is solved through calibration of the FSS, or specifically, determining the unknown boresight parameters that convert Leddar coordinate system to the camera coordinate system. Calibrating the FSS is only needed once which enables the step of applying calibration transformation in the framework in Figure 3.

Our reference coordinate system (RCS) is a right-handed Cartesian coordinate system. It has the same units as the world coordinate system (WCS), with its origin at the camera optical center, x axis along the transverse direction of the image plane, y axis along the longitudinal direction of the image plane, z axis along the camera optical view direction. The target 3D coordinates $\bar{\mathbf{P}}_t$ in RCS can be parametrized in Equation 2:

$$\bar{\mathbf{P}}_t = \begin{pmatrix} T \\ 0 \\ 0 \end{pmatrix} + (D_{it} + D_b) \begin{pmatrix} \cos(\theta_i) \sin(\psi_i) \\ \sin(\theta_i) \sin(\psi_i) \\ \cos(\psi_i) \end{pmatrix} \quad (2)$$

where $\psi_i = \psi_0 + (8.5 - i)\psi_\Delta$, $\theta_i = \theta_0$, $i = 1, 2, 3, \dots, 16$

where t is a specific time point, T the horizontal location of Leddar optical center in RCS, polar angles θ_i and ψ_i the orientation of i^{th} beam segment in RCS, D_{it} the distance measurement of i^{th} beam segment, D_b the bias of distance measurement, and $\bar{\mathbf{P}}_t$ the target 3D coordinates in RCS. We consider T , D_b , θ_i and ψ_i to be constant during pan-tilt rotation for no relative movement between Leddar and camera. Assuming all the beam segments are equiangular, angle ψ_i can be presented by an arithmetic sequence parametrized with ψ_0 and ψ_Δ (Equation 2).

The real-world coordinates of target \mathbf{P} can be converted from $\bar{\mathbf{P}}_t$ in RCS through camera extrinsic parameters in Equation 3:

$$\mathbf{P} = \mathbf{R}_t \bar{\mathbf{P}}_t + \mathbf{T}_t \quad (3)$$

where \mathbf{R}_t is the rotation matrix, \mathbf{T}_t the translation vector. \mathbf{R}_t can be characterized by Euler angles $(\alpha_t, \beta_t, \gamma_t)$ following z - y - x rotation order. Assuming the initial rotation matrix is \mathbf{R}_0 or Euler angles $(\alpha_0, \beta_0, \gamma_0)$, the temporal change of \mathbf{R}_t during a horizontal scan can be represented by rotation matrix \mathbf{R}_w in Equation 4:

$$\begin{aligned} \mathbf{R}_t &= \mathbf{R}_0 \mathbf{R}_w; & (4) \\ \mathbf{R}_w &= \begin{pmatrix} 1 - 2(u_y u_y + u_z u_z) & 2(u_x u_y - u_w u_z) & 2(u_w u_y + u_x u_z) \\ 2(u_x u_y + u_w u_z) & 1 - 2(u_x u_x + u_z u_z) & 2(u_y u_z - u_w u_x) \\ 2(u_x u_z - u_w u_y) & 2(u_y u_z - u_w u_x) & 1 - 2(u_x u_x + u_y u_y) \end{pmatrix}; \\ \mathbf{T}_t &= \begin{pmatrix} X_{ct} \\ Y_{ct} \\ Z_{ct} \end{pmatrix} = \begin{pmatrix} X_{c0} + a_1 \phi_x + a_3 \phi_z - (a_{01} \phi_x + a_{03} \phi_z) \\ Y_{c0} + b_1 \phi_x + b_3 \phi_z - (b_{01} \phi_x + b_{03} \phi_z) \\ Z_{c0} + c_1 \phi_x + c_3 \phi_z - (c_{01} \phi_x + c_{03} \phi_z) \end{pmatrix}; \end{aligned}$$

where

$$\begin{aligned} \begin{pmatrix} u_x \\ u_y \\ u_z \\ u_w \end{pmatrix} &= \begin{pmatrix} \sin(\phi_\alpha) \cos(\phi_\beta) \sin(\omega_t/2) \\ \cos(\phi_\alpha) \sin(\omega_t/2) \\ \sin(\phi_\alpha) \sin(\phi_\beta) \sin(\omega_t/2) \end{pmatrix}; \\ \mathbf{R}_t &= \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}, \mathbf{R}_0 = \begin{pmatrix} a_{01} & a_{02} & a_{03} \\ b_{01} & b_{02} & b_{03} \\ c_{01} & c_{02} & c_{03} \end{pmatrix}. \end{aligned}$$

where ϕ_α, ϕ_β define the horizontal rotation axis and w_t is the horizontal rotation angle measured by the servo. The camera optical center \mathbf{T}_t also slightly moves during the horizontal scan, whose temporal change can be parameterized by its initial position (X_{c0}, Y_{c0}, Z_{c0}) and rotation origin (ϕ_x, ϕ_z) in Equation 4.

Solving the calibration Equations 2, 3 and 4 requires measurement of \mathbf{P} and D_{it} from the same target point. However, since the exact Leddar point is invisible from the web camera, it is impossible to measure the exact 3D coordinates for \mathbf{P} in the real world. Instead, we can reduce the requirement of the 3D \mathbf{P} by finding a planar target with constant Z values (Z_0) and arbitrary X and Y values. Therefore, combining Equation 3 and 4, $\bar{\mathbf{P}}_t$ and in RCS should satisfy a planar constraint in Equation 5:

| | |
|--|-----|
| $(c_1 \ c_2 \ c_3)\bar{\mathbf{P}}_t + Z_{ct} = Z_0$ | (5) |
|--|-----|

where Z_{ct} is the Z component of \mathbf{T}_t . With Equation 2 and 5, the only two required measurements are D_{it} at multiple time points and Z_0 of the planar target.

The above four equations form a set of nonlinear calibration equations with nine unknown intrinsic terms ($T, D_b, \theta_0, \psi_0, \psi_\Delta, \phi_\alpha, \phi_\beta, \phi_x, \phi_z$) and six initial extrinsic terms ($X_{c0}, Y_{c0}, Z_{c0}, \alpha_0, \beta_0, \gamma_0$). Our solution is iterative. The initial extrinsic terms ($X_{c0}, Y_{c0}, Z_{c0}, \alpha_0, \beta_0, \gamma_0$) are solved using least-square regression of camera collinearity equations (Equation 6):

| | |
|---|-----|
| $\mathbf{x}_t = K\bar{\mathbf{P}}'_t = KR_t^T(\mathbf{P}' - \mathbf{T}_t)$ | (6) |
| $K = \begin{pmatrix} \frac{f}{\mu} & 0 & x_0 \\ 0 & \frac{f}{\mu} & y_0 \\ 0 & 0 & 1 \end{pmatrix}$ | |

given additional measurements of pixel coordinates $\mathbf{x}_{t=0}$ and the corresponding world coordinates \mathbf{P}' (details in Section 2.3). The R_t^T in Equation 6 denotes transposed R_t in Equation 3. For simplicity, the camera intrinsic parameters in Equation 6 are fixed, including camera focal length f , pixel size μ , half image width x_0 and half height y_0 in pixels. Lens distortion is not considered in this study. Combining Equation 4 and 6, both extrinsic terms ($X_{c0}, Y_{c0}, Z_{c0}, \alpha_0, \beta_0, \gamma_0$) and intrinsic ($\phi_\alpha, \phi_\beta, \phi_x, \phi_z$) can be inferred by non-linear least-square regression, and further substituted into Equation 2, 3 and 5 to finally estimate Leddar intrinsic parameters ($T, D_b, \theta_0, \psi_0, \psi_\Delta$).

After knowing all intrinsic and extrinsic parameters, locating Leddar points on camera images is feasible using Equation 2 and 6. It is also possible to roughly estimate a Leddar point \mathbf{P} given D_{it} and w_t from pan-tilt angles based on Equation 2, 3 and 4. Instead of using pan-tilt angles, we can also rely on camera global alignment in the following section for more precise inference of R_t and \mathbf{T}_t and thus estimate \mathbf{P} from Equation 2.

2.3. Calibration Experiment

This section presents an example of system calibration, with a flat wall ($Z_0 = 0$) being our calibration target (Figure 4). The left bottom corner of the wall was defined as the WCS origin. The sensors scanned the wall with a fixed tilt angle of about 15° and continuous horizontal angle from 50° to 140° (2.4° per

second). A total of 70 frames were subsampled from the video with equal interval for calibration. On the front wall was a 16×9 grid from an optical projector, for the purpose of calibrating the camera extrinsic parameters in Equation 6. As mentioned, the calibration equation (Equation 6) required measuring the WCS coordinates \mathbf{P}' of each projected circle center and extracting the corresponding pixel coordinates \mathbf{x}_t from camera frames. We used a grid of circular targets instead of a chessboard pattern to be more robust to edge detection error during automatic extraction of \mathbf{x}_t from camera images. Since the camera had limited field of view of around 50° , not all circles appeared in camera images. Therefore, an ID number was assigned to each circle to help link \mathbf{x}_t and \mathbf{P}' automatically.

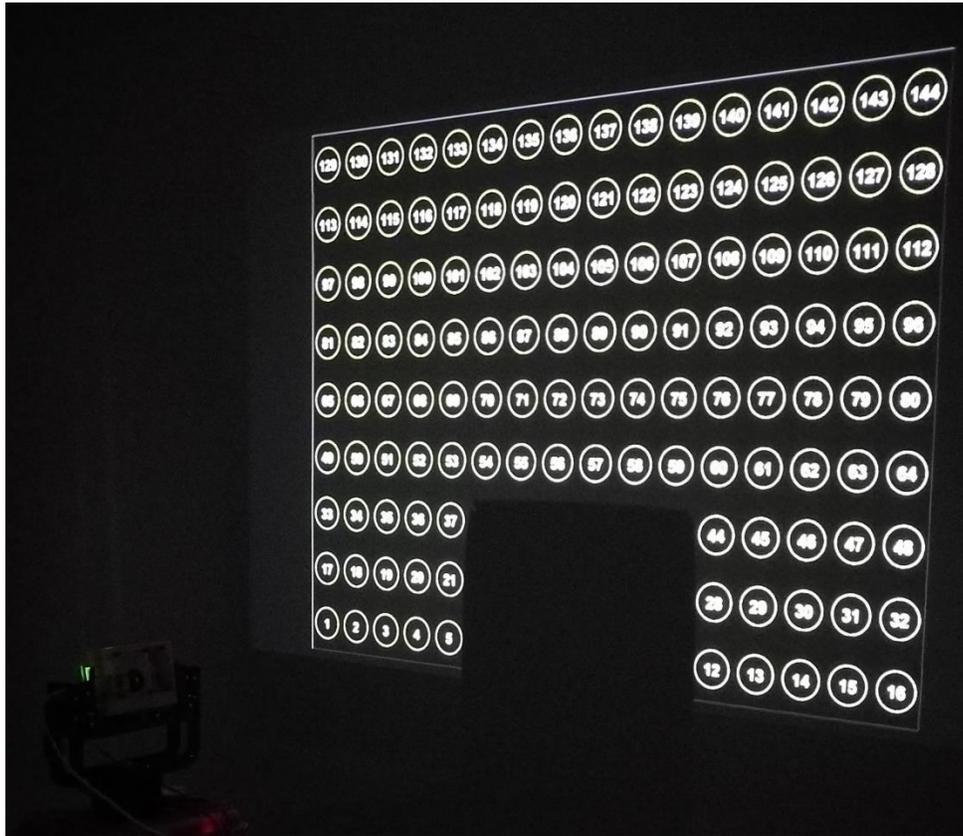


Figure 4. Experiment setup for the FSS calibration.

Assuming the projector lens had no distortion, we manually measured the \mathbf{P}' 's of the four corner circle centers on the wall and then applied bilinear interpolation to get \mathbf{P}' 's of all the 144 circle centers. The process of extracting \mathbf{x}_t was challenging because circle projection on the wall from the camera's view perspective exhibited ellipse shapes. Extracting ellipse centers is more difficult than extracting circle centers. We adopted Jia, *et al.* [40]'s Characteristic Number Ellipse Detector (CNED) to coarsely detect ellipse parameters (centers and axes). All the settings of the CNED were set as default except the CNL parameter of rejecting linearity, which was set to 10.0 instead of 3.0. Due to thickness of ellipse edges on the wall, redundant ellipses could be detected using the CNED. The next steps averaged the center and axis parameters over all redundant ellipses, edges were thinned using a morphological operator, and a finely fit ellipse [41] within each ellipse region was defined by its axis parameters. The ID number within each

ellipse region was identified by Optical Character Recognition (OCR) in MATLAB, with character set constrained to numbers 0-9. Using ellipse IDs greatly expedited the search for correspondence between \mathbf{x}_t and \mathbf{P}' from 70 camera frames. The accuracy of OCR recognition using MATLAB was about 80%. Improper IDs were later corrected by voting from the nearest four IDs and the final recognition accuracy was 100% among the 70 frames. Given the corresponding \mathbf{x}_t and \mathbf{P}' , the least-square Newton-Raphson iterative method [42-44] was applied to minimize residuals in Equations 2-6 following the aforementioned steps in Section 2.2. At the end of each iteration, the robust Huber's function [45] was applied to reduce the effect of residual outliers. Initial estimates of parameters were also required by the nonlinear Newton-Raphson method, in which $(\alpha_0, \beta_0, \gamma_0)$ were roughly set as $(0^\circ, 40^\circ, 0^\circ)$, (X_{c0}, Y_{c0}, Z_{c0}) manually measured as (1.80, 1.08, 1.09) in meters, and $(T, D_b, \theta_0, \psi_0, \psi_\Delta, \phi_\alpha, \phi_\beta, \phi_x, \phi_z)$ to be (0.03m, -0.44m, 180° , 90° , 2.5° , 0° , 0° , 0.06m, 0.00m).

2.4. Fusion-based Dense Point Cloud Recovery from FSS

Calibration is a preliminary requirement for fulfilling the framework in Figure 3. The framework focuses on generating and optimizing point clouds from multiple field scans. This section describes the suite of visual odometry algorithms adopted, with a field experiment presented in the following section. Our field scans cover a hemispherical view with four scanlines spanning a vertical angle between $0^\circ - 120^\circ$, though more scanlines can be added if desired. To avoid data overhead, we chose systematic sampling of the moving frames. Hence, our full hemispherical scan contains 150×4 moving frames with horizontal overlap of $> 90\%$ and vertical overlap of $\sim 80\%$. Aligning frames is the first problem since the extrinsic parameters in the lab calibration environment are not repeatable in a new location. The intrinsic parameters such as $(T, D_b, \theta_0, \psi_0, \psi_\Delta)$ from the Leddar sensor remain unchanged. To estimate the extrinsic parameters, target-based calibration or ground control points, while possible in some cases, would be tedious for a hemispherical view of 600 frames. Therefore, we directly use the dense and detailed photogrammetric information from the video to approximate camera poses R_t and then incorporate Leddar distance into the bundle adjustment for fine camera extrinsic parameters (R_t and T_t).

A common way of aligning multiple frames is to extract invariant features in each frame, match features between frames, and optimize the camera colinear equation (Equation 6). Frames with multiple scanlines also need iterative correction of scanline skewness caused by uneven distribution of matching features. The set of global alignments has been supported by image stitching software such as PtGUI for this study. A weakness of using off-the-shelf stitching software is the limited number of identified matched pixels, which is insufficient for bundle adjustment needs. Therefore, intensive extraction of SURF features [46] is added to the workflow in Figure 3, with outlier features filtered out using the homograph-based RANSAC algorithm [38]. The extracted SURF features are then matched between frames. Note that a 3D point can correspond to a set of matched pixels from multiple frames. The matched pixels from more than three frames are called "key pixels" here, whose features can be considered stable and will be used for the bundle adjustment later.

The global alignment using PtGUI exports Euler angles of each frame, which can be used to interpolate Euler angles $(\alpha_t, \beta_t, \gamma_t)$ or R_t at any timepoint when Leddar distances are measured. Then Leddar point clouds \mathbf{P}_t can be roughly recovered using Equations 2 and 3, assuming T_t is zero vector. Projecting the Leddar point clouds \mathbf{P}_t back to each image frame will add depth information to a few pixels (or "3D" pixels here). Then the sparse "3D" pixels can interpolate depth values of all the "key pixels" which will be input to the bundle adjustment later. Our method of interpolating depth value is region-based: (1) filtering the foreground in each image using k-means clustering method ($k=2$); (2) segmenting images using Statistical

Region Merging algorithm (level=8) [47]; and (3) implementing Inverse distance weighting (IDW) interpolation for each key pixel within each region. Compared to global interpolation, using region-based interpolation better maintains sharp boundary lines between different image regions.

Based on the 3D “key pixels”, both transformation matrix R_t and T_t and WCS coordinates P_t can be finely estimated using iterative bundle adjustment. First, P_t estimated from 3D “key pixels” are averaged among different frames and reprojected to each image frame using Equation 7:

The

$$\begin{pmatrix} x_t \\ y_t \\ z_t \end{pmatrix} = KR_t^T \left[\begin{pmatrix} X_t \\ Y_t \\ Z_t \end{pmatrix} - T_t \right] \quad (7)$$

$$\min_{\alpha_t, \beta_t, \gamma_t, X_{ct}, Y_{ct}, Z_{ct}} \left\| \begin{pmatrix} x'_t \\ y'_t \end{pmatrix} - \begin{pmatrix} x_t/z_t \\ y_t/z_t \end{pmatrix} \right\|, \text{ using nonlinear regression} \quad (8)$$

$$\begin{pmatrix} X'_t \\ Y'_t \\ Z'_t \end{pmatrix} = R_t K^{-1} \begin{pmatrix} x'_t \\ y'_t \\ 1 \end{pmatrix} + T_t \quad (9)$$

$$\min_{x_t, y_t, z_t} \left\| \begin{pmatrix} X_t \\ Y_t \\ Z_t \end{pmatrix} - \begin{pmatrix} X'_t \\ Y'_t \\ Z'_t \end{pmatrix} \right\| \text{ using ridge regression } (\lambda = 0.01) \quad (10)$$

reprojected points are normalized by Z coordinate and compared to the 2D coordinates of the “key pixels”. The least square error is minimized using nonlinear regression, with camera extrinsic parameters ($\alpha_t, \beta_t, \gamma_t, X_{ct}, Y_{ct}, Z_{ct}$) as variable (Equation 8). New P'_t corresponding to “key pixels” is estimated using Equation 9 based on the optimized camera extrinsics. Note that P_t is sensitive to small errors of “key pixels” and camera extrinsics due to the ill-posed mono-camera geometry. A robust solution to the ill-posed optimization problem is to use ridge regression [48] which partially minimizes the least square error between P_t and P'_t in Equation 10. Its regularization parameter λ is set to be 0.01. The optimized P_t from Equation 10 is again reprojected to each image frame in Equation 7, and repeat iterations from Equation 7 to Equation 10 until the error in Equation 8 is locally minimal. This iterative bundle adjustment for 3D recovery is similar to the Gold Standard Method from Hartley and Zisserman [38], but with depth information provided from sparse Leddar distance instead of dense stereo geometry. The resulting point clouds will exhibit rich details in the 2D planar direction but limited variation in the depth direction.

The 3D recovery from bundle adjustment produces point clouds for “key pixels”. The “key pixels” are essentially from SURF feature extraction and are mostly focused on corner pixels with sharp color gradient. Other “internal” foreground pixels should also be incorporated to produce complete and dense point clouds. Our method is to extract dense foreground pixels and solve Equation 9 and 10 to create optimal dense 3D points. To satisfy Equation 10, one 3D point requires at least one pair of matching pixels from two frames. We already have matching pixels defined as “key pixels” in previous steps. We need to extrapolate the matching relationship for all foreground pixels. This is a pixel-level dense matching process. First, the foreground pixels need to be subsampled at a certain interval (e.g. 10 pixels in this study) to avoid data overhead. Then the disparities of “key pixels” between current frame and one matched frame are calculated. The disparities are used to interpolate a disparity map for all foreground pixels in the current frame. Given a disparity map, the foreground pixel location in the matched frame can be estimated. This dense matching step between a pair of matched frames is repeated for all the matching frames. Each set of matched pixels from multiple frames is given one unique ID, corresponding to one unique P_t . Finally,

using Equation 9 and 10, the densely matched sets of pixels can produce dense point clouds. The final RGB colors of dense point clouds are the average RGBs from matched pixels.

2.5. Application: Tracking Autumn Leaf Drop Processes with the FSS

The FSS was used to track canopy changes during an autumn season in 2018. Our experiment site was located in an area of cottonwood (poplar) stands in Lethbridge, Canada (49°41'45.2"N, 112°51'54.0"W). The FSS was mounted on a tripod surrounded by six poplar trees within 20 m, including *Populus angustifolia*, *P. deltoides* and their hybrid, *P. x acuminata* [Zanewich et al.,]. The irregular shapes of poplar trees increased the difficulty of depth information recovery but the rich texture of the scene facilitated feature extraction, which supports frame alignment. The camera lens filter was replaced to block near-infrared light and enable natural-colored images. Multiple scanlines were collected, each corresponding to 360° horizontal rotation at a speed of 2.4° per second. Four scanlines were selected to cover the upper hemispherical view of the scene for further processing. All scans were aligned, optimized, and densified to create colored point clouds using the methods in Section 2.4. The same scene was scanned with a Teledyne Optech ILRIS HD (1535 nm) TLS as a benchmark. The scanning angle was 360°x80°, with the small zenith angle between 0-10° not scanned. Only last returns were recorded and point spacing was 3.2 cm at a distance of 20 m from the TLS. A total of 30 ILRIS TLS scans were collected in 30 minutes with three scanlines covering the entire upper hemispherical view. The TLS scans were co-registered by the Iterative Closest Point (ICP) algorithm into one hemispherical scan with an average accuracy below 1.3 cm. The same scanning and processing activities were repeated on September 9th (09/09), September 17th (09/17), October 1st (10/01) and October 17th (10/17) during the autumn defoliation period in 2018, to evaluate the reusability of our static scanning system in a temporal monitoring context. Hemispherical photos based the Digital Hemispherical Photography (DHP) methods were also captured on 09/09, 09/17 and 10/17 for benchmarking purposes.

Canopy vertical volume profile and Plant Area Index (PAI) were extracted from the FSS point clouds to evaluate the capabilities of 3D canopy detection and canopy attribution. The volume profile was defined as a total volume of voxels at each height, where a unit voxel was $0.1 \times 0.1 \times 0.1 \text{m}^3$ and a height slice was 0.1m. The PAI was calculated based on a path length distribution (PATH) model [7,14]. Specifically, the PATH model consisted of two equations (Equation 11 and 12) with the gap fraction $\overline{P(\theta)}$ and path distribution p_l to be the only two inputs. To calculate angle-specific gap fractions, a hemispherical image from either FSS or TLS point clouds was first converted to a black and white binary image under a fisheye perspective. The fisheye binary image was equally sliced into 28 rings representing a zenith angle between 15-69° and a ring width of 4°. The overlap between two neighboring rings is 2°. The gap fraction \overline{P} was defined as the ratio of the “hole” pixel numbers to the “filled” pixel numbers within a ring slice. The “filled” pixels represented the overall canopy area and was generated based on image morphological smoothing. The path distribution p_l was defined as the probability density function (PDF) of the optical path length within the crown area, with l representing the within-crown path length. The l ranged between 0-1, scaled by the maximum value l_{max} . The p_l was approximated by the histogram of the l , normalized by the total histogram area. With both the gap fraction $\overline{P(\theta)}$ and path distribution p_l extracted from crown area, the integral equation (Equation 11) was solvable based on any root-finding algorithm and the $FAVD \cdot l_{max}$ could be estimated, where $PAVD$ stands for the plant area volume density and $G(\theta)$ the leaf angle distribution. The $G(\theta)$ was set 0.5 in this study corresponding to a spherical leaf angle distribution [49]. The $PAVD \cdot l_{max}$ was then input to the Equation 12 to determine the $PAI_{true}(\theta)$ at a specific zenith angle θ . The final PAI value was a weighted sum of $PAI_{true}(\theta)$ over all zenith angles (Equation 13) [14].

$$\overline{P(\theta)} = \int_0^1 e^{-G(\theta) \cdot (FAVD \cdot l_{max}) \cdot l} p_l dl, \text{ where } \int_0^1 p_l dl = 1 \quad (11)$$

$$PAI_{true}(\theta) = \int_0^1 \cos(\theta) \cdot (FAVD \cdot l_{max}) \cdot l \cdot p_l dl \quad (12)$$

$$PAI_{true} = \frac{\sum_{\theta} PAI_{true}(\theta) \cdot \sin(\theta)}{\sum_{\theta} \sin(\theta)} \quad (13)$$

An important portion of canopy was foliage and the corresponding index was LAI. The LAI was directly related to photosynthetic processes and carbon productivity, and was a more sensitive index than PAI to reflect seasonal biomass changes. We estimated LAI values by contrasting leaf-on and leaf-off gap fraction values, as illustrated in Equation 14, where N was the number of pixels and P was short for the gap fraction $\overline{P(\theta)}$. Specifically, N_{wood} , N_{leaf} and N_{hull} are the numbers of wood, leaf and canopy pixels, respectively, and P_{off} , P_{on} , and P_{leaf} were the gap fractions of leaf-off, leaf-on, leaf-only canopy, respectively. Based on Equation 14, P_{leaf} was a simple ratio of P_{off} to P_{on} (Equation 15). With known P_{leaf} , LAI was estimated in a similar manner with PAI using Equation 11-13, except for replacing $\overline{P(\theta)}$ (namely P_{on}) with P_{leaf} .

$$P_{off} = 1 - \frac{N_{wood}}{N_{hull}}, P_{on} = 1 - \frac{N_{wood} + N_{leaf}}{N_{hull}}, P_{leaf} = 1 - \frac{N_{leaf}}{N_{hull} - N_{wood}} \quad (14) \quad 3.$$

$$P_{leaf} = \frac{P_{on}}{P_{off}} \quad (15)$$

Results and Discussion

3.1. Calibration

An example frame among the 70 frames from the camera video is shown in Figure 5a. The original 720p frame is cropped to 1280x500 to be compact. Only part of the 16x9 circle grid is within the field of view. The frame shows no blurry effect, implying that an FPS of 25 is sufficient to match a rotation speed of 2.4° per second. The purple band in the frame is the near-infrared LED light from Leddar, since the camera lens has no IR filter. The visibility of the LED light provides an intuitive way of validating calibration accuracy: the footprints of all the 16 Leddar beams after calibration should fall within the purple area. The ellipses on the wall are detected with the CNED method, shown in green in Figure 5b. It is clear that a few incomplete ellipses are skipped and many redundant ellipses are created. This is a preliminary step of approximating ellipse ranges and locations. Fine ellipses after edge thinning and geometrical fitting are shown in random colors in Figure 5c overlaid by the edge image. Edge noise is inevitable but have limited impact on the ellipse fitting results. Each ellipse center is marked as a green cross. The integer number inside each ellipse is the circle ID predicted with the OCR and posterior voting method. The OCR recognition confidence is also placed under each ellipse ID as a decimal number. Ellipses with low OCR confidence are removed, such as the No.142 and No.80. The remaining 33 ellipses still satisfy the minimal requirement of having four control points for the camera collinearity equation. Note that all the 70 frames have been inspected to have four or more ellipses at the beginning. After calibration, point-based footprints of the Leddar beams are projected on the example frame. The Leddar points basically fall within the LED light zone, except for the first beam on less illuminated area.

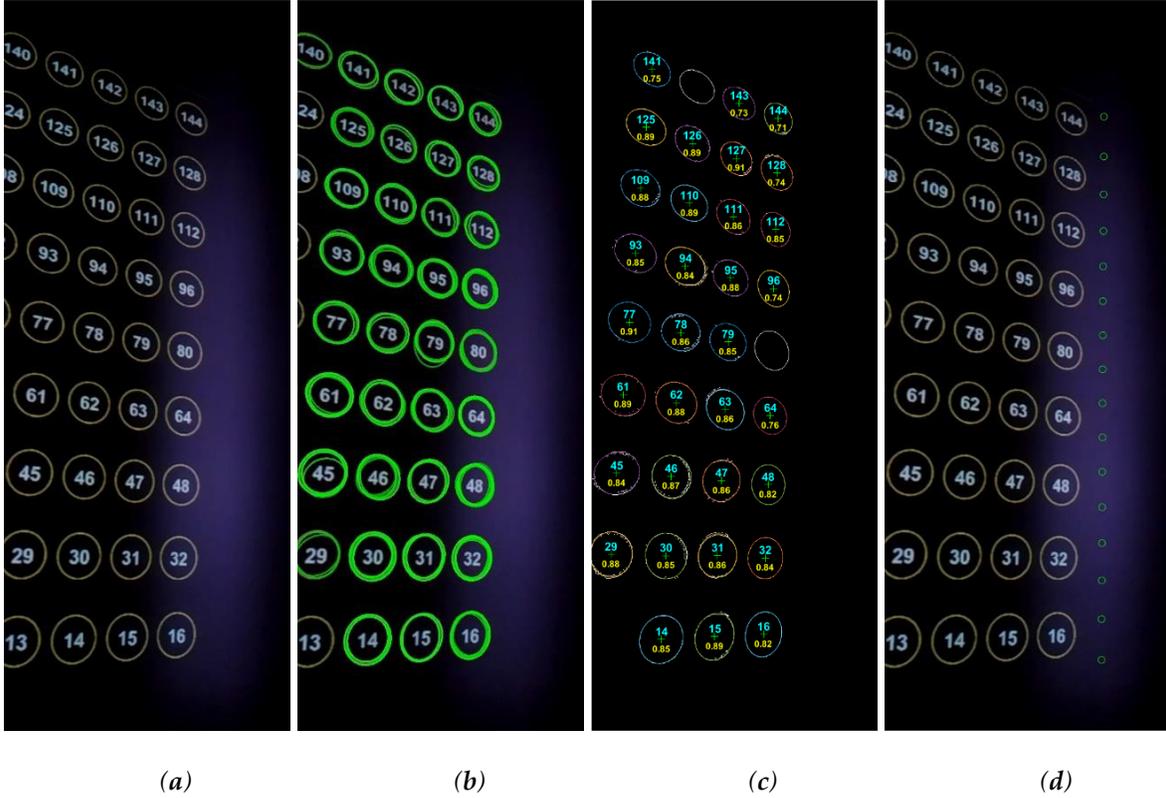


Figure 5. Calibration processing of an example frame: (a) circle grid (yellow) and LED light (purple) from camera view, (b) ellipse detection (green) by CNED method, (c) fine ellipse fitting (random color), ellipse ID from OCR (cyan) and OCR confidence (yellow), and (d) calibrated beam points reprojected to the frame image (green).

The point clouds after calibration of Leddar distances and poses are displayed in Figure 6a, with the X-Y projection approximating a planar shape and the X-Z projection linear shape. The entire point clouds contain 16 beams, each creating 70 points along the sweeping direction. Several points overlap when Leddar is still static at the beginning or the end. The sweeping pattern of each beam on the X-Y plane is not a straight line. This is because Leddar has a fixed tilt angle of about 15° upwards. A few lines are not smooth and their noise is not systematic for all lines, probably less likely due to the servo movement but rather Leddar distance measurement instability. The standard error is 1.03 pixels for optimizing the camera collinearity equation (Equation 6), 3.84 pixels for optimizing the temporal rotation equation (Equation 4 and 6), and finally 9.7 mm in WCS for solving Leddar distance equation (Equation 2, 3 and 5). Equation 4 constrains the rotation matrix to a fixed rotation axis. Without Equation 4, solving Equation 6 for each frame separately is still feasible and yields a standard error of 2.34 pixels. However, the retrieved camera extrinsic parameters, such as the camera center locations T_t shown as the green dots in Figure 6b, lose physical meaning and present irregular movements. Therefore, applying Equation 4 accounts for the real camera movement shown as the white arc points in Figure 6b and thus enables more reliable calibration parameters. The final estimate of intrinsic parameters ($T, D_b, \theta_0, \psi_0, \psi_\Delta, \phi_\alpha, \phi_\beta, \phi_x, \phi_z$) are 0.070 m, -0.428 m, 180.84° , 89.39° , 3.32° , -0.69° , 12.19° , 0.027 m and -0.0006 m.

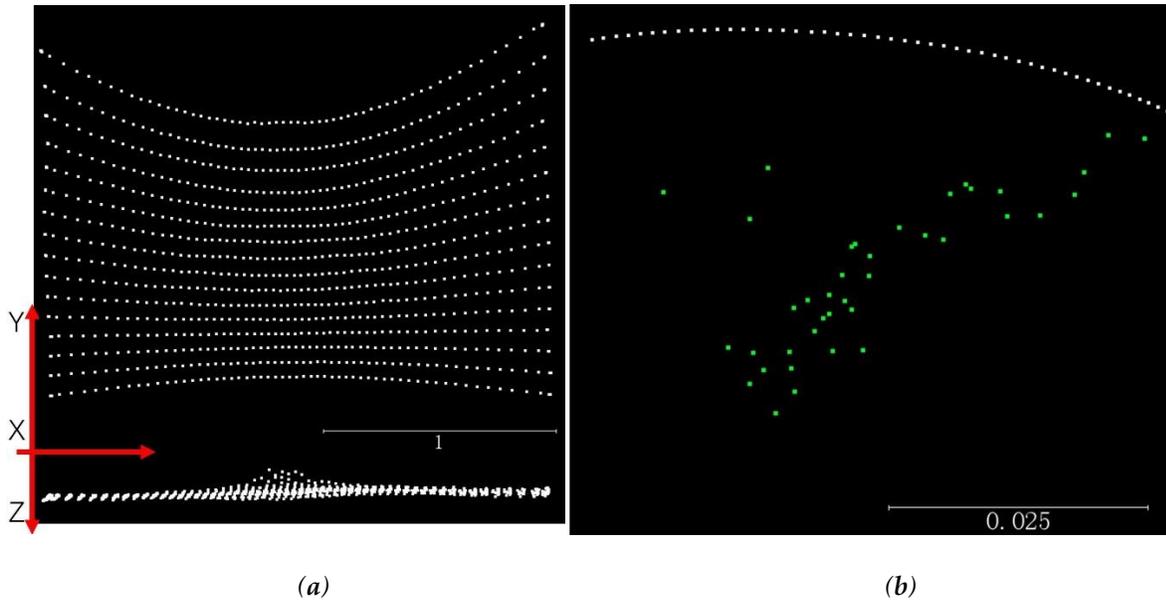


Figure 6. Calibrated Leddar points and camera trajectory: (a) calibrated Leddar points of a flat wall on X-Y plane (above) and on X-Z plane (below), and (b) camera trajectory with rotation constraints (above) and without constraints (green below).

3.2. Fusion-based Dense 3D Recovery

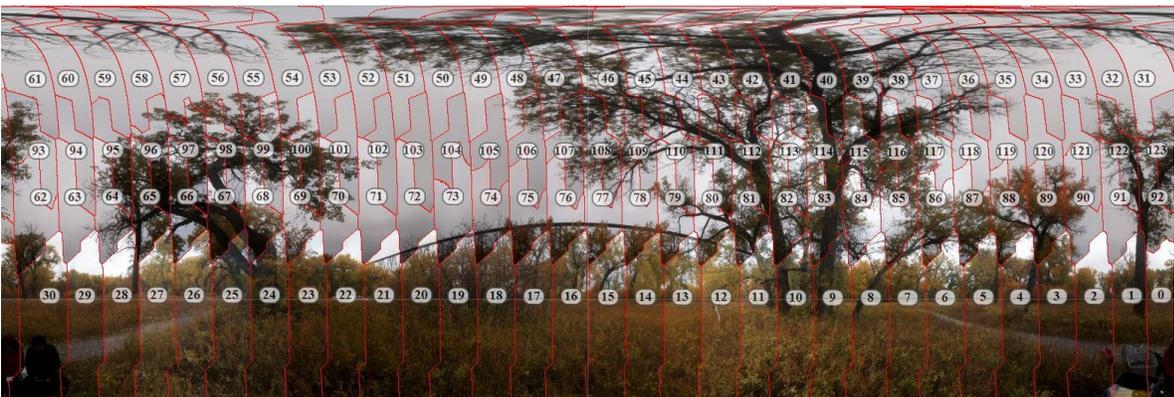
The success of point cloud recovery hinges greatly on the quality of aligning video frames, because camera poses determine the general form and structure of point clouds. Our video scans of poplar trees on four different dates are aligned based on automatic solutions provided from PtGUI, including image matching, feature extraction, feature matching, horizon correction, bundle adjustment and image mosaicking. Example alignment results for the 10/01 and 10/17 videos are visualized as $360^\circ \times 120^\circ$ spherical panoramas under Equirectangular projection in Figure 7a,b. The trees displayed leaf-on conditions on 10/01 and were defoliated completely on 10/17. Both scenes were centered on a railway viaduct and the lower part of the panorama was discarded. No obvious alignment gap or inconsistencies were found from the two images. The processing results of the leaf-on scene are visualized in Figure 7c-h. Figure 7c shows alignment of separate images and their seamlines in PtGUI without mosaicking and color blending. The colors of individual tiles in Figure 7c differ from each other due to sunlight variation during scan. Yet based on visual inspection, the alignment of tiles is seldom affected by the color difference, indicating strong robustness of PtGUI's feature extraction algorithms. The alignment errors estimated from bundle adjustment in PtGUI are 4.5, 3.0, 2.5, 2.5 and 2.2 pixels for the scenes of 09/09, 09/17, 10/01 and 10/17, respectively. The relatively large error of the 09/09 is due to various factors such as windy conditions, thick canopy and cloudy sky.



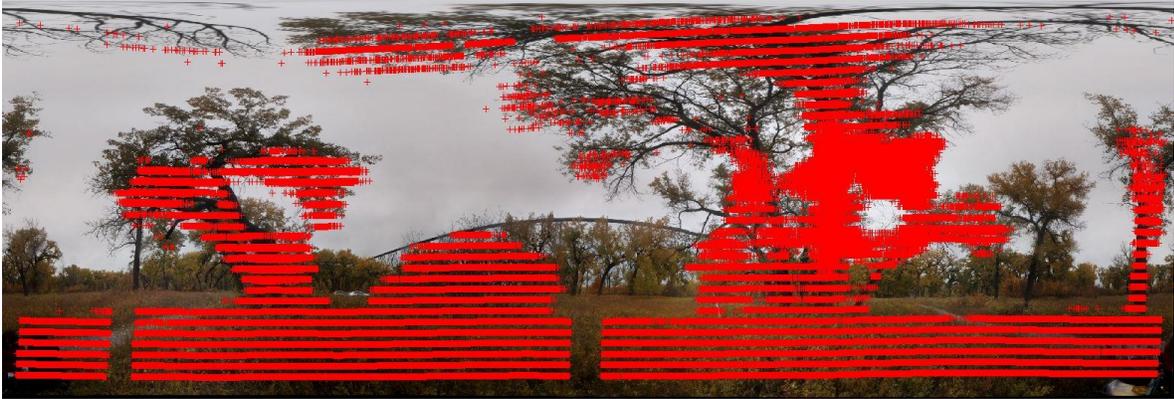
(a)



(b)



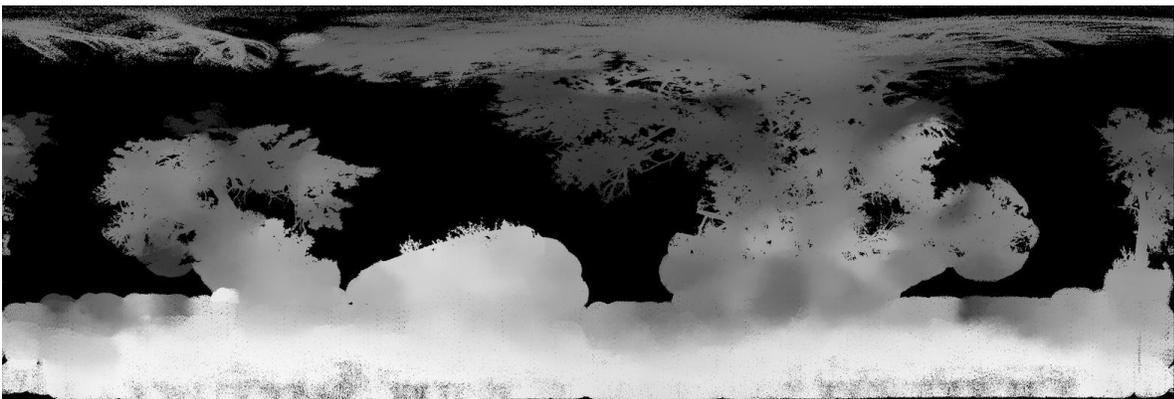
(c)



(d)



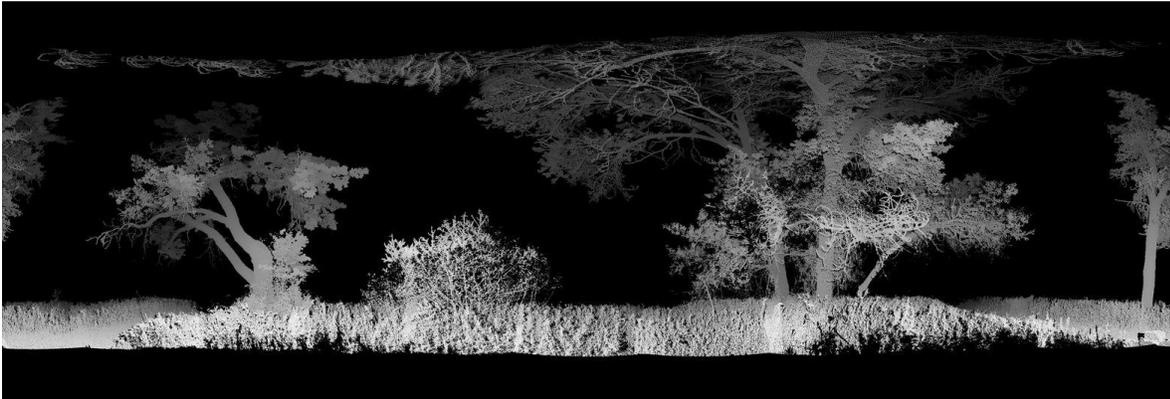
(e)



(f)



(g)



(h)

Figure 7. Hemispherical view of processing results: (a) image global alignment for the 10/01 scan, (b) image global alignment for the 10/17 scan, (c) global alignment layout in PtGUI software with image IDs and seamlines for the 10/01 scan, (d) Leddar-only point clouds (red cross) reprojected to the hemispherical image, (e) RGB colors from fusion-based point clouds, (f) depth image from fusion-based RGB point clouds, (g) depth image from Leddar-only point clouds (point size enlarged for clearer visualization), and (h) depth image from TLS scans. (e), (f), (g) and (h) are all using hemispherical projection.

Leddar points are reprojected as the red crosses in Figure 7d and overlaid with the panorama view of the 10/01 scene, after applying Leddar intrinsic parameters from calibration and camera extrinsic parameters from PtGUI alignment. The Leddar points capture the basic scene structure nearby except for upper canopy, distant ground and thin branches. The minimum, average and maximum detection range of Leddar in this scene is 1.64, 6.45 and 14.17m, respectively. The Leddar point clouds have obvious gaps between the beams and on the ground due to missing signals. This problem of Leddar data sparsity limits potential applications such as tree surveying and object detection, unless photographic information is integrated. Therefore, the iterative bundle adjustment is applied at the point cloud level to minimize the disagreement between Leddar reprojected pixels, camera pixels, and camera extrinsic parameters. Iterations of the bundle adjustment error, measured in pixels, is plotted for the four scanning dates in Figure

8. The initial error of bundle adjustment can be greater than 8 pixels but will converge to a level comparable to PtGUI alignment error. The final error from bundle adjustment is 2.8, 1.9, 2.0 and 1.8 pixels for 09/09, 09/17, 10/01, and 10/17, respectively.

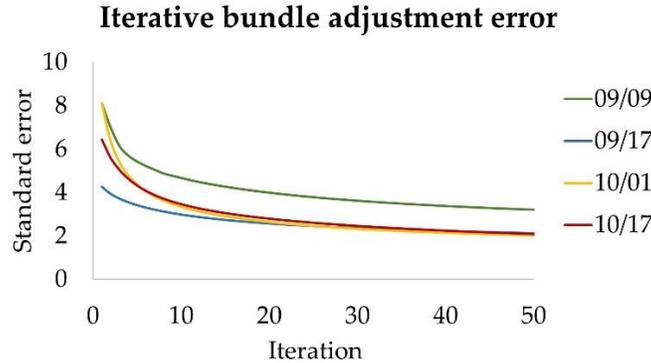


Figure 8. Fusion error convergence with iterations (in pixels) on four defoliating dates in 2018.

The fusion-based point clouds after image background removal, iterative bundle adjustment, and dense matching recovery are reprojected into two panorama images shown as Figure 7e-f. Figure 7e displays reprojected pixels with RGB colors, and Figure 7f is the corresponding depth image with nearer objects showing brighter colors. The reprojection from point clouds to a hemispherical-view image is not simply one point per pixel, considering the previous dense recovery process has a subsampling rate of 10 pixels per point. Therefore, each point has a buffer of 10 pixels in a hemispherical image. Similarly, reprojecting TLS point clouds into a hemispherical depth image in Figure 7h also needs to consider the footprint of each TLS laser beam. The scanning spacing of each Ilris HD beam (1600 μ rad) is set to be the footprint size according to the suggestions in [12]. The scanning spacing corresponds to a constant pixel size of 1.6 and thus each reprojected pixel of the hemispherical images is dilated by a factor of 1.6.

In contrast to the Leddar reprojection image in Figure 7d, the image in Figure 7e not only captures rich 2D details but also covers a reasonable extent due to the region-based interpolation. The main problem of the fusion-based point clouds is false interpolation. The problem can be illustrated when comparing specific tree point clouds extracted from TLS, fusion-based point clouds, and Leddar point clouds in Figure 9a-c. The TLS point clouds clearly exhibit branch-level details with warmer colors representing higher laser intensity. The fusion-based point clouds have distinguishable stem colors and noisy branches, still highly detailed compared to the obscure Leddar point clouds. Yet the fusion-based point clouds overfill the gaps between branches and also falsely incorporate pixels from remote shrubs. This is inevitable since region-based interpolation and bundle adjustment can mitigate but not eradicate the problem of coarse and sparse depth measurement from Leddar. The depth image in Figure 7f displays a strong smoothing effect compared to the depth image reprojected from TLS in Figure 7h, but are much more detailed than the Leddar-only point clouds in Figure 7g with indiscernible sparse points.

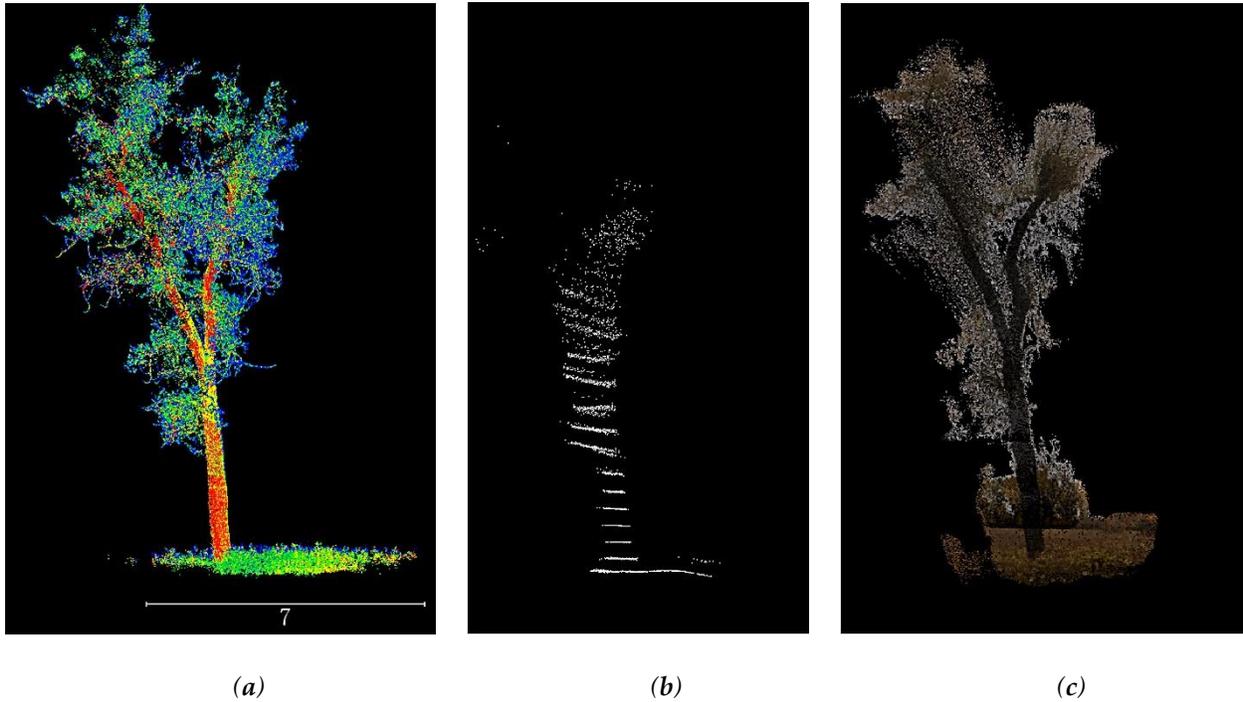


Figure 9. Example tree point clouds from (a) TLS scans, (b) Leddar-only point clouds and (c). fusion-based dense point clouds.

3.3. Tracking Changes of Canopy Vertical Volume Profile, PAI and LAI

Vertical volume profiles from TLS, fusion-based point clouds, and Leddar point clouds can be contrasted with Figure 10a-d. Both profiles from the fusion-based point clouds and the Leddar point clouds are correlated with the TLS profiles, regardless of scanning date. The r^2 of profiles over the maximum tree height range between Leddar and TLS stays around 0.3 from the first three leaf-on scenes and jumps to 0.48 from the last leaf-off scene. In contrast, the r^2 of profiles between the fusion-based and TLS are around 0.65 from the leaf-on scenes, constantly higher than 0.52 from the leaf-off scene. This is because thick crowns and leafy understory lead to Leddar signal loss but do not affect photography-based interpolation. With defoliated trees in the last scene, photography-based interpolation tends to incorporate more false pixels from areas beyond the canopy, thus resulting in lower accuracy. It is also noteworthy that the fusion-based point clouds demonstrate higher frequency than TLS point clouds in the middle crown area due to the overfilling effect, but also thinner near upper crown due to the loss of Leddar points. Note that the TLS has a slightly narrower scanning view than the FSS, with part of upper crown and ground not sampled by scans. The profile difference around upper crown and ground can be greater than observed in Figure 10. This problem of profile distortion might be due to the imperfect hemispherical stitching process.

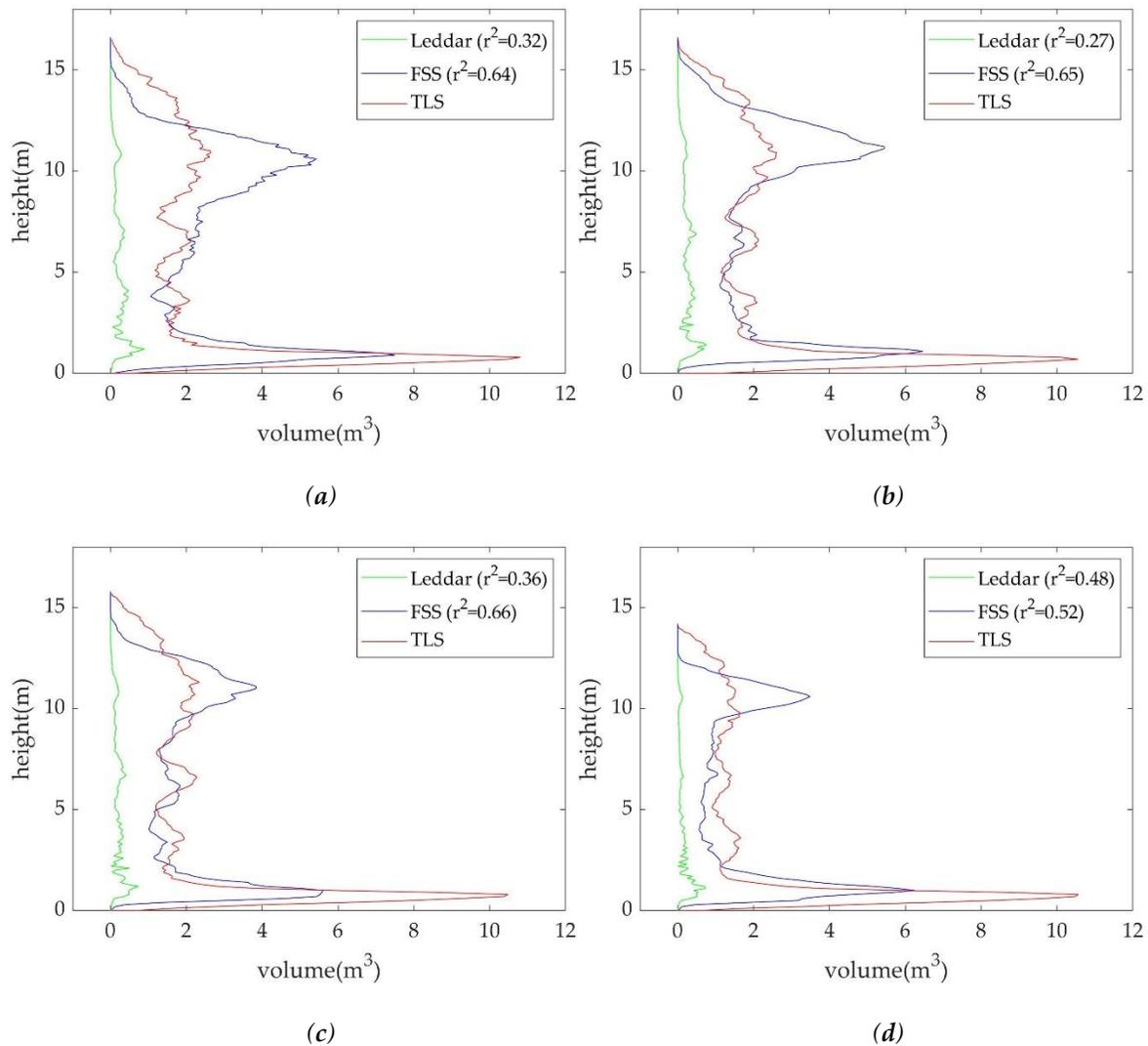


Figure 10. Vertical volume profiles from TLS, fusion-based and Leddar-only point clouds on (a) 09/09, (b) 09/17, (c) 10/01, and (d) 10/17, 2018, where horizontal axis denotes volume of voxels with a unit voxel of $0.1 \times 0.1 \times 0.1 \text{ m}^3$, and vertical axis denotes height in meters.

The benefits of synthesizing both 3D and color information make FSS a potentially valuable complement to conventional LAI or PAI surveying tools, such as digital hemispherical photography (DHP). Figure 11 compares the fisheye image from FSS with the DHP photo from the same site. The canopy shapes between the two images are visually identical, seen from Figure 11a and Figure 11b. The FSS, in addition, captures depth information such as the image in Figure 11c, with a benchmarking TLS depth image provided in Figure 11d. Note that the upper crown area was not scanned with the TLS due to the field of view constraint. The availability of depth images enables FSS to calculate true PAI and LAI based on the PATH model. Different PAI and LAI estimates based on FSS, TLS and DHP methods, and based on PATH and non-PATH methods are contrasted in Figure 12, with bars denoting PAI and crosses denoting LAI. The non-PATH method relies on the LAI models from the Hemisfer software [50,51] which combines Lang

[52]'s LAD (leaf angle distribution) function, Lang and Xiang [53]'s clumping correction, and Schleppi, Conedera, Sedivy and Thimonier [50] non-linearity correction model. The non-PATH method focuses on the RGB images from DHP or FSS, or the depth images from TLS, whereas the PATH method additionally needs point cloud input from FSS or TLS.

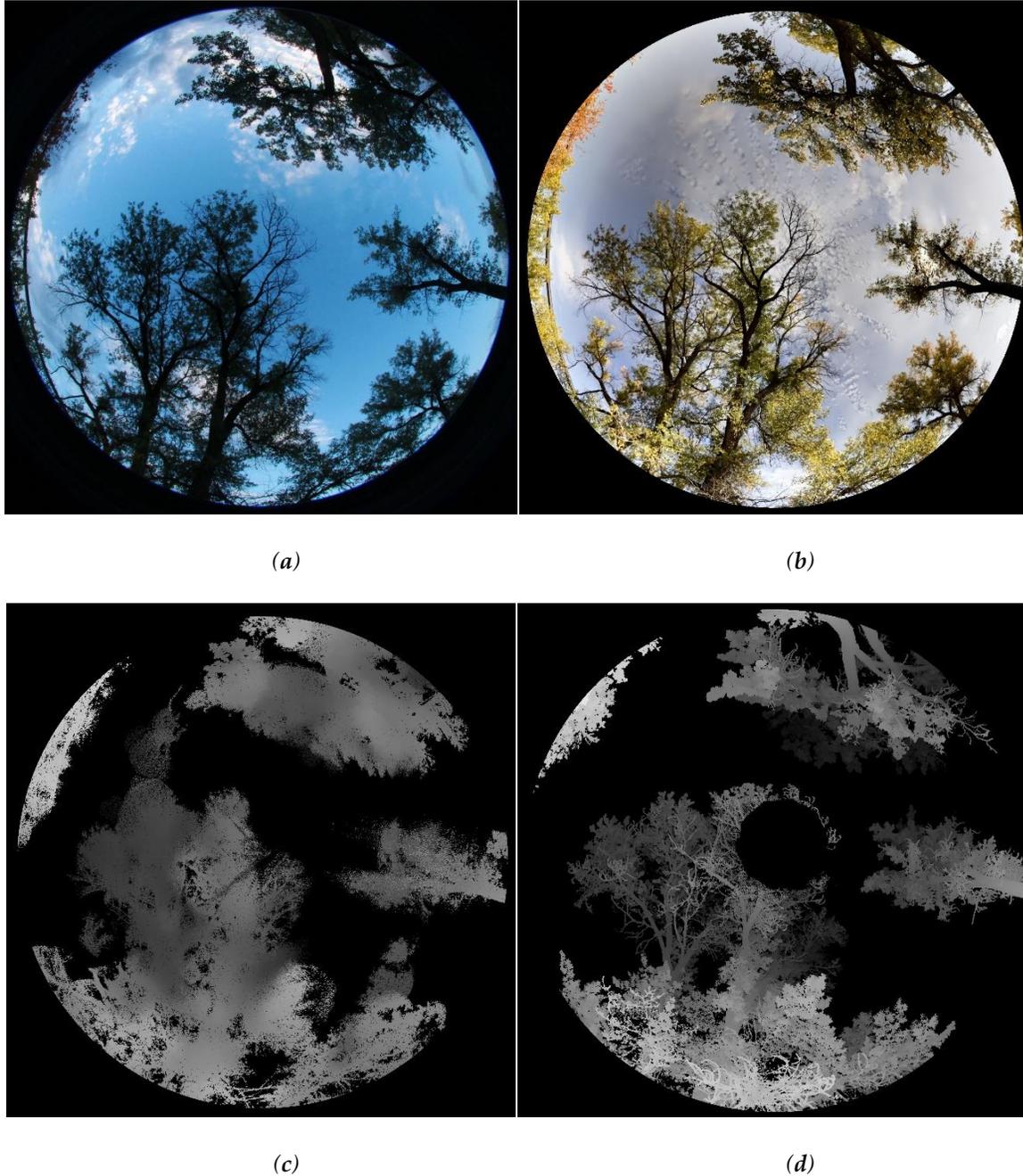


Figure 11. Fisheye-view images compiled from the 09/09 datasets based on (a) DHP, (b) FSS, (c) FSS depth, and (d) TLS depth.

For the non-PATH methods in Figure 12, PAI and LAI values generally decline with the defoliation dates, with all the LAI values reaching zero level on the leaf-off date 10/10, except that the PAI and LAI value from non-PATH FSS increases on 10/01. The incorrect increase implies the instability of using image-only methods. The possible cause of the incorrect increase is FSS's underestimation of PAI and LAI from the 09/09 and 10/01 FSS images, in contrast to the PAI and LAI values from DHP and TLS. Strong spectral reflectance from the sunlight are observed in the 09/09 and 10/01 FSS images and a small portion of canopy pixels in the FSS image displays a similar color as the sky background. These bright canopy pixels were not successfully identified as leaf area, causing the underestimation effect. The DHP method does not have the underestimation issue since the DHP images were captured near dusk. The TLS method does not have the stability issue as FSS, since depth images are used instead of color images. The TLS method, however, has an issue of overestimating PAI. The leaf-off PAI from TLS is 32% higher than from DHP, compared to the average 3% overestimation of leaf-on PAI from TLS. The overestimation issue of TLS has two typical causes. The depth images, especially the leaf-off ones, contain ghost points or misaligned points around thin branches and twigs. Gaps smaller than the beam width of TLS are also not differentiable from the depth images [12]. The TLS's overestimation of twig PAI leads to an underestimation of LAI by 26% relative to DHP.

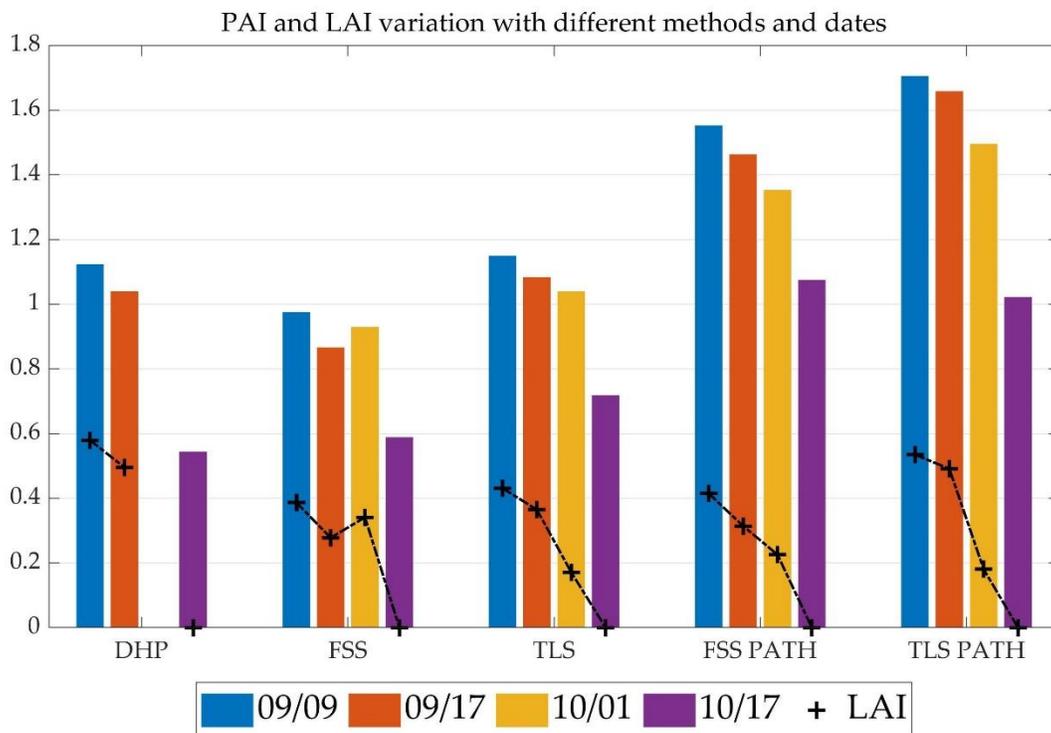


Figure 12. Comparing different methods of LAI estimation on four scanning dates. The colored bars represent PAI, and crosses for the associated LAI. The 10/01 DHP dataset is not available. The DHP and the FSS methods are based on RGB images and the TLS based on depth images.

With the PATH model applied to TLS and FSS, the PAI estimates are about 30-45% greater than the non-PATH PAIs. The PATH PAI estimation from FSS does not have the problem of PAI increase on 10/01, indicating the importance of incorporating depth correction. The PATH LAI estimates are also greater than

the non-PATH by 14% on average, except for the 10/01 FSS LAI anomaly. Considering the optical image methods usually underestimate true PAI or LAI [7] by 20%-60%, it is assumed that the PATH model is a closer approximation of the true PAI or LAI values.

It is important to understand why the PATH model usually outputs higher PAI (or LAI) values than the classic geometrical-optical model. Indeed, the $PAI_{true}(\theta)$ solved from the PATH model does not have a simple analytic form, due to various forms of p_l . However, if we simply assume p_l is constantly 1, or equivalently, the within-crown path length distribution is uniform, the PATH model then has an analytic solution of PAI, which is basically a Lambert W function of gap fraction $\overline{P}(\theta)$ (Figure 13). The traditional LAI model (effective LAI) using Beer's law is also contrasted in Figure 13. It clearly shows that the PATH PAI is consistently greater than non-PATH PAI, especially when gap fraction is small. It is also noteworthy that the PATH model might be overly sensitive to near-zero gap fraction changes. The upper bound of PAI based on the PATH model is $-\frac{\cos(\theta)}{G(\theta)} \frac{l_{max}}{l_{min}} \ln \overline{P}(\theta)$ and lower bound is $-\frac{\cos(\theta)}{G(\theta)} \frac{l_{min}}{l_{max}} \ln \overline{P}(\theta)$. The wide ranges of PAI indicates strong flexibility of the PATH model, but it is also important for future studies to examine what are the rigorous PAI bounds based on different forms of the p_l functions, and what can be a suitable analytical form of p_l functions or PAI functions with smoother gap fraction sensitivity.

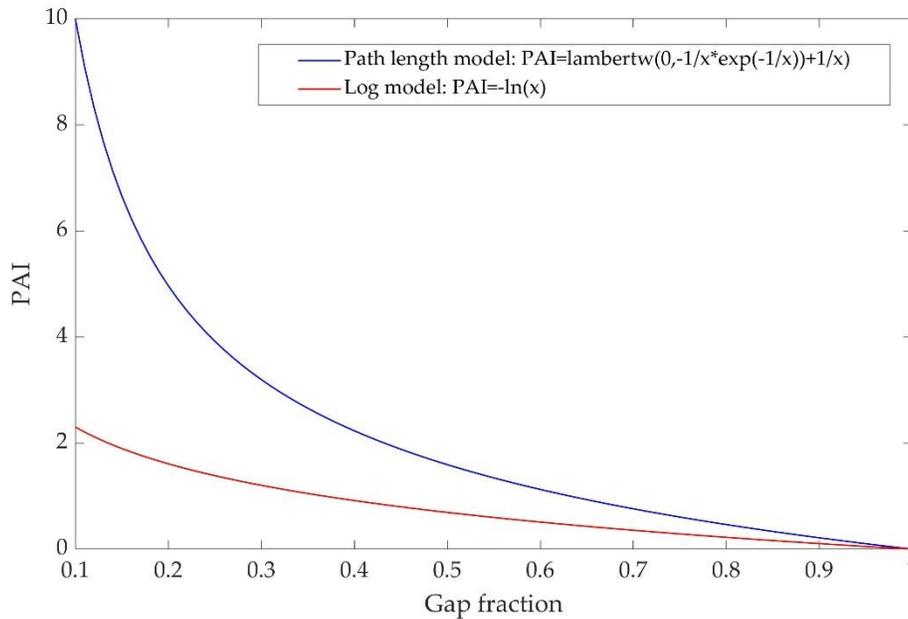


Figure 13. Relationship between gap fraction (x) and PAI (or LAI). The red curve shows PAI values from the path length distribution model, compared to the blue curve from the simple Beer's law model.

4. Conclusions

Timely monitoring of canopy characteristics is needed to understand the spatiotemporal variation of biomass in a forest ecosystem, and to evaluate carbon budgets as part of forest stand reporting. The advent of low-cost multi-beam LiDAR sensors, Leddar in particular, has presented many successful object tracking applications. Yet the Leddar sensor is not comparable to TLS in sampling 3D details due to its limited FOV and point resolution. This limitation was mitigated in this study by constructing a low-cost 3D fusion scanning system, FSS, integrating Leddar, camera and pan-tilt robotics. A framework of integration was developed, generally including (1) plane-based physical calibration converting Leddar distance into 3D

point and locating Leddar point from images, (2) global image alignment obtaining panorama and coarse camera poses, (3) iterative bundle adjustment optimizing camera poses using both Leddar distance and corresponding pixels at point cloud level, and (4) dense point cloud recovery based on dense matching and interpolation. The calibration error of Leddar points was 9.7 mm at a distance of ~ 1 m. The set of fusion-based methods was applied to recover hemispherical colored point clouds from multi-temporal poplar tree scans during the autumn defoliation period. The bundle adjustment error was 1-3 pixels, indicating strong agreement between image and Leddar projection on the X-Y plane. Yet uncertainty exists on the depth (Z) direction due to the coarse resolution of Leddar distance. Final fusion-based point clouds were compared to the TLS scans collected on the same spot and date. The vertical profile volumes between TLS and FSS point clouds had an r^2 of 0.5-0.7 over the maximum tree height range, which varies with leaf cover conditions and exceeds the r^2 of 0.3-0.5 between TLS and pure Leddar point clouds. PAI and LAI metrics were also extracted from FSS, TLS and DHP for leaf-on and leaf-off dates. Using only image data, the PAI and LAI tended to be underestimated with FSS and overestimated with TLS. With the point cloud PATH model, both PAI and LAI from FSS or TLS were corrected to approach their assumed true values. By combining both color and depth information, the FSS has demonstrated versatility and great potential for the application of canopy foliage monitoring.

The FSS system was primarily developed for a static scanning of environment. For environmental applications such as crown measurement and biomass delineation, low-cost sensor systems such as the FSS (as built) does not parallel TLS and DHP's resolution or precision at present, yet the demand for gross mensuration should not be overlooked and sensor hardware upgrading is inevitable. The core contribution of this study is a holistic calibrating and fusing scheme for a low resolution multi-sensor platform. The advantage of FSS should be clear. It is more portable and low-cost compared to a conventional TLS, and has a higher frequency, greater detection range and FOV compared to many indoor-oriented LED or flash LiDAR systems. It is therefore suitable to be deployed in high numbers into sensor networks for broad-scale environmental monitoring. Adapting the FSS to other mobile system such as UAV is also feasible, following a similar framework of calibration, pose approximation, bundle adjustment and densification, except the need for external pose information from GPS and IMU and a dedicated image alignment method. A better viewing geometry such as stereoscopy from a mobile system could greatly improve the precision of 3D recovery especially in the depth direction. A bright future for cost-effective fusion-based 3D canopy monitoring systems is anticipated.

Author Contributions: conceptualization, C.H. and Z.X.; methodology, Z.X.; formal analysis, Z.X.; investigation, C.H., S.R. and Z.X.; data curation, C.B., F.X., D.P., E.J., and Z.X.; writing—original draft preparation, Z.X.; writing—review and editing, C.H.; supervision, C.H.; project administration, C.H.; funding acquisition, C.H.

Funding: This research was funded by the S.G.S. International Tuition Award and the Dean's Scholarship from the University of Lethbridge, Campus Alberta Innovates Program (CAIP), and NSERC Discovery Grants Program.

Acknowledgments: Zhouxin Xi would like to thank Laura Chasmer, Derek Peddle and Craig Coburn from the University of Lethbridge, and Richard Fournier from the Université de Sherbrooke for the invaluable comments and supports.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. van der Sande, M.T.; Zuidema, P.A.; Sterck, F. Explaining biomass growth of tropical canopy trees: the importance of sapwood. *Oecologia* **2015**, *177*, 1145-1155, doi:10.1007/s00442-015-3220-y.
2. Sumida, A.; Watanabe, T.; Miyaura, T. Interannual variability of leaf area index of an evergreen conifer stand was affected by carry-over effects from recent climate conditions. *Scientific Reports* **2018**, *8*, doi:10.1038/s41598-018-31672-3.
3. Kim, J.; Ryu, Y.; Jiang, C.; Hwang, Y. Continuous observation of vegetation canopy dynamics using an integrated low-cost, near-surface remote sensing system. *Agricultural and Forest Meteorology* **2019**, *264*, 164-177, doi:<https://doi.org/10.1016/j.agrformet.2018.09.014>.
4. de Wit, C.T. *Photosynthesis of leaf canopies*; Pudoc: 1965.
5. Ross, J. *The radiation regime and architecture of plant stands*; Junk Publishers: The Hague, 1981.
6. Zheng, G.; Moskal, L.M. Retrieving leaf area index (LAI) using remote sensing: Theories, methods and sensors. *Sensors* **2009**, *9*, 2719-2745.
7. Yan, G.; Hu, R.; Luo, J.; Weiss, M.; Jiang, H.; Mu, X.; Xie, D.; Zhang, W. Review of indirect optical measurements of leaf area index: Recent advances, challenges, and perspectives. *Agricultural and Forest Meteorology* **2019**, *265*, 390-411, doi:<https://doi.org/10.1016/j.agrformet.2018.11.033>.
8. Zhao, K.; García, M.; Liu, S.; Guo, Q.; Chen, G.; Zhang, X.; Zhou, Y.; Meng, X. Terrestrial lidar remote sensing of forests: Maximum likelihood estimates of canopy profile, leaf area index, and leaf angle distribution. *Agricultural and Forest Meteorology* **2015**, *209-210*, 100-113, doi:<http://dx.doi.org/10.1016/j.agrformet.2015.03.008>.
9. Li, Y.; Su, Y.; Hu, T.; Xu, G.; Guo, Q. Retrieving 2-D leaf angle distributions for deciduous trees from terrestrial laser scanner data. *IEEE Transactions on Geoscience and Remote Sensing* **2018**, *56*, 4945-4955.
10. Zhu, X.; Skidmore, A.K.; Wang, T.; Liu, J.; Darvishzadeh, R.; Shi, Y.; Premier, J.; Heurich, M. Improving leaf area index (LAI) estimation by correcting for clumping and woody effects using terrestrial laser scanning. *Agricultural and Forest Meteorology* **2018**, *263*, 276-286, doi:<https://doi.org/10.1016/j.agrformet.2018.08.026>.
11. Hopkinson, C.; Lovell, J.; Chasmer, L.; Jupp, D.; Kljun, N.; van Gorsel, E. Integrating terrestrial and airborne lidar to calibrate a 3D canopy model of effective leaf area index. *Remote Sensing of Environment* **2013**, *136*, 301-314.
12. Hancock, S.; Essery, R.; Reid, T.; Carle, J.; Baxter, R.; Rutter, N.; Huntley, B. Characterising forest gap fraction with terrestrial lidar and photography: An examination of relative limitations. *Agricultural and Forest Meteorology* **2014**, *189*, 105-114.
13. Calders, K.; Armston, J.; Newnham, G.; Herold, M.; Goodwin, N. Implications of sensor configuration and topography on vertical plant profiles derived from terrestrial LiDAR. *Agricultural and Forest Meteorology* **2014**, *194*, 104-117, doi:<http://dx.doi.org/10.1016/j.agrformet.2014.03.022>.
14. Hu, R.; Yan, G.; Mu, X.; Luo, J. Indirect measurement of leaf area index on the basis of path length distribution. *Remote Sensing of Environment* **2014**, *155*, 239-247, doi:<https://doi.org/10.1016/j.rse.2014.08.032>.
15. Hu, R.; Bournez, E.; Cheng, S.; Jiang, H.; Nerry, F.; Landes, T.; Saudreau, M.; Kastendeuch, P.; Najjar, G.; Colin, J., et al. Estimating the leaf area of an individual tree in urban areas using terrestrial laser scanner and path length distribution model. *ISPRS Journal of Photogrammetry and Remote Sensing* **2018**, *144*, 357-368, doi:<https://doi.org/10.1016/j.isprsjprs.2018.07.015>.
16. Zhu, Z.; Liu, J. Unsupervised extrinsic parameters calibration for multi-bem LiDARs. In Proceedings of Proc. 2nd Int. Conf. Computer Science and Electronics Engineering; pp. 1110-1113.
17. Olivier, P. *Leddar Optical Time-of-Flight Sensing Technology*. Engineering and Manufacturing, L.I., Ed. Quebec City, Quebec, 2015; p 13.
18. Gangadharan, S.; Burks, T.F.; Schueller, J.K. A comparison of approaches for citrus canopy profile generation using ultrasonic and Leddar® sensors. *Computers and Electronics in Agriculture* **2019**, *156*, 71-83.

19. Arnay, R.; Hernández-Aceituno, J.; Toledo, J.; Acosta, L. Laser and Optical Flow Fusion for a Non-Intrusive Obstacle Detection System on an Intelligent Wheelchair. *IEEE Sensors Journal* **2018**, *18*, 3799-3805.
20. Mimeault, Y.; Cantin, D. Lighting system with driver assistance capabilities. Google Patents: 2013.
21. Godejord, B. Characterization of a commercial LIDAR module for use in camera triggering system. NTNU, 2018.
22. Thakur, R. Scanning LIDAR in Advanced Driver Assistance Systems and Beyond: Building a road map for next-generation LIDAR technology. *IEEE Consumer Electronics Magazine* **2016**, *5*, 48-54.
23. Mimeault, Y. Parking management system and method using lighting system. Google Patents: 2014.
24. Hentschke, M.; Pignaton de Freitas, E.; Hennig, C.; Girardi da Veiga, I. Evaluation of Altitude Sensors for a Crop Spraying Drone. *Drones* **2018**, *2*, 25.
25. Elaksher, A.F.; Bhandari, S.; Carreon-Limones, C.A.; Lauf, R. Potential of UAV lidar systems for geospatial mapping. In Proceedings of Lidar Remote Sensing for Environmental Monitoring 2017; p. 104060L.
26. Bohren, J.; Foote, T.; Keller, J.; Kushleyev, A.; Lee, D.; Stewart, A.; Vernaza, P.; Derenick, J.; Spletzer, J.; Satterfield, B. Little ben: The ben franklin racing team's entry in the 2007 DARPA urban challenge. *Journal of Field Robotics* **2008**, *25*, 598-614.
27. Muhammad, N.; Lacroix, S. Calibration of a rotating multi-beam lidar. In Proceedings of Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on; pp. 5648-5653.
28. Atanacio-Jiménez, G.; González-Barbosa, J.-J.; Hurtado-Ramos, J.B.; Ornelas-Rodríguez, F.J.; Jiménez-Hernández, H.; García-Ramírez, T.; González-Barbosa, R. LIDAR velodyne HDL-64E calibration using pattern planes. *International Journal of Advanced Robotic Systems* **2011**, *8*, 59.
29. Levinson, J.; Thrun, S. Unsupervised Calibration for Multi-beam Lasers. In Proceedings of Experimental Robotics: The 12th International Symposium on Experimental Robotics; p. 179.
30. Sheehan, M.; Harrison, A.; Newman, P. Self-calibration for a 3D laser. *The International Journal of Robotics Research* **2012**, *31*, 675-687.
31. Li, J.; He, X.; Li, J. 2D LiDAR and camera fusion in 3D modeling of indoor environment. In Proceedings of Aerospace and Electronics Conference (NAECON), 2015 National; pp. 379-383.
32. Budge, S.E.; Badamkar, N.S.; Xie, X. Automatic registration of fused lidar/digital imagery (texel images) for three-dimensional image creation. *OPTICE* **2014**, *54*, 031105-031105, doi:10.1117/1.OE.54.3.031105.
33. Bodensteiner, C.; Hübner, W.; Jüngling, K.; Solbrig, P.; Arens, M. Monocular camera trajectory optimization using LiDAR data. In Proceedings of Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on; pp. 2018-2025.
34. Park, Y.; Yun, S.; Won, C.S.; Cho, K.; Um, K.; Sim, S. Calibration between color camera and 3D LIDAR instruments with a polygonal planar board. *Sensors* **2014**, *14*, 5333-5353.
35. Zhou, L.; Deng, Z. Extrinsic calibration of a camera and a lidar based on decoupling the rotation from the translation. In Proceedings of Intelligent Vehicles Symposium (IV), 2012 IEEE; pp. 642-648.
36. Fremont, V.; Bonnifait, P. Extrinsic calibration between a multi-layer lidar and a camera. In Proceedings of Multisensor Fusion and Integration for Intelligent Systems, 2008. MFI 2008. IEEE International Conference on; pp. 214-219.
37. Debattisti, S.; Mazzei, L.; Panciroli, M. Automated extrinsic laser and camera inter-calibration using triangular targets. In Proceedings of Intelligent Vehicles Symposium (IV), 2013 IEEE; pp. 696-701.
38. Hartley, R.; Zisserman, A. *Multiple view geometry in computer vision*; Cambridge university press: 2003.
39. De Silva, V.; Roche, J.; Kondoz, A. Robust fusion of LiDAR and wide-angle camera data for autonomous mobile robots. *Sensors* **2018**, *18*, 2730.
40. Jia, Q.; Fan, X.; Luo, Z.; Song, L.; Qiu, T. A fast ellipse detector using projective invariant pruning. *IEEE Transactions on Image Processing* **2017**, *26*, 3665-3679.

41. Fitzgibbon, A.W.; Pilu, M.; Fisher, R.B. Direct least squares fitting of ellipses. In Proceedings of Proceedings of 13th International Conference on Pattern Recognition, 25-29 Aug. 1996; pp. 253-257 vol.251.
42. Newton, I. *De analysi per aequationes numero terminorum infinitas*; 1711.
43. Newton, I.; Colson, J. *The Method of Fluxions and Infinite Series; with Its Application to the Geometry of Curve-lines... Translated from the Author's Latin Original Not Yet Made Publick. To which is Subjoin'd a Perpetual Comment Upon the Whole Work... by J. Colson*; 1736.
44. Ypma, T.J. Historical development of the Newton–Raphson method. *SIAM review* **1995**, *37*, 531-551.
45. Huber, P.J. *Robust statistics*; Wiley: New York, 1981; pp. ix, 308 p.
46. Bay, H.; Tuytelaars, T.; Van Gool, L. Surf: Speeded up robust features. In Proceedings of European conference on computer vision; pp. 404-417.
47. Nock, R.; Nielsen, F. Statistical region merging. *IEEE Transactions on pattern analysis and machine intelligence* **2004**, *26*, 1452-1458.
48. Hoerl, A.E.; Kennard, R.W. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* **1970**, *12*, 55-67.
49. Chen, Y.; Zhang, W.; Hu, R.; Qi, J.; Shao, J.; Li, D.; Wan, P.; Qiao, C.; Shen, A.; Yan, G. Estimation of forest leaf area index using terrestrial laser scanning data and path length distribution model in open-canopy forests. *Agricultural and Forest Meteorology* **2018**, *263*, 323-333, doi:<https://doi.org/10.1016/j.agrformet.2018.09.006>.
50. Schleppei, P.; Conedera, M.; Sedivy, I.; Thimonier, A. Correcting non-linearity and slope effects in the estimation of the leaf area index of forests from hemispherical photographs. *Agricultural and Forest Meteorology* **2007**, *144*, 236-242.
51. Thimonier, A.; Sedivy, I.; Schleppei, P. Estimating leaf area index in different types of mature forest stands in Switzerland: a comparison of methods. *European Journal of Forest Research* **2010**, *129*, 543-562.
52. Lang, A. Simplified estimate of leaf area index from transmittance of the sun's beam. *Agricultural and Forest Meteorology* **1987**, *41*, 179-186.
53. Lang, A.; Xiang, Y. Estimation of leaf area index from transmission of direct sunlight in discontinuous canopies. *Agricultural and Forest Meteorology* **1986**, *37*, 229-243.

Appendix 10: Deriving Climate Change Projections from Regional Climate Models

Sauchyn, D. 2020. DERIVING CLIMATE CHANGE PROJECTIONS FROM REGIONAL CLIMATE MODELS. Final Sub Report to the AIEES WIP. Prairie Adaptation Research Collaborative, University of Regina

Deriving Climate Change Projections from Regional Climate Models

Final Sub Report - April 2020
Prairie Adaptation Research Collaborative, University of Regina



The objective of the collaborative research project was to separate climatic and anthropogenic influences on the Castle River watershed through archive and contemporary image analysis, field studies and model simulation. Researchers from the Prairie Adaptation Research Collaborative (PARC), at the University of Regina, contributed to evaluating the hydrological response of the montane to alpine transition zones of the Castle headwaters. This contribution was achieved by delivering 1) data sets required for hydrological simulations, and 2) reconstructing the flow of the Castle River over past centuries from tree-ring chronologies. PARC also provided a full weather station (wind speed and direction, wind gusts, solar radiation, water content, air temperature, relative humidity, precipitation) at no capital cost. This weather station had been deployed on the Waldron Ranch between Lundbreck and Longview. It was disassembled and transported to the University of Lethbridge field research station near the Castle Mountain Ski Resort.

Deriving Climate Change Projections from Regional Climate Models

PARC staff generated climate change projections in support of the research of Dr. Stefan Kienzle, who simulated the hydrology of the Castle River watershed using the ACRU model. For future projections of basin hydrology, Dr. Kienzle ran the ACRU model using future values of the following daily climate variables:

- Mean temperature (°C)
- Maximum temperature (°C)
- Minimum temperature (°C)
- Total precipitation (mm)
- Average relative humidity (%)
- Total solar radiation (MJ/m²)
- Average wind speed (m/s)

PARC Research Associate Ms. Yuliya Andreichuk provided Dr. Kienzle with this future climatology. She created scripts for extracting and processing raw data from 10 Regional Climate Models (RCMs), which simulate climate at much high resolution (25-50 km) than Global Climate Models (GCMs), which are the most common source of climate change scenarios. Because RCMs have limited geographic extent (e.g., North America), boundary conditions are derived from a GCM. Thus a run of a RCM represents the dynamically downscaling of GCM output. From the data repository the North American Regional Climate Change Assessment Program (NARCCAP), we downloaded output from nine RCM_gcm pairs:

- CRCM_ccsm
- CRCM_cgcm3
- ECP2_gfdl
- HRM3_gfdl
- HRM3_hadcm3
- MM5I_ccsm
- MM5I_hadcm3
- RCM3_cgcm3
- RCM3_gfdl

All the NARCAAP experiments were based on the SRES A2 emission scenario (at the higher end of the range of SRES scenarios). Barrow and Sauchyn (2017) previously evaluated the capacity of these RCMS to simulate the historical climate of western Canada. On the basis of their analysis, Yuliya developed an algorithm and scripts for bias correction of the RCM data. Yuliya supplemented the NARCCAP RCM data with climate projections from a run version 4 of the Canadian Regional Climate Model (CRCM4). The emission pathway was RCP 8.5 and boundary conditions were from the Community Earth System Model (CESM). Output from CRCM4 is at a resolution of 25 km as illustrated in Figure 1 for the Oldman River Basin.

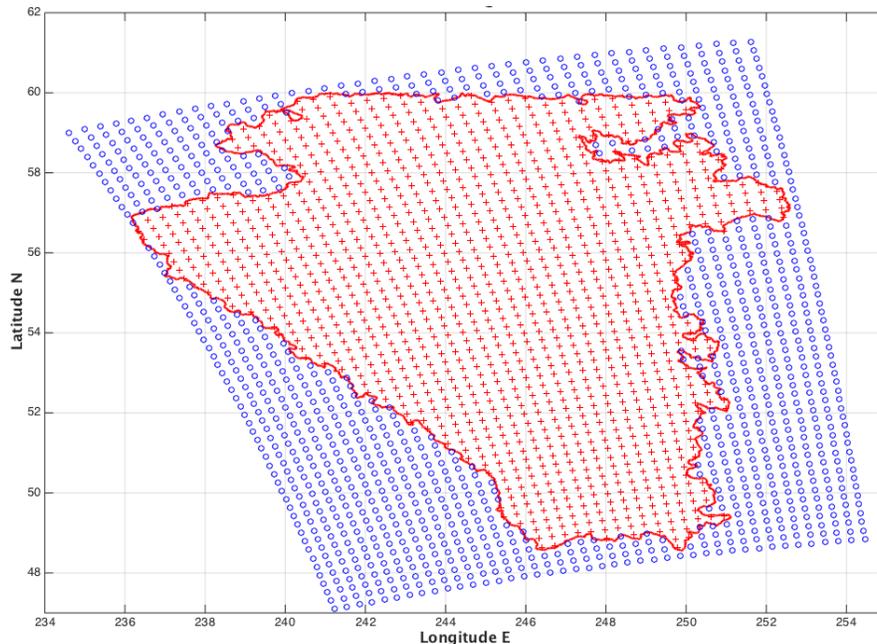


Figure 1: The boundary of the Oldman River Basin superimposed on the 25 km grid of the Canadian

Regional Climate Model version 4.

Paleohydrology of the Castle River Basin

The projection of future climate, and therefore hydrology, is subject to considerable uncertainty. Barrow and Sauchyn (2019) determined that in western Canada the major source of this uncertainty is the natural internal variability of the regional climate regime. Discrepancies among climate models and greenhouse gas emission scenarios are responsible for relatively less uncertainty. Historical weather data record natural variability but also the influence of anthropogenic climate change. Tree rings are a proxy source of pre-industrial climate data of annual and seasonal resolution.

PARC PhD student Ms. Samantha Kerr investigated the relationship between tree ring-width chronologies from moisture-sensitive trees in and near the Castle River basin, instrumental streamflow records, and total basin runoff simulated by the Canadian Regional Climate Model version 5 (CRCM5). Her research is documented in a separate report (Kerr, 2018). The mean annual flow of the Castle River at Beaver Mines was reconstructed from 1150 to 2010 using tree-ring data from multiple species and proxies (i.e. earlywood, latewood, and annual ring-width). The proxy streamflow data were downscaled stochastically to weekly estimates using methods developed by Sauchyn and Ilich (2017). The spectral analysis of this long times series revealed statistically significant modes of variability at inter-annual and decadal scales, associated with the strong influence of coupled ocean-atmosphere oscillations in the Pacific Ocean basin on the hydroclimate of western Canada.

Samantha compared reconstructed river flow to hydrological data derived from a run of the CRCM5 driven by reanalysis data (historical data derived from observations and weather models). This comparison enabled an evaluation of the ability of a climate model to replicate the hydroclimate of 1979-2010, the period of overlap between the tree-ring records, weather observations and output from the CRCM. Validation of the capacity of the CRCM to simulate historical climatic variability provides confidence that RCMs are the source of reliable future estimates of the hydroclimate of the Castle River basin.

References

- Barrow, EB and DJ Sauchyn. 2017. An Analysis of the Performance of RCMs in Simulating Current Climate over Western Canada, *The International Journal of Climatology*. DOI: 10.1002/joc.5028
- Barrow, EB and DJ Sauchyn. 2019. Uncertainty in climate projections and time of emergence of climate signals in western Canada. *The International Journal of Climatology*, <https://doi.org/10.1002/joc.6079>
- Kerr, Samantha A. 2018. Tree-ring Reconstructed Streamflow of the Castle River and Comparison to Reanalysis Total Runoff (mrro) for CRCM5, PARC report for the Castle Watershed Project.
- Sauchyn, David and Ilich, Nesa. 2017. Nine hundred years of weekly streamflows: Stochastic downscaling of ensemble tree-ring reconstructions. *Water Resources Research*, DOI: 10.1002/2017WR021585

Appendix 11: ACRU water resources model setup, simulation and soil mapping for the Castle Headwaters

Kienzle. S 2020. ACRU WATER RESOURCES MODEL SETUP AND SIMULATION FOR THE OLDMAN RIVER BASIN HEADWATERS. Final Sub Report to the AIEES WIP. Department of Geography and Environment, University of Lethbridge.

ACRU water resources model setup and simulation for the Castle Headwaters

Kienzle, S.

Table of Contents

| | | |
|---------|---|-----|
| 1 | Analysis of future water yields into the Oldman, St. Mary and Waterton Reservoirs | 225 |
| 2 | Objectives | 226 |
| 3 | Methods..... | 227 |
| 3.1 | Study Area: Three Research Watersheds | 227 |
| 3.1.1 | The Oldman Reservoir Watershed (ORW)..... | 227 |
| 3.1.2 | The St. Mary Reservoir Watershed (SMRW)..... | 228 |
| 3.1.3 | The Waterton Reservoir Watershed (WRW) | 229 |
| 3.2 | Hydrological simulations of impacts of climate change | 230 |
| 3.3 | The ACRU agro-hydrological modelling system..... | 232 |
| 3.4 | Modelling Approach | 233 |
| 3.4.1 | Data sets required..... | 234 |
| 3.4.1.1 | Historical climate data..... | 234 |
| 3.4.1.2 | Future climate data | 234 |
| 3.4.1.3 | Hydrological Response Units (HRUs)..... | 238 |
| 3.5 | Verification Methods | 239 |
| 3.5.1 | Validation Statistics..... | 239 |
| 3.5.1.1 | Difference of Mean Annual Flow | 239 |
| 3.5.1.2 | Difference in Standard Deviation | 240 |
| 3.5.1.3 | Coefficient of Determination | 240 |
| 3.5.1.4 | Slope of the Regression Line | 241 |
| 3.5.1.5 | RMSE-observations standard deviation ratio (RSR)..... | 241 |
| 3.5.1.6 | Nash-Sutcliffe Coefficient of Efficiency (NSE) | 242 |
| 3.5.1.7 | Graphical validation | 242 |
| 3.5.2 | Validation of Simulations for Climate Change Impacts Analysis | 242 |
| 4 | Results and Discussion..... | 243 |
| 4.1 | Verification Analysis..... | 243 |
| 4.1.1 | Difference of Mean Annual Flow | 243 |

| | | |
|-------|---|-----|
| 4.1.2 | Difference in Standard Deviation..... | 244 |
| 4.1.3 | Coefficient of Determination | 244 |
| 4.1.4 | Slope of the Regression Line..... | 244 |
| 4.1.5 | RMSE-observations standard deviation ratio (RSR)..... | 244 |
| 4.1.6 | Nash-Sutcliffe Coefficient of Efficiency (NSE) | 245 |
| 4.2 | Water Yield Analysis | 249 |
| 4.2.1 | Spatial and temporal trends of water yields | 249 |
| 4.3 | Exceedance probability analysis of water yield | 250 |
| 5 | Conclusions | 252 |
| 6 | References | 253 |
| 7 | MAPPING FUNDAMENTAL SOIL PROPERTIES IN THE WAST CASTLE WATERSHED, ALBERTA | 257 |
| 7.1 | Methods..... | 257 |
| 7.2 | Results..... | 258 |
| 7.3 | Conclusions | 262 |
| 7.4 | References | 263 |
| 8 | Other info..... | 264 |
| 8.1 | MSc Theses completed for this project: | 264 |
| 8.2 | Public and conference presentations | 264 |

1 Analysis of future water yields into the Oldman, St. Mary and Waterton Reservoirs

The Oldman River Basin (ORB) is in a semiarid zone located in the southern part of the province of Alberta, Canada (Figure 1). The watershed is 28,270 km² large and has a population of approximately 235,000, of which 100,000 live in the City of Lethbridge (Oldman Watershed Council, 2019). Tributaries to the Oldman River, such as the Waterton, Castle River, Belly or St. Mary River, originate in the Rocky Mountains to the west and south of the ORB. The prairie climate of the region is associated with a very low water yield potential due to the low annual precipitation of less than 450mm per year. Consequently, the potential of agricultural production on good prairie soils and adequate heat units during the growing season are limited by water availability. Consequently, the development of an irrigation distribution system was key to the economic and social development of this arid region. Water from the rivers originating in the Rocky Mountains are

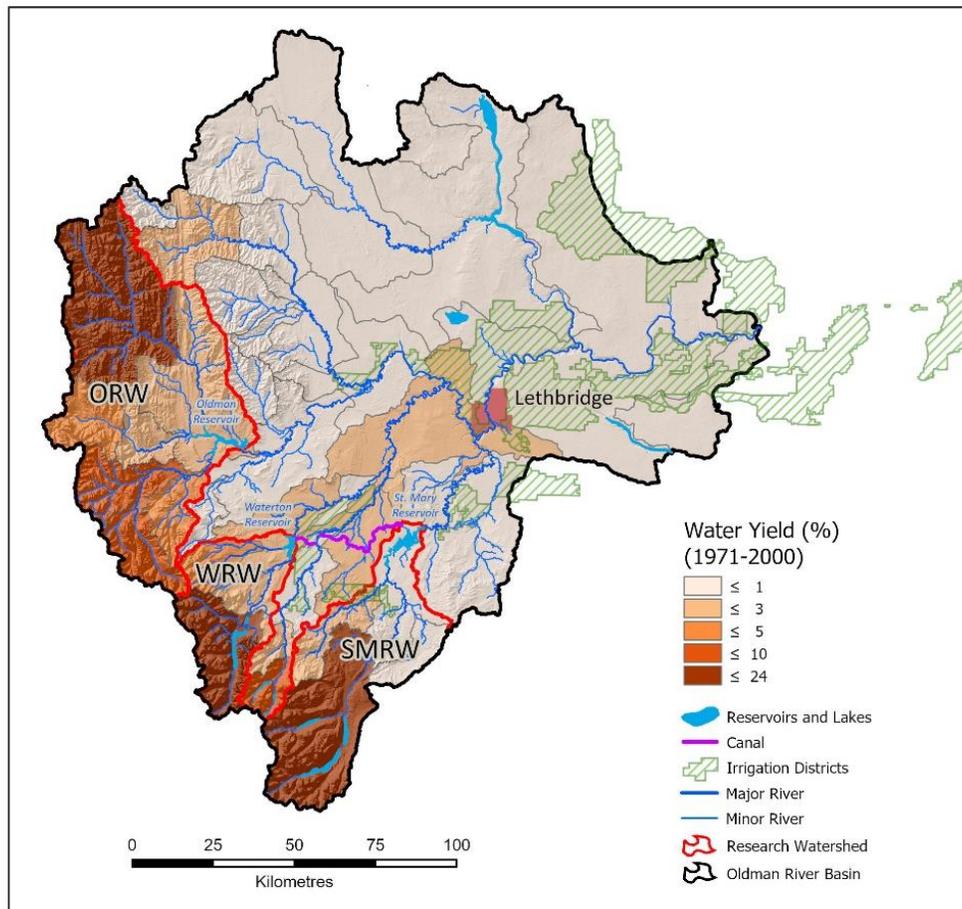


Figure 1: The Oldman River Basin with major reservoirs, three research watersheds, and water yields for the period 1971-2000. ORW = Oldman Reservoir Watershed; WRW = Waterton Reservoir Watershed; St. Mary Reservoir Watershed.

diverted through a series of canals and pipelines to semi-arid prairie areas. Irrigation storage reservoirs capture water mostly during the spring and store it for later use through a large and modern network of irrigation canals. This provides protection from water scarcity during periods of low river flow, and also helps to mitigate extreme floods.

The water resources in the Canadian Prairie region are largely influenced by snow as well as diminishing glaciers and ice fields in the Rocky Mountains. The Rocky Mountains are a considerable source of freshwater (Figure 1), contributing to 50-80% of the spring flows (Barnett et al., 2005; Mueller & Kienzle, 2011). Streamflow has already declined in the northern regions of the Rocky Mountains, with a potential 10% less flow by 2050 (Rood, Samuelson, Weber, & Wywrot, 2005). A diminishing snowpack in the Rocky Mountains could mean altering the magnitude and the duration of peak spring flows for southern Alberta (Rood et al., 2008). In addition to increasing temperatures and decreasing volumes of snowpack, earlier snowmelt and decrease in summer soil moisture can also be expected (Barnett et al., 2005).

The economic vitality of the region is reflected by the amount of water provided for irrigated agriculture. Southern Alberta encompasses 60% of irrigated agriculture in Canada (Russenberger, Bjornlund, & Xu, 2012). More than 70% of the licensed water withdrawals are used to irrigate approximately 4000 km² of land within this region (Russenberger et al., 2012; Schindler & Donahue, 2006). According to AIPA (2019), Alberta's irrigated agriculture contributes approximately \$ 3.6 billion to the provincial domestic product annually. Crops grown are cereals, oilseeds, forages and specialty crops. Forage crops significantly support Alberta's strong beef industry. Specialty crops such as potatoes and sugar beets are processed locally and are shipped globally. As the president of the AIPA stated in a support letter to this project (2016): *"...streamflow into the Oldman and St. Mary reservoirs [...] supplies some 750,000 acres of irrigated land in Alberta. Because climate change will influence the timing and volume of this runoff that supplies water for communities, irrigation, recreation, business and habitat, we will need the tools [...] to help us foresee and be adaptive to new and changing water supply realities."* The objectives of this project meet demands by both the AIPA and *Alberta's Water Strategy* Task to secure *"Future Water Supply and Watershed Management"*.

2 Objectives

The objective of this research project was to simulate the watersheds of the two largest irrigation reservoirs, the Oldman Reservoir (OR) and the St. Mary Reservoir (SMR), for historical (1951-2017) and future (2041-2070) water yields. As the SMR has historically received 59% of the inflow from the St. Mary River and 41% from the Waterton and Belly rivers via the Belly-St. Mary Canal, the Waterton River Watershed (WR) needed to be included in the simulations. These three watersheds represent about 28% of the Oldman River Basin, but, according to Kienzle and Mueller, provide approximately 80% of the annual water yield (1971-2000) of the entire basin (Figure 2). As the soils in the ORB headwaters are not well characterized and poorly understood, a sub-objective is to create a series of soil maps characterizing soil depth and soil hydrological parameters, initially only for the Castle River watershed as a pilot study.

3 Methods

3.1 Study Area: Three Research Watersheds

3.1.1 The Oldman Reservoir Watershed (ORW)

The Oldman River Dam was completed in 1991. The Oldman Reservoir, with a capacity of 490 million m³, is the largest in southern Alberta. The Oldman Reservoir Watershed (ORW) is defined here as the watershed upstream of the Dam Wall (Figure 2). It receives 34% of the total average (1971-2000) water yield of the ORB (Kienzle & Mueller, 2013). It is comprised of alpine landscapes located on the eastern slopes of Alberta's Rocky Mountains and gradually transitions into a foothills region. It is 4,380 km² in size, with elevations ranging from 1037 m to 3151 m. The western part of the watershed consists mainly of coniferous and deciduous forests and barren rock. In contrast, the eastern part of the watershed predominantly consists of grassland and agricultural areas.

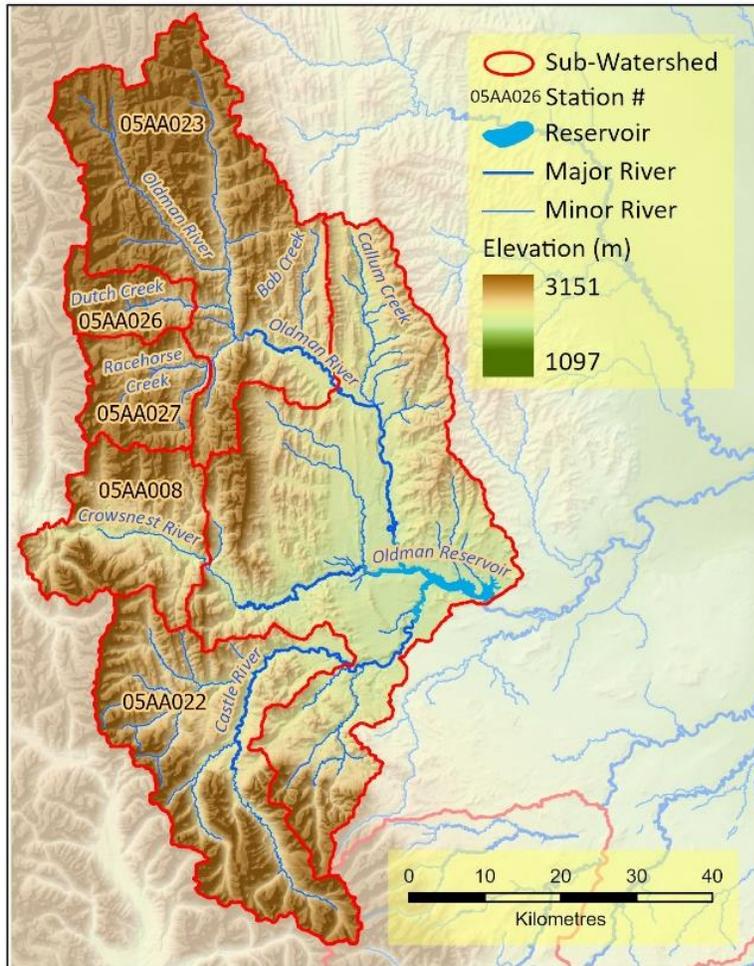


Figure 2: The Oldman Reservoir Watershed and its sub-watersheds with Canadian Water Survey station numbers.

According to the available Parameter-elevation Regressions on Independent Slopes Model (PRISM) surfaces developed by Daly et al. (1994, 2002, 2008), the mean annual precipitation (1971-2000) is 755 mm, ranging from a maximum of over 1600 mm at the Continental Divide to just over 500 mm at the reservoir. Of this precipitation, about 2/3 is falling as snow near the Continental Divide, while only about 1/3 is falling as snow (Kienzle and Clark, 2020).

The tributaries that originate in the Canadian Rocky Mountains are (from south to north): the Castle River, the Crowsnest River, Racehorse Creek, Dutch Creek, and the Oldman River, which all eventually flow into the Oldman Reservoir.

3.1.2 The St. Mary Reservoir Watershed (SMRW)

The St. Mary River Dam was completed in 1951. The St. Mary Reservoir, defined here as the watershed

upstream of the Dam Wall (Figure 3), has a capacity of 396 million m³. It receives 25.4% of the total average (1971-2000) water yield of the ORB (Kienzle & Mueller, 2013). It is comprised in its upper reaches of Glacier National Park in Montana. It is 2,286 km² in size, with elevations ranging from 1,088 m to 3,025 m.

According to the available Parameter-elevation Regressions on Independent Slopes Model (PRISM) surfaces developed by Daly (2006; 2008; 1997), the mean annual precipitation (1971-2000) is 755 mm, ranging from a maximum of over 1600 mm at the Continental Divide to about 450 mm at the reservoir. Of this precipitation, about 2/3 is falling as snow near the Continental Divide, while only about 1/3 is falling as snow (Kienzle and Clark, 2020: albertaclimaterecords.com).

About 55% of the watershed is situated in Montana, where about 90% of the water flowing into the St. Mary Reservoir is produced. International treaties are governing the partition of water that flows into Alberta across the Waterton River, the Belly River, and the St. Mary River. The St. Mary River Irrigation District (SMRID) delivers water through 2060 km of canals and pipelines to approximately 372,000 acres of land between Lethbridge and Medicine Hat.

With a full supply level volume of 396 million m³, the St. Mary Reservoir has a live storage of 350 million m³. From the mean annual total (1980-1986) inflow of 1047 million m³, 650 million m³ (62%) came from the St. Mary River, while 397 million m³ (38%) originated from the Waterton Reservoir.

3.1.3 The Waterton Reservoir Watershed (WRW)

The Waterton Reservoir was established by the Alberta Government in 1964 and has a surface area of 9.6 km² and has a full storage volume of 111.2 million m³. It diverts about 20m³/s of water to the SMR via a canal, which flows into the Belly River, from where it is further diverted about 1 km downstream to a canal connecting to the SMR further east (Figure 4).

The watershed area is 1,273 km², of which 244 km² (19.2%) in the USA, being part of Glacier National Park in Montana. The elevation ranges from 1,175 m at the reservoir to 3,119 m at the peak, part of the continental divide. No major towns exist in the watershed. The upper Waterton River watershed receives the highest amounts of precipitation in Alberta, with values at Cameron Lake recorded to be over 1600 mm per year. This watershed produces 21.3 % of the mean annual streamflow (1971-2000) leaving the ORB. The Waterton River, originating at the headwaters in the Glacier-Waterton Peace Park in Montana, has two main tributaries, the Blakiston Creek and the Drywood Creek.

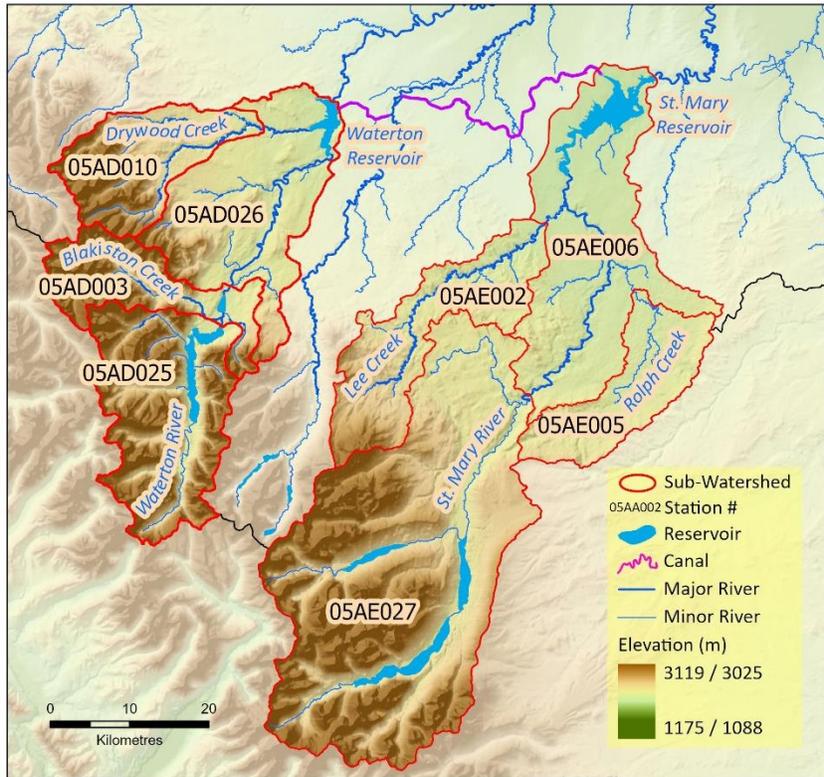


Figure 3: The Waterton Watershed (left) and St. Mary Watershed (right) and their sub-watersheds with Canadian Water Survey station numbers. The purple canal conveys water from the Waterton Reservoir to the St. Mary Reservoir.

Table 3.1: Main characteristics of the three study watersheds

| Watershed | OW | SMW | WR |
|---|------|------|------|
| Full level capacity of Reservoir (in million m ³) | 490 | 357 | 114 |
| Surface area at full supply level (in km ²) | 24.2 | 37.5 | 9.6 |
| Watershed area (in km ²) | 4390 | 2250 | 1273 |
| Watershed area on US territory (%) | 0 | 55.5 | 19.2 |
| Average depth at full supply level (m) | 20.2 | 9.5 | 11.9 |
| Mean Annual Water Yield (1971-2000) | 1542 | 717 | 741 |

3.2 Hydrological simulations of impacts of climate change

The hydrological response to climate change has been studied through the application of watershed-scale hydrological models driven by GCM or RCM-derived scenarios of future climate (e.g. Loukas et al., 2002; Nurmohamed et al., 2007). Physically-based, spatially distributed hydrological models are the only effective means to assess the impacts of climate change on hydrological response, as they are able to

capture the spatial variability of hydrological processes throughout complex watersheds (Bathurst et al., 2004; Warburton, 2010). The ACRU agro-hydrological modelling system (Schulze, 1995, updated; Smithers and Schulze, 1995) is a physical-conceptual, distributed hydrological modelling system designed to be responsive to changes in land cover and climate (Figure 4), and has proven to be providing meaningful and reliable simulations of water resources under conditions of climate change (Warburton, 2010; Kienzle et al., 2012; Nemeth et al., 2012).

Many other hydrological models exist and are used in Canada, including the UBC Watershed Model (Pipes and Quick, 1977), the WATFLOOD model (Kouwen, 2014), HSPF (Hydrological Simulation Program in Fortran, Bicknell et al., 1993), SWAT (Neitsch et al., 2002), the Cold Regions Hydrological Model, CRHM (Pomeroy et al., 2007), or the Distributed Hydrology Soil Vegetation Model, DHSVM (Wigmosta et al., 2002). Each model has certain strengths and weaknesses. The UBC Watershed Model is quite simplistic and does not integrate land cover information. It is designed to work well in alpine and montane regions, which renders it less suitable for prairie regions. WATFLOOD is designed for very large watersheds, such as the Saskatchewan River Basin or the Mackenzie River Basin, and the spatial resolution is typically coarse (i.e. 100 km²). HSPF, on the other hand, is a model that required extensive calibration, which means that it is very data intensive to set up. HSPF lacks evapotranspiration routines, which are critical to simulate climate change impacts. The Distributed Hydrology Soil Vegetation Model (DHSVM) is a research tool that has only a small fraction of input and output variables when compared to ACRU, and very limited support is available, as it is predominantly a research tool. The Cold Regions Hydrological Model (CRHM) is not a classical operational model, but the leading model used in cold regions research environments, where highly detailed input data are available to test hydrological processes.

While the SWAT model and HSPF use a static lapse rate to estimate elevation adjusted air temperatures (Neitsch et al., 2002), others, such as the UBC model or ACRU, adjusts daily minimum and maximum air temperatures using monthly lapse rates. ACRU is the only model where the lapse rate corrected temperatures are further adjusted as a function of daily incoming solar radiation (representing exposition) and land cover. This is considered to be critical for the separation of rain and snow, evapotranspiration and snow melt, and allows for more realistic estimation of local hydrological behaviour.

ACRU, the model of choice, was applied to simulate historical (1951-2017) and future (2041-2070) water yields for each of the three study watersheds. It is emphasized that model validation is the most important part of the hydrological modelling framework. The approach followed here had five steps:

- 1) Set up the ACRU model based on best available data, which includes elevation and terrain derivatives such as the sloped area under-estimation factor and daily solar radiation, climate variables such as minimum and maximum temperatures, precipitation, solar radiation, and estimates of relative humidity and wind speed, a range of land cover variables, a wide range of soil variables, and a spatial delineation of hundreds of hydrological response units for each watershed under consideration.
- 2) A period lasting at least ten years and which includes wet, dry and normal years is chosen for the verification analysis. During the verification process, variables impossible to measure such as the proportion of outflow from an HRU on any given day, or the baseflow recession index, are adjusted

within physically meaningful ranges to match available observations. A wide range of objective functions were used to statistically describe the success of the simulations quantitatively.

- 3) Once verification analyses were completed, ACRU was run for the entire simulation period (here: 1951 to 2017). Validation statistics are again computed for a wide range of objective functions to evaluate the quality of the simulations.
- 4) Finally, ACRU is applied with the same bio-physical parameter sets for each watershed to simulate future streamflows, where the only assumed change is a different set of daily climate time series.
- 5) A wide range of descriptive statistics are used to show changes, including the calculation of five historical climate normal (1951-1980, 1961-1990, 1971-2000, 1981-2010, 1991-2017) and one future climate normal (2041-2070) for water yields and selected hydro-climatological parameters.

3.3 The ACRU agro-hydrological modelling system

ACRU is a spatially distributed physical-based hydrological modelling system designed to simulate total evaporation, soil water, groundwater and reservoir storages, land cover and abstraction impacts, snow water dynamics and streamflow at a daily time step. Having its roots in agro-hydrology, ACRU has been designed to be responsive to changes in land use and climate, with specific variables governing the atmosphere-plant-soil water interfaces. Surface runoff and infiltration are computed using a modified SCS equation (Schmidt and Schulze, 1987), where the daily runoff depth is proportional to the antecedent soil moisture content. ACRU has been applied for water resource assessments (e.g. Everson, 2001; Kienzle et al., 1997; Schulze et al., 2004; Martinez et al., 2008), flood estimation (Smithers et al., 1997, 2001), land use impacts (Kienzle and Schulze, 1991; Tarboton and Schulze, 1993; Kienzle and Schmidt, 2008, Schmidt et al., 2009), and climate change impacts (New and Schulze, 1996; New, 2002; Schulze et al., 2004; Warburton et al., 2010; Forbes et al., 2011; Kienzle et al., 2011), and often requires extensive GIS pre-processing (Kienzle, 1993, 1996; Schulze et al., 1990, Kienzle, 2010). Thus, ACRU was successfully applied in various climate change impacts studies under a wide variety of climates. Comprehensive model manuals are available (Schulze, 1995; Smithers and Schulze, 1995). ACRU is also well described by Kiker et al. (2006). Kienzle (1993) described the link between ACRU, GIS and decision support systems.

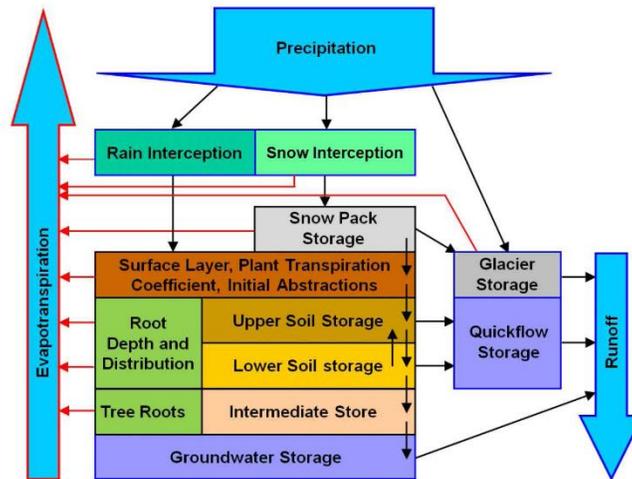


Figure 4: The ACRU agro-hydrological modelling system

Although the emphasis of ACRU is agro-hydrological modelling, the model has been successfully applied in the Canadian Rocky Mountains, because ACRU has a set of particular routines that make it superior over other locally used hydrological models, such as SWAT or CRHM. ACRU is particularly well suited as an operational model for three reasons: its emphasis on agro-hydrological model means that the soil-plant-atmosphere interface is particularly geared to simulating both evaporation and transpiration, both critical components of the daily hydrological budget. The addition of specific mountain hydrology routines optimize the use of sparse hydro-climatological and bio-physical datasets that are typically available from government monitoring networks. The special routines developed over the recent years include the realistic spatial distribution of hydro-climatological time series across entire watersheds, the downscaling of regional climate models, or the area-adjustment for sloped areas.

3.4 Modelling Approach

As is the case with every integrated/multipurpose hydrological modelling system applied to simulate hydrological responses in large and heterogeneous watersheds, ACRU requires considerable spatial information, *inter alia*, on topography, a wide range of bio-physical and hydroclimatological parameters. The spatial organization of ACRU is usually based on hydrological response units (HRUs), each having a unique combination of elevation, land cover, and climate.

In Alberta, not all required variables are available, in which case they are estimated within physically meaningful ranges, based either on available literature, complex GIS analysis, or local expert knowledge. In a typical application, however, variables that remain unknown or are uncertain, such as soil depth, flow routing variables or baseflow parameters, require calibration against streamflow hydrographs. The output of the ACRU model consists of daily time series of 52 variables for each HRU, including air temperature, snow water equivalent, soil water deficit and surplus, groundwater recharge, groundwater flow, and streamflow, as well as water use by vegetation, and evaporation from wet surfaces. From the time series, frequency analyses of any variable can be carried out using exceedance probability plots, which provide information on the percentage of time a certain value, e.g. flood, soil moisture or low

flow, is exceeded.

3.4.1 Data sets required

Motivated by the lack of readily available hydro-climatological data for the calculation of potential and actual evapotranspiration, soil moisture deficits, groundwater recharge, and crop modelling, massive processing of hydro-climatological data, including solar radiation, sunshine hours, wind, and relative humidity, were created for the entire province of Alberta. Recently available climate data sets of daily T_{min}, T_{max} and precipitation for the period 1950 to 2017, at a spatial resolution of 10 by 10 km, were processed to be directly available for input into ACRU. These important databases enable the calculation of daily and spatially explicit estimates of potential evapotranspiration (using Penman-Monteith).

3.4.1.1 Historical climate data

In order to provide a daily climate time series required for spatially consistent distributed hydrological modelling, daily climate surfaces are required to be interpolated from daily point-source data (Jolly et al. 2005). The National Land and Water Information Service (NLWIS) released the daily 10-km gridded climate dataset to cover the period 1950 to 2017. The analysis was achieved by collaboration between Environment Canada, Custom Climate Services, Natural Resources Canada, and the Australian National University. The dataset consists of 10 km by 10 km grids cells for daily minimum and maximum temperatures and precipitation for the Canadian landmass south of 60° N and was therefore suitable for this study. Grid values were interpolated using records from daily climate stations from the Canadian Climate Data Archives (Environment Canada, 2012) using ANUSPLIN, a tri-variate thin plate smoothing spline surface fitting method (Hutchinson 2004; Hutchinson et al. 2009; Hopkinson et al. 2011; Newlands et al. 2011), with latitude, longitude, and elevation as co-variables. In comparison with other interpolation techniques, this method performs well when interpolating noisy climate data across complex terrain (Hutchinson and Gessler 1994; McKenney et al. 2006). Each 68-year time series had a unique location identifier, which was associated with a specific latitude and longitude. As it impossible to create square grids on the curved Earth surface, the point locations were spatially gridded by calculating Thiessen polygons. There are not true 10km by 10km squares, but are rather polygons where any location inside the polygon is closest to associated point holding the climate time series. The Thiessen polygons are referred here as 10-km climate grids to illustrate their origin. Anderson (2016) has shown that the 10-km climate grids are superior for hydrological simulations over the traditional methods of using a few single climate stations to drive the hydrological model, because the tradition method of using “driver stations” has a major disadvantage. This shortcoming is that often entire sub-watersheds have one climate input, resulting in either an entire sub-watershed receiving precipitation or not, consequently over-estimating precipitation events.

3.4.1.2 Future climate data

A climate scenario can be defined as “*a coherent, internally consistent and plausible description of a possible future state of the world...*” (Parry, 2002). They are “*plausible representations of the future that are consistent with assumptions about future emissions of greenhouse gases and other pollutants*”

and with our understanding of the effect of increased atmospheric concentrations of these gases on global climate” (IPCC-TGICA, 2007). Climate scenarios are not predictions of future climate. They are typically used to assess impacts of climate change, address vulnerability to change and to develop adaptation strategies or actions.

Climate scenarios are dependent on emission assumptions. The Intergovernmental Panel on Climate Change (IPCC) published a set of emissions scenarios for use in climate change studies (SRES = Special Report on Emissions Scenarios (Nakicenovic, et al., 2000)). The SRES team has developed conditions to represent future evolution of demographic, social, economic, technological, and environmental developments. Out of six marker scenarios, climate modellers have focussed on three of these scenarios, referred to as the low (B1 = stabilization of CO₂ at 550 ppm), intermediate (A1B = stabilization at 750 ppm) and high (A2 = CO₂ does not stabilize) scenarios.

The North American Regional Climate Change Assessment Program (NARCCAP) is an international program to provide regionally meaningful simulations of climate change (Mearns et al., 2009). NARCCAP researchers are running a set of regional climate models (RCMs) over a large area covering the conterminous United States and most of Canada. As RCM have significantly higher spatial resolution than GCMs, they offer a higher topographic complexity and allow the simulation of finer scale atmospheric dynamics. Consequently, RCMs provide a more adequate method of producing the climatic information required for regional impact studies (Poitras et al. 2011; Barrow and Sauchyn 2017). The RCMs receive inputs from a set of atmosphere-ocean general circulation models and are driven by the A2 emissions scenario. The A2 emissions scenario was one of the 'marker' scenarios developed through the IPCC. It also happens to be the most realistic scenario based on the global greenhouse gas reduction efforts (or the lack thereof) in the past two decades. RCMs, with a spatial resolution ranging from about 25 to 50 km, provide much more spatial detail than a GCM, which have spatial resolution between about 125 to 600 km, by one to two orders of magnitude, and therefore represent the topography and associated atmospheric dynamics much better. Consequently, they provide much more adequate climate information for regional impact studies (Poitras et al. 2011; Barrow and Sauchyn 2017). The RCMs were simulated for the historical base period 1971-2000 and the future climate normal period 2041-2070.

Three Regional Climate Models (RCMs) representing future climate projections for the period 2041 to 2070 for relatively wetter/cooler to relatively warmer/drier conditions have already been bias-corrected and downscaled to match the existing 1951 to 2017 climate grids, so that future simulations can be based on a consistent database. As these RCMs have only been processed for Alberta, they needed to be bias corrected for the US portion of the ORB to enable future simulations on the St. Mary and Waterton Rivers.

Three RCMs were chosen to represent the projected envelope of possible future climates. From a large cloud of possible climate projections, those were chosen that represent a median scenario (half of the other scenarios are relatively wetter or drier, or relatively cooler or warmer), and two scenarios representing the opposite ends of future hydrological behaviours: a relatively cooler and relatively wetter scenario, resulting in relatively high water yield, and a relatively warmer and relatively drier scenario, resulting in relatively low water yields.

- The **Cooler/Wetter Scenario** is based on the *Regional Climate Model Version 3* (RCM3), based on the *Third-Generation Coupled GCM* (cgcm3). The RCM3 was developed at the University of Carolina-Santa Cruz, USA.
- The **Median Scenario** is based on the *Canadian RCM* (CRCM), also driven by the *Third-Generation Coupled GCM* (cgcm3).
- The **Warmer/Drier Scenario** is based on the *Hadley Regional Model 3* (HRM3), which is driven by the *Geophysical Fluid Dynamics Laboratory GCM* (gfdl). The HRM3 was developed at the Hadley Centre, UK, while the gfdl was developed at Princeton University, USA.

As an RCM is still much larger than the 10 km climate grid, the RCMs were further downscaled to match the spatial extent of each 10 km climate grid. The downscaling of GCM output from global to watershed scale is typically carried out using either statistical or dynamical downscaling techniques. Previous studies have compared the two downscaling techniques (Lapp et al., 2009; Murphy, 1999, 2000; Teutschbein et al., 2011; Wilby et al., 1997). The statistical downscaling method utilizes statistics in order to build relationships between the large-scale data set against regionally available variables (Beckers et al., 2009; Fowler et al., 2007; Wilby et al., 2004). In contrast, the dynamical downscaling method focused on the development of RCMs (Teutschbein & Seibert, 2010). The extraction of local climate information using GCM output data as a boundary condition allows the local scale climate processes to be captured (Teutschbein & Seibert, 2010; Wilby et al., 2002). Despite the recent progress in the development of RCMs (Teutschbein & Seibert, 2010, 2012), RCM data sets often have biases due to the definition and spatial averaging within GCM grid cells. Therefore, Teutschbein and Seibert (2010, 2012) recommended to apply a bias correction technique to an ensemble of RCM simulations, because previous studies have shown that uncorrected RCM simulations are a large source of uncertainty in modelling hydrological response to climate change. Consequently, this study used bias-corrected RCM data.

As the RCMs have one data period covering the historical period ranging from 1971-2000, the RCMs climate time series can be downscaled using the Delta Method while being bias corrected, thus spatially matching the observed data's spatial resolution. The primary reasons for matching the spatial resolution of the observed 10 km climate grid time series are

- (a) to replicate the seasonality and magnitude of the 1971-2000 climate normal characteristics,
- (b) to apply the established parameter input files based on the historical period to the future RCM data sets, thus preserving the localized downscaling bias corrections, and
- (c) to enable the comparison of the three future RCM projections used in this study.

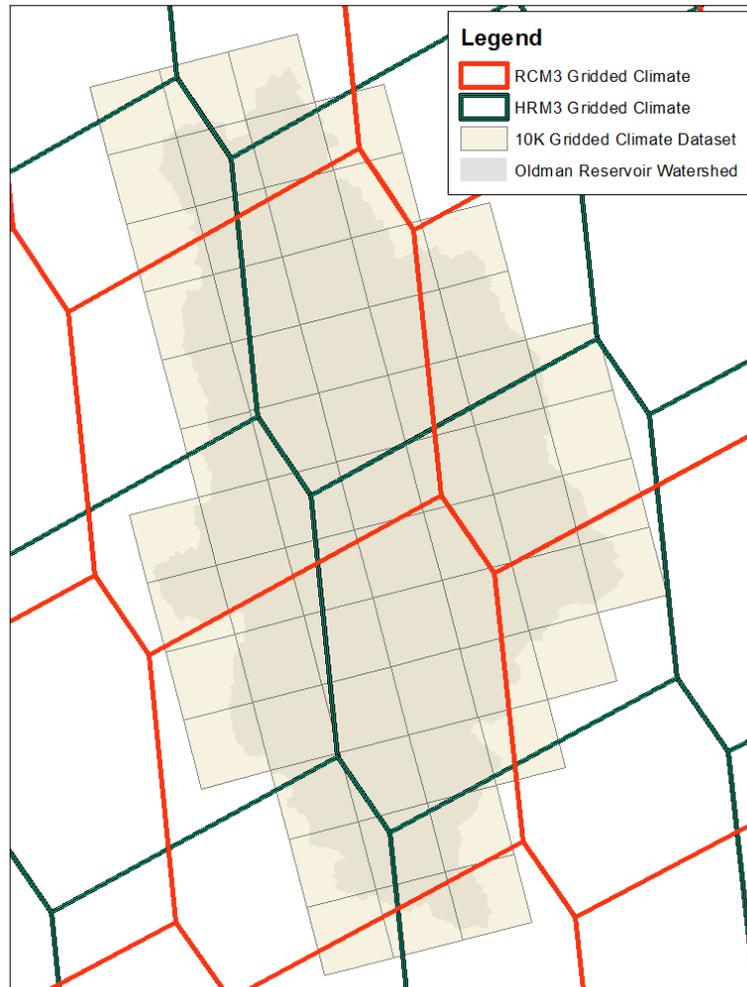


Figure 5: Comparison of the 10 km climate grid with two RCMs for the area of the Oldman Reservoir Watershed. The polygons are Thiessen polygons constructed from gridded RCM points. Two regional climate data: 1) HRM3 (warmer/drier) climate projection driven by GFDL and 2) RCM3 (cooler/wetter) climate projection driven by CGCM3 (from Bonifacio, 2016).

Since the regional climate model data and hydro-climatological data are available in spatially separated time series, it was important to spatially downscale the regional climate data using an area-weighting ratio based on the spatial overlay of the 10K climate grids and the RCM climate grids. The downscaling and bias correction procedures were as follows:

- 1) Create Thiessen Polygons for each RCM grid point (Figure 5). This was necessary, as RCMs are not available as true square grids, but rather as point locations. A square grid cannot be constructed to fit the Earth's sphere.
- 2) Overlay the 10 km climate gridded dataset and the RCM Thiessen Polygons in a GIS (Figure 5)
- 3) Calculate the intersecting areas ("areaUnion") created during the spatial overlay process.

- 4) Calculate the monthly average for minimum and maximum temperature and precipitation for each 10Km climate grid and each RCM polygon, resulting in spatially averaged areas for determination of correction factors (variable P_{10K-i} and T_{10K-i} , and P_{RCM-i} and T_{RCM-i} , with $i=1-12$).
- 5) Calculate the temperature differences and precipitation ratios for each month between the RCM value and the 10 km Climate Grid as follows:

$$P_{ratio-i} = P_{10K-i} / P_{RCM-i}$$

$$T_{ratio-i} = T_{10K-i} - T_{RCM-i}$$
- 6) Calculate the percent area (perArea) of each 10K Climate Grid, complete or partial as follows:

$$P_{perArea} = P_{areaUNION} / P_{areaRCM-i}$$

$$T_{perArea} = T_{areaUNION} / T_{areaRCM-i}$$
- 7) The correction factors are calculated as follows:

$$P_{perRatio} = P_{perArea} * P_{ratio-i}$$

$$T_{perRatio} = T_{perArea} + T_{ratio-i}$$
- 8) The RCM-based daily precipitation time series (P_{RCM}) and daily temperature time series (T_{RCM}) for each 10 km climate grid are calculated using the equation:

$$P_{RCMcorr} = P_{RCM} * P_{ratio-i}$$

$$T_{RCMcorr} = T_{RCM} + T_{ratio-i}$$

3.4.1.3 Hydrological Response Units (HRUs)

Hydrological Response Units (HRUs) are spatial units with relatively homogeneous hydrological behaviour that can be delineated based on a selection of similar physical characteristics (Flügel, 1995). Watershed boundaries were based on PFRA watershed boundaries (Figure 1). A digital elevation model (DEM) was obtained from AltaLIS Ltd. (www.altalis.com), a major source of spatial data and imagery in Alberta, with a spatial resolution of 10m grid cells (equivalent to 1:20,000 scale maps in accuracy). For the creation of HRUs, terrain derivatives such as sloped area under-estimation factor and solar radiation were calculated and then resampled to a 100 m grid cell size. The 100 m DEM was classified into 100 m interval elevation bands, and for elevation above 2000 m into 200 m interval elevation bands. The 10 km climate grids were included for the HRU delineation, so that the HRUs included in each 10 km grid cell could be fed by an individual climate time series. This allows for more realistic hydrological simulations, because the response for an e.g. 2000 km size watershed is driven by 20 individual climate time series, with different daily precipitation inputs, ranging possibly from zero to extremely high. The landcover shapefile was reclassified into eight categories of land cover classes. Lastly, the mean annual solar radiation was calculated using the Solar Radiation tool in ArcGIS 10.5, aggregated from quarter-hour intervals for the entire year for each 100 m grid cell. The annual output was then reclassified into four quartiles.

For each study watershed, spatial overlay analysis of the sub-watershed boundaries, the 10 km climate grids, the elevation bands, the eight land cover layers, and the four annual solar radiation classes resulted in many HRUs. The HRUs were then analysed for their area, and HRUs with an area smaller than 1 km² were iteratively aggregated into neighbouring HRUs until the smallest HRU was at least 1 km² in size.

The final number of HRUs were in 1706 HRUs for the ORB, 608 HRUs for the WRB, and 545 HRUs for the

SMRB. Each HRU was then parameterized to have a unique combination of hydrological variables, most of which were derived by GIS overlay analysis. The area of each HRU was calculated based on its true, sloped area, as the planimetric area derived from ArcGIS is under-estimated in steeply sloped terrain, which otherwise would have affected interception volumes, soil moisture storages, groundwater recharge rates, actual evapotranspiration volumes, and runoff coefficients (Kienzle 2010).

3.5 Verification Methods

There are currently no universally accepted guidelines for evaluating hydrological models in climate change studies (Moriassi et al., 2012). Here, a combination of evaluation approaches were used to rigorously quantify the ACRU model performance. The key objective of hydrological simulations is for simulated values to "mimic" as closely as possible corresponding observed values, where the simulated values should correspond to observed values as closely as possible on a 1:1 basis, where means and variances of simulated values are conserved when compared with means variances of observed values, where simulated and observed values exhibit a close association with one another, where there is no systematic under- or over-simulation error, i.e. no bias, between simulated and observed trends, and that the streamflow probabilities are similar. The degree of closeness of simulated and observed streamflow values can be measured by a number of specific goodness-of-fit criteria, known as objective functions.

The resulting validation statistics are classed into four categories (excellent, good, satisfactory, unsatisfactory) so as to allow for easier evaluation of the overall simulation success. The classifications are based on Schulze and Smithers (1995) for the percent difference of simulated and observed water yields, and on recommendations by Moriassi et al. (2007).

The objective functions tested included (Tables 4.1 to 4.4):

- The percentage difference between the sum of simulated daily flows and observed daily flows, which is equivalent to the Percent Bias (PBIAS).
- The percentage difference between standard deviations of simulated daily flows and observed daily flows.
- The coefficient of determination (r^2) for both daily and monthly flows.
- The regression coefficient (slope as a ratio),
- The ratio of the root mean square error to the standard deviation of measured data (RSR), and
- the Nash-Sutcliffe efficiency coefficient (NSE), where values for daily (monthly) simulations range from over 0.75 for excellent to less than 0.5 for unsatisfactory (Moriassi et al., 2007).

3.5.1 Validation Statistics

3.5.1.1 Difference of Mean Annual Flow

The key variable is simulation of Mean Annual Flow (MAF), shown in **Equation 1**, using the Schulze and Smithers (1995) classification. According to Schulze and Smithers (1995), a difference of less than 5% is characterized as excellent, and one over 15% is seen as unsatisfactory. The equivalent PBIAS value (**Equation 2**) is classified using Moriassi et al. (2007), where values under 10% are classified as excellent,

and values over 25% are classed as unsatisfactory.

$$MAF_{Diff} = \left(\left(\sum_{i=1}^n P_i - \sum_{i=1}^n O_i * 100 \right) - 100 \right)$$

where MAF_{Diff} is the difference summed streamflow during the observation period, expressed as a percent, and P_i is the simulated, or predicted, daily streamflow, and O_i is observed daily streamflow.

The Percent Bias (PBIAS) is calculated using the same input variables as the equation for Mean Annual Flow Differences:

$$PBIAS = \left[\frac{\sum_{i=1}^n (O_i - P_i) * 100}{\sum_{i=1}^n (O_i)} \right]$$

3.5.1.2 Difference in Standard Deviation

The difference in standard deviation is important for hydrological simulations as this objective function provides a measure to evaluate how well the model mimics the range and frequencies of often highly variable streamflow volumes (**Equation 3**). Following recommendations by Schulze and Smithers (1995), values less than 5% are characterized as excellent, and the aim is keep the difference below 15%. In order to categorize four evaluation classes, differences below 15% are considered to be good, values under 25 are considered satisfactory, and differences over 25% are considered unsatisfactory.

$$STDEV_{Diff} = \left((STDEV_{Sim} - STDEV_{Obs}) * 100 \right) - 100$$

where $STDEV_{Diff}$ is the difference between simulated ($STDEV_{Sim}$) and observed ($STDEV_{Obs}$) standard deviations, expressed as a percentage.

3.5.1.3 Coefficient of Determination

The coefficient of determination (r^2) describes the degree of collinearity between simulated and measured data, thus describing the proportion of the variance in measured data explained by the model. The coefficient of determination (**Equation 4**) ranges from 0 to 1, with higher values indicating less error variance. According to Santhi et al. (2001) and Van Liew et al. (2003), r^2 values greater than 0.5 are considered acceptable. However, here the evaluation criteria by Schulze and Smither (1995) are applied, which are more stringent. Although the r^2 has been commonly used for model validation, the coefficient of determination is considered to be oversensitive to high extreme values (floods) and insensitive to accurate simulation of low flows (Legates and McCabe, 1999).

$$r^2 = \left[\frac{\sum_{i=1}^n (O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^n (O_i - \bar{O})^2} \sqrt{\sum_{i=1}^n (P_i - \bar{P})^2}} \right]$$

where O_i is the i^{th} observation of daily streamflow, P_i is the i^{th} simulated value of daily streamflow, \bar{O} is the mean of observed data, \bar{P} is the mean of simulated data, and n is the total number of observations.

3.5.1.4 Slope of the Regression Line

The slope of the regression line (b) indicates the relative relationship between simulated and measured values (**Equation 5**). A slope of 1 for simulated against observed streamflows indicates that the model perfectly reproduces the magnitudes of measured data. A slope greater than 1 indicates over-simulation of daily streamflows, and a slope less than 1 reveals under-simulation of daily streamflows. The slope is highly sensitive to very high values (flood events). The slope is commonly examined under the assumption that measured and simulated values are linearly related.

$$b = \frac{\sum_{i=1}^n (O_i - \bar{O})(P_i - \bar{P})}{\sum_{i=1}^n (O_i - \bar{O})^2}$$

where b is the slope of the regression line, expressed as a ratio, where O_i is the i^{th} observation of daily streamflow, P_i is the i^{th} simulated value of daily streamflow, \bar{O} is the mean of observed data, \bar{P} is the mean of simulated data, and n is the total number of observations. Slope values are rated according to Schulze and Smithers (1995), and classed as “Excellent” when the slope is greater than or equal to 0.9, and unsatisfactory when it is less than 0.6.

3.5.1.5 RMSE-observations standard deviation ratio (RSR)

The RMSE-observations standard deviation ratio (RSR) is recommended as a model evaluation statistic (e.g. Singh et al., 2004) and normalizes the RMSE by using the observed standard (**Equation 6**). A value of zero indicates perfect model simulation. The smaller RSR is, the better is the simulation. RSR is rated here following Moriasi et al. (2007) to be excellent when values are less than or equal 0.5, and unsatisfactory when values are greater than 0.7.

$$RSR = \frac{RMSE}{STDEV_{obs}} = \left[\frac{\sum_{i=1}^n (O_i - P_i)^2}{\sqrt{\sum_{i=1}^n (O_i - \bar{O})^2}} \right]$$

where $RMSE$ is the root-mean-square-error, $STDEV_{obs}$ is the standard deviation of observed streamflows, O_i is the i^{th} observation of daily streamflow, P_i is the i^{th} simulated value of daily streamflow, \bar{O} is the mean of observed data, \bar{P} is the mean of simulated data, and n is the total number

of observations.

3.5.1.6 Nash-Sutcliffe Coefficient of Efficiency (NSE)

Nash and Sutcliffe (1970) characterise the Nash-Sutcliffe efficiency coefficient (NSE) as a normalized statistic that describes the relative magnitude of the residual variance, referred to as “noise”, and compares it to the measured data variance, referred to as “information”. The NSE indicates how well the plot of observed versus simulated data fits the 1:1 line. NSE is computed as shown in **Equation 7**:

$$NSE = 1 - \left[\frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \right]$$

where O_i is the i^{th} observation of daily streamflow, P_i is the i^{th} simulated value of daily streamflow, \bar{O} Ymean is the mean of observed data, and n is the total number of observations. NSE values range from $-\infty$ and 1.0, with $NSE = 1$ being the optimal value. According to Moriasi et al. (2007), NSE values greater than or equal to 0.0 are generally viewed as acceptable levels of performance, whereas values below 0.0 indicate that the mean observed value is a better predictor than the simulated value, which indicates unacceptable performance. However, following the guidelines suggested by Moriasi et al. (2007), NSE values must be greater than or equal to 0.5 to be considered satisfactory (Tables 4.1 to 4.4). Sevatt and Dezetter (1991) proposed that the NSE is the best objective function for reflecting the overall fit of a hydrograph.

3.5.1.7 Graphical validation

These statistical methods are supplemented by a wide range of graphical comparisons of simulated and observed values, including: (a) scatterplots of daily and monthly streamflows, (b) exceedance probability curves of daily and monthly streamflows, (c) seasonal simulated and observed streamflows, and (d) daily simulated and observed hydrographs for the entire simulation period.

3.5.2 Validation of Simulations for Climate Change Impacts Analysis

The validation statistics are reported twice for all simulated sub-watersheds: once for the calibration period, which varies for the three sub-watersheds due to available observed streamflow records. For the OR, the calibration period is 1971-1980, for the SMR, it is 1956-1970, and for the WR, it is 1976-1990. The validation statistics are also reported for the period 1971-2000, as this the period for which the RCMs were bias-corrected to statistically match this climate normal period.

Table 3.4: Summary of Watershed stations used for verification and validation analyses

| Watershed ID | OR 1 | OR 2 | OR 3 | OR 4 | OR 5 | SMR 1 | SMR 2 | SMR 3 | WR 1 | WR 2 | WR 3 |
|-----------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| HYDAT Station # | 05AA022 | 05AA008 | 05AA027 | 05AA026 | 05AA023 | 05AE027 | 05AE005 | 05AE002 | 05AD003 | 05AD010 | 05AD008 |

| HYDAT Station Name | Castle River | Crowsnest River | Racehorse Creek | Dutch Creek | Oldman River | St. Mary River | Rolph Creek | Lee Creek | Waterton River | Drywood Creek | Waterton River |
|---|--------------|-----------------|-----------------|-------------|--------------|----------------|-------------|-----------|----------------|---------------|----------------|
| Calibration Period | 1971 - 1980 | | | | | 1956-1970 | | | 1976 - 1990 | | |
| Area (km ²) | 820.3 | 401.5 | 215.9 | 142.4 | 1091.0 | 1206.4 | 222.4 | 312.3 | 611.6 | 238.3 | 882.5 |
| Areal proportion relative to all 3 watersheds (%) | 13.3 | 6.5 | 3.5 | 2.3 | 17.8 | 19.6 | 3.6 | 5.1 | 10.0 | 3.9 | 14.4 |
| Areal proportion to watershed (%) | 30.7 | 15.0 | 8.1 | 5.3 | 40.8 | 69.3 | 12.8 | 17.9 | 35.3 | 13.8 | 50.9 |
| Average simulated Water Yield for 1971-2000 (million m ³) | 463 | 141 | 76.6 | 52.0 | 336 | 571.2 | 7.8 | 52.6 | 512 | 90.9 | 741.2 |

A total of 11 HYDAT gauging stations were available for the verification of simulated streamflow (Table 3.4): five for the OR, and three each for the SMR and WR watersheds. The verification statistics can be classified into groups of acceptance, based on Schulze and Smithers (1995) and Moriasi et al. (2007).

4 Results AND DISCUSSION

4.1 Verification Analysis

A summary of the validation statistics for all 11 watersheds are provided for both the respective calibration periods and the common baseline period 1971-2000 (Tables 4.1 and 4.2). In Tables 4.3 and 4.4, the validation statistics for all 11 simulated sub-watersheds were averaged by weighting the individual proportions by the average water yield proportion for the period 1971 to 2000, resulting in a “water yield proportion-weighted average” as a measure for overall model performance for all 11 sub-watersheds. Table 4.4 shows that daily simulated streamflows compared very well with observed streamflow records. With an overall under-simulation of streamflow by 2.1%, an overall average Nash-Sutcliffe coefficient of efficiency of 0.66 for daily flows, and a coefficient of determination for monthly flows of 0.79, simulations were of a quality to allow the simulation of future climate projections with confidence.

4.1.1 Difference of Mean Annual Flow

Table 4.1 provides the verification statistics for the calibration period, which varies for each of the three research watersheds. The following conclusions can be derived from the verification results. Nine out of 11 comparisons were classified “Excellent”, being within 5%. One sub-watershed was over-simulated by 9% , while one was over-simulated by 11%. and only two

stations had errors in water yield simulation of about 9% and 11%. Tables 4.5 and 4.6 reveal that the two watersheds not rated “Excellent” contributed only about 3% and 10% of the runoff of the respective watershed.

4.1.2 Difference in Standard Deviation

The comparison of simulated against observed standard deviation in daily streamflows resulted in “Unsatisfactory” to “Excellent” simulations (Tables 4.1 and 4.2). The general trend is that, especially in watersheds with otherwise high rated simulations, the simulated streamflow is under-simulated. That is to be expected, as several high flood events, such as the 1975 and the 1995 floods, are not represented in the climate records and are, consequently, not simulated to be floods. As the standard deviation uses the square of the simulation errors, large differences skew the validation results. For the underlying research question, an under-simulation of standard deviations of daily flows is presumed to be not a major problem.

4.1.3 Coefficient of Determination

The coefficients of determination between daily simulated and observed flow for the 11 sub-watersheds ranges from “Excellent” to “Unsatisfactory”. While the results of two watersheds are deemed unsatisfactory, they represent less than 10% of the total study area of 7950 km² (Table 4.3). The coefficients of determination between monthly simulated and observed flow are higher (Table 4.3), where the area-weighted coefficients of determination for 11 gauged sub-watersheds reveal an increase from 0.68 for daily flows to 0.80 for monthly flows. As the objective of this research was to determine climate change impacts on annual water yield rather than daily streamflow, the modelling results can be interpreted as being acceptable.

4.1.4 Slope of the Regression Line

Table 4.1 reveals that the slopes range from 0.66 to 0.88, indicating a systematic under-simulation of daily streamflows. As the regression line is highly influenced by a few large outliers, such as during large floods, this objective function is sensitive to those outliers. Similarly to the difference in standard deviation, an under-simulation of rare flood events is to be expected.

4.1.5 RMSE-observations standard deviation ratio (RSR)

The water yield proportion-weighted average for all three research watersheds is rated, with a value of 0.57, as good (Table 4.4). Only Lee Creek (RSR = 0.8) and Rolph Creek (RSR = 0.95) have RSR values

significantly over 0.7. It is noted that Rolph Creek is a small tributary contributing 1.1% of the watershed's water yield (Table 4.6).

4.1.6 Nash-Sutcliffe Coefficient of Efficiency (NSE)

For the 11 simulated sub-watersheds, the NSE ranges from "Unsatisfactory" to "Excellent", with a water yield proportion-weighted average of 0.73 for the calibration periods, and 0.66 for simulation of the baseline period 1971-2000 (Tables 4.3 and 4.4). Therefore, the streamflow simulations into the three respective reservoirs are categorized as "Good". Again, both Lee Creek and Rolph Creek simulations are unsatisfactory, to the most part because of the low and variable streamflows.

Table 4.1: Validation statistics for three research watersheds for the respective calibration periods.

| Validation Statistic | Excellent | Good | Satisfactory | Unsatisfactory | OR 1 | OR 2 | OR 3 | OR 4 | OR 5 | SMR 1 | SMR 2 | SMR 3 | WR 1 | WR 2 | WR 3 |
|----------------------------------|--|---------|--------------|----------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| HYDAT Station # | | | | | 05AA022 | 05AA008 | 05AA027 | 05AA026 | 05AA023 | 05AE027 | 05AE005 | 05AE002 | 05AD003 | 05AD010 | 05AD008 |
| HRU # | | | | | 314 | 492 | 613 | 716 | 1111 | 341 | 391 | 463 | 256 | 376 | 608 |
| n (daily) | OM:1971-1980 SMR: 1956-1970 WR = 1976-1990 | | | | 3653 | 3653 | 2126 | 2126 | 3288 | 3644 | 3644 | 5479 | 5113 | 5113 | 5752 |
| % Difference of Mean Annual Flow | <= 5 | <= 10 | <= 15 | > 15 | 0.29 | -2.39 | 0.79 | 9.16 | -1.53 | -2.1 | -11.1 | -2.8 | -2.13 | 4.09 | -1.65 |
| Difference in standard deviation | <= 5 | <= 15 | <= 25 | > 25 | -14.94 | -4.31 | -12.57 | -4.32 | -12.74 | -30.8 | 10.1 | -31.3 | -2.8 | 12.36 | -17.14 |
| r ² daily flows | >= 0.80 | >= 0.70 | >= 0.60 | < 0.60 | 0.77 | 0.77 | 0.69 | 0.62 | 0.68 | 0.75 | 0.36 | 0.37 | 0.82 | 0.55 | 0.78 |
| r ² monthly flows | >= 0.85 | >= 0.75 | >= 0.65 | < 0.65 | 0.89 | 0.84 | 0.84 | 0.78 | 0.83 | 0.85 | 0.54 | 0.65 | 0.87 | 0.75 | 0.85 |
| Slope of the regression line | >= 0.9 | >= 0.8 | >= 0.6 | < 0.6 | 0.75 | 0.84 | 0.74 | 0.75 | 0.72 | 0.72 | 0.66 | 0.72 | 0.88 | 0.66 | 0.74 |
| Percent Bias | <= 10% | <= 15 | <= 25 | > 25 | -0.29 | 2.39 | -0.79 | -9.16 | 1.53 | 2.1 | 11.2 | 2.9 | 2.13 | -4.17 | 1.65 |
| RSR | <= 0.50 | <= 0.60 | <= 0.70 | > 0.70 | 0.48 | 0.49 | 0.54 | 0.64 | 0.56 | 0.5 | 0.95 | 0.8 | 0.43 | 0.68 | 0.48 |
| NSE | >= 0.75 | >= 0.65 | >= 0.50 | < 0.50 | 0.77 | 0.76 | 0.71 | 0.58 | 0.68 | 0.75 | 0.1 | 0.36 | 0.82 | 0.53 | 0.77 |

Table 4.2: Validation statistics for three research watersheds for the period 1971-2000.

| Validation Statistic | Excellent | Good | Satisfactory | Unsatisfactory | OR 1 | OR 2 | OR 3 | OR 4 | OR 5 | SMR 1 | SMR 2 | SMR 3 | WR 1 | WR 2 |
|----------------------------------|-----------|---------|--------------|----------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| HYDAT Station # | | | | | 05AA022 | 05AA008 | 05AA027 | 05AA026 | 05AA023 | 05AE027 | 05AE005 | 05AE002 | 05AD003 | 05AD010 |
| HRU # | | | | | 314 | 492 | 613 | 716 | 1111 | 341 | 391 | 463 | 256 | 376 |
| n | 1971-2000 | | | | | | | | | 10957 | 7141 | 10957 | 10958 | 10593 |
| % Difference of Mean Annual Flow | <= 5 | <= 10 | <= 15 | > 15 | 0.51 | -3.14 | -5.53 | 4.24 | -5.74 | 0.07 | 5.5 | 4 | -6.15 | 4.46 |
| Difference in standard deviation | <= 5 | <= 15 | <= 25 | > 25 | -15.2 | -9.29 | -20.1 | -9.9 | -23.54 | -12.3 | -25.5 | -32.4 | -13.32 | -50.0 |
| r ² daily flows | >= 0.80 | >= 0.70 | >= 0.60 | < 0.60 | 0.73 | 0.72 | 0.68 | 0.64 | 0.57 | 0.68 | 0.2 | 0.28 | 0.77 | 0.39 |
| r ² monthly flows | >= 0.85 | >= 0.75 | >= 0.65 | < 0.65 | 0.85 | 0.8 | 0.81 | 0.8 | 0.71 | 0.76 | 0.28 | 0.53 | 0.86 | 0.68 |
| Slope of the regression line | >= 0.9 | >= 0.8 | >= 0.6 | < 0.6 | 0.72 | 0.77 | 0.66 | 0.72 | 0.58 | 0.72 | 0.39 | 0.36 | 0.76 | 0.37 |
| Percent Bias | <= 10% | <= 15 | <= 25 | > 25 | -0.51 | 3.14 | 5.53 | -4.24 | 5.74 | -0.07 | -5.49 | -4 | 6.15 | 4.46 |
| RSR | <= 0.50 | <= 0.60 | <= 0.70 | > 0.70 | 0.52 | 0.53 | 0.57 | 0.61 | 0.65 | 0.57 | 0.98 | 0.86 | 0.48 | 0.78 |
| NSE | >= 0.75 | >= 0.65 | >= 0.50 | < 0.50 | 0.73 | 0.59 | 0.67 | 0.63 | 0.57 | 0.68 | 0.03 | 0.26 | 0.77 | 0.39 |

Table 4.3: Water yield proportion-weighted average validation statistics for three research watersheds (OR, SMR, WR) for the respective calibration periods.

| Validation Statistic | Excellent | Good | Satisfactory | Unsatisfactory | All | OR | SMR | WR |
|----------------------------------|-----------|---------|--------------|----------------|--------|--------|--------|-------|
| % Difference of Mean Annual Flow | <= 5 | <= 10 | <= 15 | > 15 | -0.98 | -0.17 | -2.27 | -1.19 |
| Difference in standard deviation | <= 5 | <= 15 | <= 25 | > 25 | -14.13 | -12.16 | -30.34 | -0.51 |
| r ² daily flows | >= 0.80 | >= 0.70 | >= 0.60 | < 0.60 | 0.74 | 0.73 | 0.71 | 0.78 |
| r ² monthly flows | >= 0.85 | >= 0.75 | >= 0.65 | < 0.65 | 0.85 | 0.86 | 0.83 | 0.85 |
| Slope of the regression line | >= 0.9 | >= 0.8 | >= 0.6 | < 0.6 | 0.77 | 0.75 | 0.72 | 0.85 |
| Percent Bias | <= 10% | <= 15 | <= 25 | > 25 | 0.97 | 0.17 | 2.28 | 1.18 |
| RSR | <= 0.50 | <= 0.60 | <= 0.70 | > 0.70 | 0.51 | 0.52 | 0.53 | 0.47 |
| NSE | >= 0.75 | >= 0.65 | >= 0.50 | < 0.50 | 0.73 | 0.73 | 0.71 | 0.78 |

Table 4.4: Water yield proportion-weighted average validation statistics for three research watersheds (OR, SMR, WR) for the baseline period 1971 – 2000.

| Validation Statistic | Excellent | Good | Satisfactory | Unsatisfactory | All | OR | SMR | WR |
|----------------------------------|-----------|---------|--------------|----------------|--------|--------|--------|--------|
| % Difference of Mean Annual Flow | <= 5 | <= 10 | <= 15 | > 15 | -2.10 | -2.19 | 0.46 | -4.55 |
| Difference in standard deviation | <= 5 | <= 15 | <= 25 | > 25 | -16.68 | -17.14 | -14.14 | -18.85 |
| r ² daily flows | >= 0.80 | >= 0.70 | >= 0.60 | < 0.60 | 0.67 | 0.67 | 0.64 | 0.71 |
| r ² monthly flows | >= 0.85 | >= 0.75 | >= 0.65 | < 0.65 | 0.79 | 0.79 | 0.73 | 0.83 |
| Slope of the regression line | >= 0.9 | >= 0.8 | >= 0.6 | < 0.6 | 0.69 | 0.68 | 0.69 | 0.70 |
| Percent Bias | <= 10% | <= 15 | <= 25 | > 25 | 2.45 | 2.19 | -0.46 | 4.55 |
| RSR | <= 0.50 | <= 0.60 | <= 0.70 | > 0.70 | 0.56 | 0.57 | 0.60 | 0.53 |
| NSE | >= 0.75 | >= 0.65 | >= 0.50 | < 0.50 | 0.66 | 0.65 | 0.64 | 0.71 |

4.2 Water Yield Analysis

Water yield is here defined as the total annual runoff volume from a watershed.

4.2.1 Spatial and temporal trends of water yields

Table 4.5: Water yield summary for the Oldman Reservoir watershed and five sub-watersheds for the Baseline (1971-2000) and Future (2041-2070) periods.

| HYDAT Station # | 05AA022 | 05AA008 | 05AA027 | 05AA026 | 05AA023 | Oldman Reservoir |
|---|--------------|-----------------|-----------------|-------------|--------------|------------------|
| HYDAT Station Name | Castle River | Crowsnest River | Racehorse Creek | Dutch Creek | Oldman River | Total Watershed |
| HRU # | 314 | 492 | 613 | 716 | 1111 | 1706 |
| Area (km ²) | 820.3 | 401.5 | 215.9 | 142.4 | 1091.0 | 4390.0 |
| Areal Proportion of Total Watershed (%) | 18.7 | 9.1 | 4.9 | 3.2 | 24.9 | 100.0 |
| Baseline (1971-2000) | | | | | | |
| Observed Water Yield (million m ³) | 460 | 145 | 71.9 | 54.2 | 395 | - |
| Simulated Water Yield (million m ³) | 463 | 141 | 76.6 | 52.0 | 336 | 1542 |
| Difference (%) | 0.5 | -3.1 | -6.1 | -4.1 | -7.3 | - |
| Water Yield Proportion relative to WR (%) | 30.0 | 9.1 | 5.0 | 3.4 | 21.8 | 100.0 |
| Future (2041-2070) | | | | | | |
| Simulated Water Yield: CRCM (million m ³) | 481.0 | 135.0 | 63.9 | 46.9 | 324 | 1482 |
| Difference relative to Simulated 1971-2000 (%) | 3.9 | -4.3 | -16.6 | -9.8 | -3.6 | -3.9 |
| Simulated Water Yield: HRM3 (million m ³) | 439.0 | 132.0 | 66.6 | 47.0 | 348.0 | 1482 |
| Difference relative to Simulated 1971-2000 (%) | -5.2 | -6.4 | -13.1 | -9.6 | 3.6 | -3.9 |
| Simulated Water Yield: RCM3 (million m ³) | 544.0 | 171.0 | 85.3 | 60.5 | 439.4 | 1867 |
| Difference relative to Simulated 1971-2000 (%) | 17.5 | 21.3 | 11.4 | 16.3 | 30.8 | 21.1 |

Table 4.6: Water yield summary for the St. Mary Reservoir watershed and three sub-watersheds for the Baseline (1971-2000) and Future (2041-2070) periods.

| HYDAT Station # | 05AE027 | 05AE005 | 05AE002 | St. Mary Reservoir |
|---|-------------|-------------|-----------|--------------------|
| HYDAT Station Name | St. Mary R. | Rolph Creek | Lee Creek | Total Watershed |
| HRU # | 341 | 391 | 463 | 545 |
| Area (km ²) | 1206.4 | 222.4 | 312.3 | 2286.2 |
| Areal Proportion of Total Watershed (%) | 52.8 | 9.7 | 13.7 | 100.0 |
| Baseline (1971-2000) | | | | |
| Observed Water Yield (million m ³) | 570.3 | 6.6 | 50.5 | - |
| Simulated Water Yield (million m ³) | 571.2 | 7.8 | 52.6 | 716.5 |
| Difference (%) | 0.2 | 18.4 | 4.0 | - |
| Water Yield Proportion relative to WR (%) | 79.7 | 1.1 | 7.3 | 100.0 |
| Future (2041-2070) | | | | |
| Simulated Water Yield: CRCM (million m ³) | 540.3 | 4.9 | 42.7 | 762.4 |
| Difference relative to Simulated 1971-2000 (%) | -5.4 | -37.2 | -18.8 | 6.4 |
| Simulated Water Yield: HRM3 (million m ³) | 658.1 | 11.4 | 59.5 | 817.6 |
| Difference relative to Simulated 1971-2000 (%) | 15.2 | 46.2 | 13.1 | 14.1 |
| Simulated Water Yield: RCM3 (million m ³) | 645.0 | 10.7 | 62.9 | 807.5 |
| Difference relative to Simulated 1971-2000 (%) | 12.9 | 37.2 | 19.6 | 12.7 |

Table 4.7: Water yield summary for the Waterton Reservoir watershed and two sub-watersheds for the Baseline (1971-2000) and Future (2041-2070) periods.

| HYDAT Station # | 05AD003 | 05AD010 | Waterton Reservoir |
|---|-------------|------------|--------------------|
| HYDAT Station Name | Waterton R. | Drywood C. | Total Watershed |
| HRU # | 256 | 376 | 505 |
| Area (km ²) | 611.6 | 238.3 | 1273.7 |
| Areal Proportion of Total Watershed (%) | 48.0 | 18.7 | 100.0 |
| Baseline (1971-2000) | | | |
| Observed Water Yield (million m ³) | 541 | 92.6 | - |
| Simulated Water Yield (million m ³) | 512 | 90.9 | 741 |
| Difference (%) | -5.3 | -1.9 | - |
| Water Yield Proportion relative to WR (%) | 69.1 | 12.3 | 100.0 |
| Future (2041-2070) | | | |
| Simulated Water Yield: CRCM (million m ³) | 490 | 86 | 707 |
| Difference relative to Simulated 1971-2000 (%) | -4.4 | -5.1 | 1.1 |
| Simulated Water Yield: HRM3 (million m ³) | 467 | 91.7 | 708 |
| Difference relative to Simulated 1971-2000 (%) | -8.9 | 0.8 | -4.5 |
| Simulated Water Yield: RCM3 (million m ³) | 517 | 100.4 | 774 |
| Difference relative to Simulated 1971-2000 (%) | 1.1 | 10.5 | 4.4 |

4.3 Exceedance probability analysis of water yield

All final simulated water yields for the reservoir inflows were divided into 30-year climate normals and exceedance probabilities were plotted (Figures 4.1, 4.2, 4.3). The variation in water yield for the historical climate normal vary quite little. The simulations of future water yields reveal that the median (50% probability) water yields will stay similar in magnitude, with the exception of the RCM3 projection

for the ORB and SMRW, which are, respectively, 250 and 50 million m³ higher, and HRM3 projection, which is about 50 million m³ lower.

Focusing on the 1:10-year and 1:20-year low water yield values (exceedance probability of 10 and 5% respectively), future projection show a general trend of stable or increasing water yields. However, the CRCM projection show the lowest water yields for the 1:20-year low water yield for all three watersheds. The only other lowering of 1:20-year water yield is simulated using the HRM3 projection for the ORW (Figure 4.1).

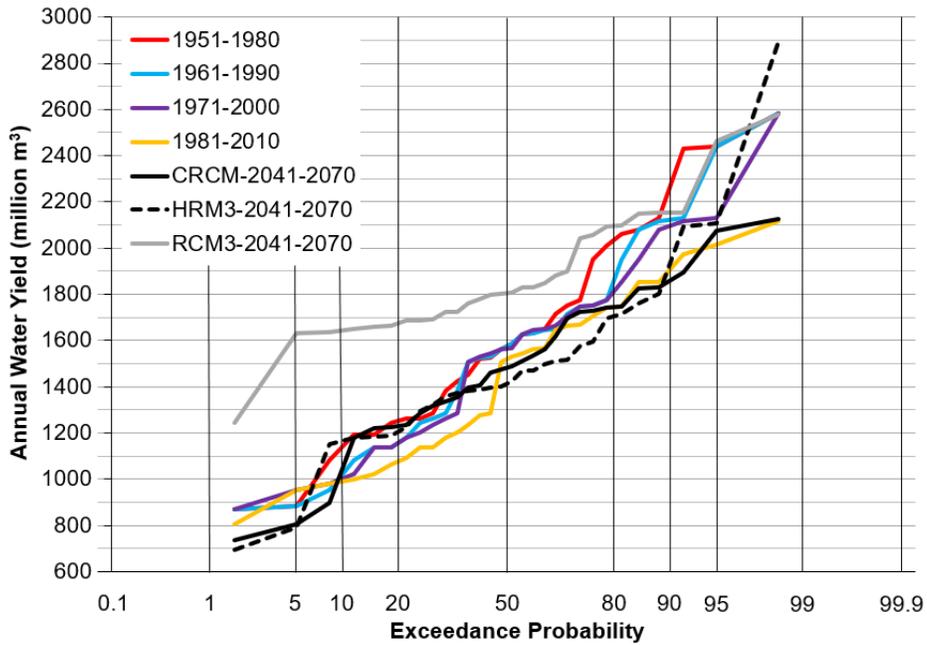


Figure 4.1: Exceedance probabilities of annual water yields for the Oldman Reservoir Watershed

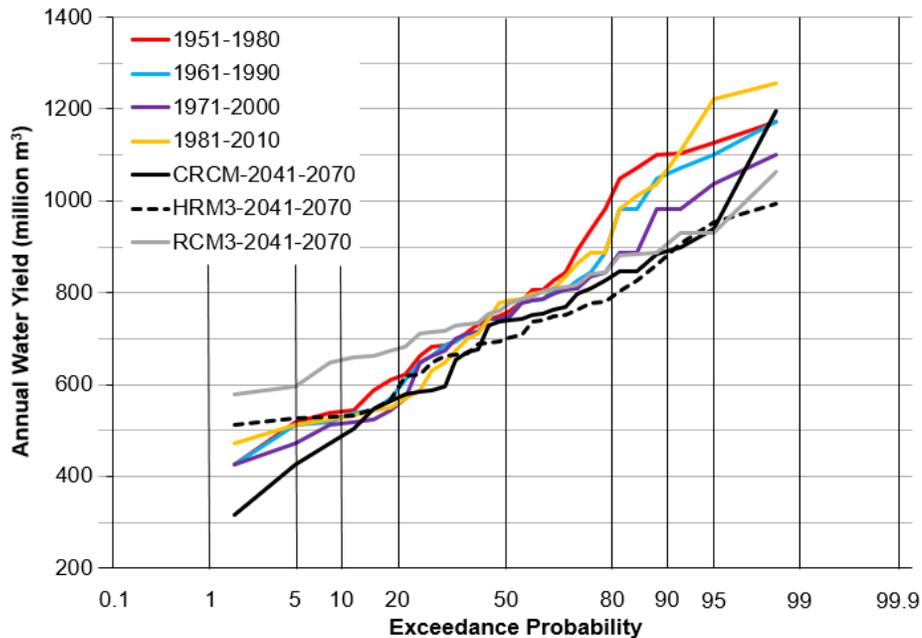


Figure 4.2: Exceedance probabilities of annual water yields for the Waterton Reservoir Watershed

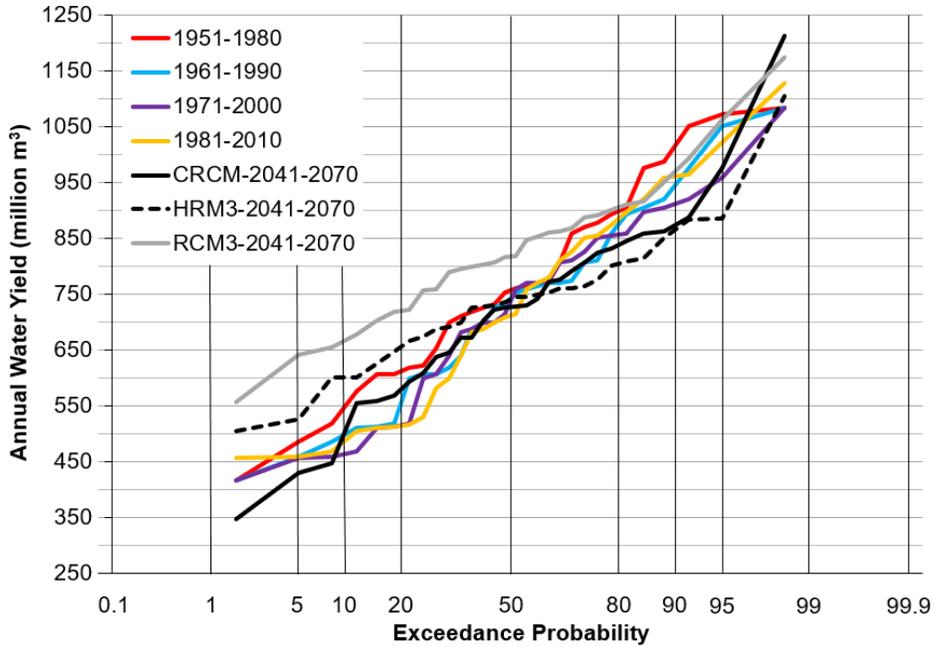


Figure 4.3: Exceedance probabilities of annual water yields for the St. Mary Reservoir Watershed

5 Conclusions

The ACRU-agrohydrological modelling system was set up and verified against 11 streamflow gauges for three watersheds (Oldman Reservoir, Waterton Reservoir, St. Mary Reservoir) upstream of three major irrigation reservoirs in southern Alberta. Verification analyses, summarized in Tables 4.3 and 4.4, reveals that the ACRU model could mimic the watershed behaviour to provide realistic annual water yields for the baseline period 1971-2000. Subsequently, the ACRU model was used with confidence to simulate future water yields under three future scenarios, ranging from relatively cool and wet to relatively warm and dry.

Resulting water yields are summarized in Tables 4.5, 4.6, and 4.7), revealing stable average future conditions for the Waterton Reservoir, with the CRM3 scenario increasing by 1.1%, the HRM3 scenario decreasing by 4.5%, and the RCM3 scenario increasing by 4.4%.

Based on Figures 4.1 to 4.3, there is potential for both increased demand for flood mitigation measures during wet years and increased frequency of dry years. Based on this analysis, future 1:20-year high and low water yields are within the order of magnitude as historical events. However, it must be stated that, going beyond this analysis, that extreme floods and droughts of magnitudes much larger than observed since 1950 must be anticipated due to the nature of climate change related weather extremes.

6 References

- Anderson T 2016: Hydrological implications of climate change on the Castle River Watershed, Alberta, Canada. MSc Thesis, University of Lethbridge.
- Barrow EB and Sauchyn DJ 2017: An analysis of the performance of RCMs in simulating current climate over Western Canada. *International Journal of Climatology* 37 (Suppl. 1): 640–58.
- Bathurst JC, Ewen J, Parkin G, O’Connell PE and Cooper JD 2004: Validation of catchment models for predicting land-use and climate change impacts. 3. Blind validation for internal and outlet responses. *J. Hydrol.* 287, 74–94.
- Beckers J, Smerdon B and Wilson M 2009: Review of hydrologic models for forest management and climate change applications in British Columbia and Alberta. FORREX Forum for Research and Extension in Natural Resources, Kamloops, BC. FORREX Series No. 25.
- Bonifacio, C 2016: Developing ACRU Utilities for modelling future water availability: A case study of the Oldman Reservoir Watershed, Alberta. MSc Thesis, University of Lethbridge.
- Bicknell BR, Imhoff JC, Kittle JL, Donigian AS and Johanson RC 1993: Hydrological Simulation Program - Fortran (HSPF): User Manual for Release 11.0. EPA/600/R-97/080, US Environmental Protection Agency, Environmental Research Laboratory, Athens, GA.
- Dent MC, Schulze RE and Angus GR 1988: Crop water requirements, deficits and water yield for irrigation planning in southern Africa. Dept. Agricultural Engineering, University of Natal, Pietermaritzburg, ACRU Report 28, WRC Report 118/1/88, Water Research Commission, Pretoria.
- Daly C 2002: Variable influence of terrain on precipitation patterns: Delineation and use of effective terrain height in PRISM. Corvallis, OR: Oregon State University. <http://www.prism.oregonstate.edu/pub/prism/docs/effectiveterrain-daly.pdf>
- Daly C, Halbleib M, Smith JI, Gibson WP, Doggett MK, Taylor JH, Curtis J, and Pasteris PA 2008: Physiographically sensitive mapping of temperature and precipitation across the conterminous United States. *International Journal of Climatology* 28(15): 2031—2064.
- Daly, C, Neilson RP and Phillips DL 1994: A statistical topographic model for mapping climatological precipitation over mountainous terrain. *Journal of Applied Meteorology* 33 (2): 140—158.
- Environment Canada 2012: Historical climate data. <http://climate.weather.gc.ca/>.
- Everson CS 2001: The water balance of a first order catchment in the montane grasslands of South Africa. *J. Hydrol.* 241, 110–123.
- Flügel WA 1995: Delineating hydrological response units by geographical information system analyses for regional hydrological modelling using PRMS/MMS in the drainage basin of the River Bröl, Germany. *Hydrological Processes*, 9(3-4), 423-436.
- Hopkinson RF, McKenney DW, Milewska EJ, Hutchinson MF, Papadopol P and Vincent LA 2011: Impact of aligning climatological day on gridded daily maximum-minimum temperature and precipitation over Canada. *Journal of Applied Meteorology and Climatology* 50: 1654—1665.
- Hutchinson MF 2004: ANUSPLIN Version 4.4 User Guide. Canberra: Fenner School of Environment and Society, Australian National University. <http://fennerschool.anu.edu.au/files/anusplin44.pdf>.
- Hutchinson MF and Gessler PE 1994: Splines—more than just a smooth interpolator. *Geoderma* 62 (1—3): 45—67.
- Hutchinson MF, McKenney DW, Lawrence K, Pedlar GH, Hopkinson RF, Milewska E and Papadopol P 2009: Development and testing of Canada-wide interpolated spatial models of daily minimum/maximum temperature and precipitation 1961—2003. *Journal of Applied Meteorology and Climatology* 48: 725—741.
- IPCC-TGICA 2007: General Guidelines on the Use of Scenario Data for Climate Impact and Adaptation

- Assessment. Version 2. Prepared by TR Carter on behalf of the Intergovernmental Panel on Climate Change, Task Group on Data and Scenario Support for Impact and Climate Assessment, 66pp.
- Kienzle SW 1993: Application of a GIS for simulating hydrological responses in developing regions. In: HydroGIS 93: Application of Geographical Information Systems in Hydrology and Water Resources Management (Proc. of the Vienna Conference, Austria, April 1993). IAHS Publications No. 211, pp. 309–318.
- Kienzle SW 1996: Using DTMs and GIS to define input variables for hydrological and geomorphological analysis. In: HydroGIS 96: Application of Geographical Information Systems in Hydrology and Water Resources Management (Proc. Of the Vienna Conference, Austria, April 1996). IAHS Publications No. 235, pp. 183–190.
- Kienzle SW 2008: A new temperature based method to separate rain and snow. *Hydrol. Processes* 22, 5067–5085.
- Kienzle SW 2010: Effects of area under-estimations of sloped mountain terrain on simulated hydrological behaviour. *Hydrol. Processes* 25: 1212-1227.
- Kienzle SW and Clark C 2020: Alberta Climate Records – Visualizing climate change past, present and future. albertaclimaterecords.com (accessed Jan. 2020)
- Kienzle SW and Mueller M 2013: Mapping Alberta’s Surface Water Resources for the period 1971-2000. *The Canadian Geographer* 57(4): 506-518.
- Kienzle SW and Nemeth MW, Byrne JM and MacDonald RJ 2012: Simulating the hydrological impacts of climate change in the upper North Saskatchewan River basin, Alberta, Canada. *Journal of Hydrology* 412-413: 76-89.
- Kienzle SW and Schmidt J 2008: Hydrological impacts of irrigated agriculture in the Manuherikia Catchment, Otago, New Zealand. *J. Hydrol. (NZ)* 47, 67–83.
- Kienzle SW and Schulze RE 1991: The simulation of the effect of afforestation on shallow ground water in deep sandy soils. *Water SA* 18, 265–272.
- Kienzle SW, Lorenz SA and Schulze RE 1997: Hydrology and Water Quality of the Mgeni Catchment. Water Research Commission, Pretoria, Report TT87/97, pp. 1–88.
- Kiker GA, Clark DJ, Martinez CJ and Schulze RE 2006: A Java-based, object-oriented modeling system for southern African hydrology. *Trans. ASABE* 49, 1419–1433.
- Kouwen N 2014: WATFLOOD Hydrological Model Routing & Flow Forecasting System: Reference Manual. University of Waterloo, Waterloo, Ontario, Canada. 267 pp.
- Lapp S, Toth B and Sauchyn DJ 2009: Constructing scenarios of future climate and water supply for the SSRB: Use and limitations for vulnerability assessment. *Prairie Forum*, 34(1), 153-180.
- Lapp S, Sauchyn DJ and Toth B 2009: Constructing scenarios of future climate and water supply for the SSRB: Use and limitations for vulnerability assessment. *Prairie Forum*, 34(1), 153-180.
- Legates DR and McCabe GJ 1999: Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resources Res.* 35(1): 233-241.
- Loukas A, Vasiliades L and Dalezios NR 2002: Potential climate change impacts on flood producing mechanisms in southern British Columbia, Canada using the CGCMA1 simulation results. *J. Hydrol.* 259, 163–188.
- Martinez CJ, Campbell KL, Annable MD and Kiker GA 2008: An object-oriented hydrologic model for humid, shallow water-table environments. *J. Hydrol.* 351, 368–381.
- McKenney DW, Pedlar JH, Papadopol P and Hutchinson MF 2006: The development of 1901—2000 historical monthly climate models for Canada and the United States. *Agricultural and Forest Meteorology* 138(1-4): 69—81.
- Mearns LO, Gutowski WJ, Jones R, Leung LY, McGinnis S, Nunes AMB and Qian Y 2009: A regional climate change assessment program for North America. *EOS*, Vol. 90, No. 36, 8 September 2009, pp.

311-312.

- Moriasi D, Arnold JG, Van Liew MW, Bingner RL, Harmel RD and Veith TL 2007: Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE*, 50(3), 885-900.
- Moriasi D, Wilson BN, Douglas-Makin KR, Arnold JG and Gowda PH 2012: Hydrologic and water quality models: use, calibration, and validation. *Transactions of the ASABE*. 55(4): 1241-1247.
- Murphy J 1999: An Evaluation of Statistical and Dynamical Techniques for Downscaling Local Climate. *Journal of Climate*, 12(8), 2256.
- Murphy J 2000: Predictions of climate change over Europe using statistical and dynamical downscaling techniques. *International Journal of Climatology*, 20(5), 489-501.
- Nakicenovic N, Alcamo J, Davis G, de Vries B, Fenhann J, Gaffin S 2000: Special Report on Emissions Scenarios: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK, 599pp.
- Neitsch SL, Arnold JG, Kiniry JR, Srinivasan R and Williams JR 2002: Soil and Water Assessment Tool – User’s Manual, Version 2000. Temple, Texas: USDA–ARS Grassland, Soil and Water Research Laboratory
- Nemeth MW, Kienzle SW and Byrne JM 2012: Multi-variable verification of hydrological processes in the upper North Saskatchewan River basin, Alberta, Canada. *Hydrological Sciences Journal* 57(1): 84-102.
- New MG 2002: Climate change and water resources in the southwestern Cape, South Africa. *S. Afr. J. Sci.* 98, 1–8. New, M.G., Schulze, R.E., 1996. Hydrologic sensitivity to climate change in the Langrivier catchment, Stellenbosch, South Africa and some implications for geomorphic processes. *Z. Geomorphol.* 107, 11–34.
- New MG and Schulze RE 1996: Hydrologic sensitivity to climate change in the Langrivier catchment, Stellenbosch, South Africa and some implications for geomorphic processes. *Zeitschrift für Geomorphologie, Supplementband 107*: 11-34.
- Newlands KN, Davidson A, Howard A and Hill H 2011: Validation and inter-comparison of three methodologies for interpolating daily precipitation and temperature across Canada. *Environmetrics* 22(2): 205—223.
- Nurmohamed R, Naipal S and De Smedt F 2007: Modeling hydrological response of the Upper Suiname river basin to climate change. *J. Spatial Hydrol.* 7 (1), 1–22.
- Oldman Watershed Council 2019: Watershed Information. Retrieved from <https://oldmanwatershed.ca/annualreports> (accessed Jan. 2019)
- Parry M 2002: Scenarios for climate impact and adaptation assessment. *Global Environmental Change* 12: 149-153.
- Pipes A and Quick MC 1977: UBC Watershed model users guide. Department of Civil Engineering, University of British Columbia, Vancouver, British Columbia, Canada.
- Poitras V, Sushama L, Seglenieks F, Khaliq FN and Soulis E 2011: Projected Changes to Streamflow Characteristics over Western Canada as Simulated by the Canadian RCM. *Journal of Hydrometeorology* 12: 1395–413.
- Pomeroy JW, Gray DM, Brown T, Hedstrom NR, Quinton WL, Granger RJ and Carey S 2007: The cold regions hydrological model: a platform for basing process representation and model structure on physical evidence. *Hydrological Processes* 21: pp. 2650–2667
- Santhi C, Arnold JG, Williams JR, Dugas WA, Srinivasan R and Hauck LM 2001: Validation of the SWAT model on a large river basin with point and nonpoint sources. *J. American Water Resources Assoc.* 37(5): 1169-1188.
- Sevat E and Dezetter A 1991: Selection of calibration objective functions in the context of rainfall-runoff modeling in a Sudanese savannah area. *Hydrological Sci. J.* 36(4): 307-330.
- Schmidt J, Kienzle SW, Srinivasan MS 2009: Estimating increased evapotranspiration losses caused by

- irrigated agriculture as part of the water balance of the Orari Catchment, Canterbury, New Zealand. *J. Hydrol. (NZ)* 48, 89–94.
- Schmidt EJ, Schulze RE 1987: Flood Volume and Peak Discharge from Small Catchments in Southern Africa, based on the SCS Technique. Water Research Commission, Pretoria, Technical Transfer Report TT/3/87, 164pp.
- Schulze RE 1995: Hydrology and Agro-hydrology: A Text to Accompany the ACRU 3.00 Agro-hydrological Modelling System. Report TT69/95. Water Research Commission, Pretoria, RSA, p. 125.
- Schulze RE, Lorentz SA, Kienzle SW and Perks L 2004: Case Study 3: Modelling the impacts of land use and climate change on hydrological response in the mixed, underdeveloped/developed Mgeni catchment, Southern Africa. In: Kabat P, Claussen M, Dirmeyer PA, Gash JHC, Bravo de Guenni L, Meybeck M, Pielke Sr RA, Vörösmarty CJ, Hutjes RWA, Lütkenmeier S (Eds.), *Vegetation, Water, Humans and the Climate: A New Perspective on an Interactive System*. Springer, Heidelberg, Germany.
- Smithers J, Schulze RE 1995: ACRU Agrohydrological Modelling System User Manual. Water Research Commission, Report TT 70/95, Water Research Commission, Pretoria, Republic of South Africa.
- Smithers JC, Schulze RE, Kienzle SW 1997: Design flood estimation using a modelling approach: A case study using the ACRU model. In: Rosbjerg D, Boutayeb N, Gustard A, Kundzewicz ZW, Rasmussen PF (Eds.), *Sustainability of Water Resources under Increasing Uncertainty*. IAHS Publication vol. 240, pp. 277–286.
- Smithers JC, Schulze RE, Pike A, Jewitt GPW 2001: A hydrological perspective in the February 2000 floods: a case study in the Sabie river catchment. *Water SA* 27, 325–332.
- Tarboton KC and Schulze RE 1993: Hydrological consequences of development scenarios for the Mgeni catchment. In: Lorentz SA, Kienzle SW, Dent MC (Eds.), *Hydrology in Developing Regions. . . The Road Ahead*. Proceedings of the Sixth South African National Hydrological Symposium, University of Natal, Pietermaritzburg, Department of Agricultural Engineering, 297–304. ISBN: 1-86840-044-1.
- Teutschbein C and Seibert J 2012: Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *J. Hydrol.* 456-457, 12-29.
- Teutschbein C and Seibert J 2010: Regional climate models for hydrological impact studies at the catchment scale: a review of recent modelling strategies. *Geogr. Compass* 4(7), 834-860.
- Van Liew MW, Arnold JG and Garbrecht JD 2003: Hydrologic simulation on agricultural watersheds: Choosing between two models. *Trans. ASAE* 46(6): 1539-1551.
- Warburton ML, Schulze RE and Jewitt GPG 2010: Confirmation of ACRU model results for applications in land use and climate change studies. *Hydrol. Earth Syst. Sci.* 14, 2399–2414.
- Wilby RL, Charles SP, Zorita E, Timbal B, Whetton P and Mearns LO 2004: Guidelines for use of climate scenarios developed from statistical downscaling methods.
- Wilby RL and Wigley TML 1997: Downscaling general circulation model output: a review of methods and limitations. *Progress in Physical Geography*, 21(4), 530-548.
- Wigmosta MS, Nijssen B, Storck P and Lettenmaier DP 2002: The Distributed Hydrology Soil Vegetation Model, In *Mathematical Models of Small Watershed Hydrology and Applications*, VP Singh, DK Frevert, eds., Water Resource Publications, Littleton, CO., p. 7-42.

7 MAPPING FUNDAMENTAL SOIL PROPERTIES IN THE WAST CASTLE WATERSHED, ALBERTA

In Rocky Mountain watershed there is no to extreme limited soil property information available. Soil information is important for environmental research and other activities in the watershed. Soil hydrological properties govern the processes of infiltration, transpiration, percolation, and groundwater recharge. While soils are an essential element which determine plant available water and runoff processes, mountain soils are poorly studied, and the quantification of soil hydrological parameters remain to be highly uncertain.

Soil measurements, soil sampling and subsequent lab analyses were carried out at 131 randomly sampled sites in the Castle River Watershed. Data were compiled for the position and magnitude of hydrologically important soil properties: soil restrictive depth, soil texture, soil organic matter, and solum depth. Stepwise linear regression was used to find relationships between soil and environmental relationships (land cover, terrain, solar radiation). The linear regression models are used to create maps of predicted spatial distribution of soil properties throughout the entire watershed. The final objective is comparing sampled and modelled data to soil survey and globally predicted soil information to validate sampled data and determine the improvement of knowledge gained. These objectives lead to the overall objective of increasing the knowledge of soil properties to aid environmental research in the study area.

It was found that soils are deeper than previously thought, that soils are generally very sandy, and that the soils generally contain high levels of organic matter. All soils are highly variable, which resulted in relatively weak statistical results from the regression models. Models reveal that vegetation type, vegetation height and density, terrain slope, terrain position and geology are the most important variables in predicting soil properties.

An attempt was made to create detailed soils maps for the study area based on the established regression models.

7.1 Methods

In order to ensure that representative soil sampling was carried out, a pre-study was carried out using GIS overlay analysis to divide the Castle River Watershed into soil units, based on perceived factors for soil genesis: Land cover, wetness index as an index for terrain position, and solar radiation. Within the resulting soil units, a GIS selected randomly located soil sampling locations, which were then visited using a GPS, and soil measurements and samples were taken as close to the locations as feasible (Figure 7.1). Soil lab analyses were carried out following methods in *Soil Sampling and Methods of Analysis* by Carter & Gregorich (2008). Soil texture and soil organic matter were compared against professional soil labs for independent quality control.

Field measurements included restrictive soil depth, rooting depth, and depth of the A-horizon. Samples were collected from both the A and the B horizons of sample sites. Percent sand, silt, and clay were measured in the lab for each horizon.

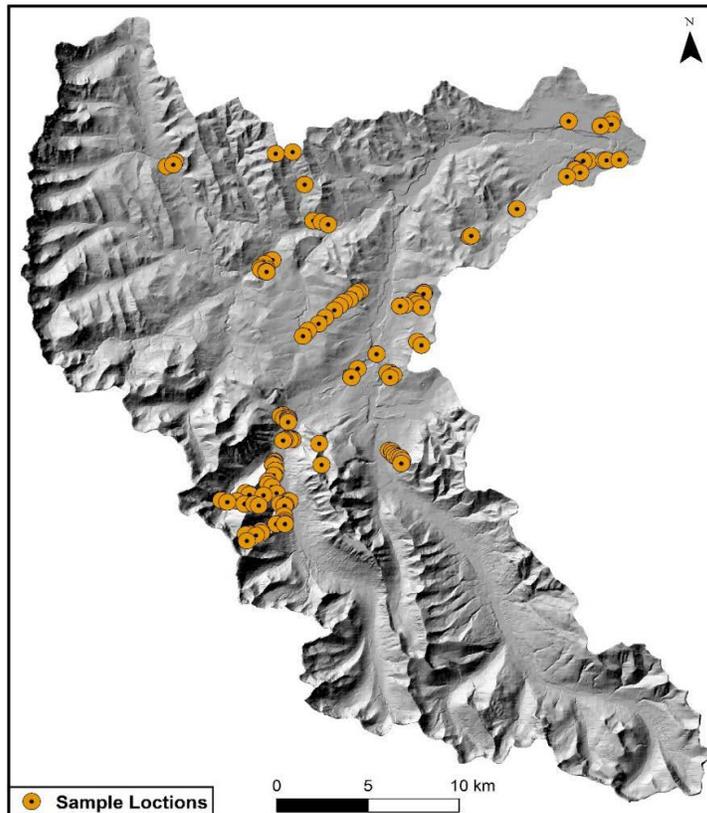


Figure 7.1: The Castle River Watershed and 131 soil sampling sites

Stepwise multiple linear regression was the statistical method chosen to analyze relationships between soil properties (dependent variables) and environmental characteristics (independent variables). Stepwise linear regression was chosen because it has been used to model soil property spatial distribution (Grunwald, 2009). It was established that stepwise multiple regression performed similarly to Random Forest and other methods (Forkuor et al., 2017; Hitziger and Lieb, 2014; Zhang et al., 2012). Independent variables were chosen based on literature and possible influential contributions to soil property formation (Table 7.1).

Based on the best regression model, established for soil depth, sand, silt and clay content, and organic matter content, stepwise linear regression models were established, from which spatial models were constructed in ArcGIS to map each soil variable across the Castle River Watershed.

7.2 Results

Field measurements and laboratory analyses results are summarized in Table 7.2 (Deering, 2019). Soil restrictive depth sampling resulted in an average of 74 cm with a minimum of 3 cm and a maximum of 182 cm, and 50 % of the data deeper than 63 cm. This information exceeded pre-field operation expectations that the average soil restrictive depth would be approximately 50 cm. Thus, soils were deeper than expected. This is of great importance for water storage understanding because, in theory, deeper soils have the capacity to hold more water than shallower soils (Shukla, 2014, p. 95). No distinct spatial patterns are seen by viewing the data in relation to its spatial location alone.

Table 7.1: List of dependent and independent variables used for the stepwise regression analysis (Deering, 2019)

| Dependent Variables | Data type | Independent variables | Data Type |
|----------------------------|------------------|-------------------------------------|------------------|
| Restrictive Depth | Ratio | Topographic Wetness Index | Ratio |
| Mineral BD | Ratio | Slope (%) | Ratio |
| Litter BD | Ratio | Elevation (m) | Ratio |
| %S, %Si, %Cl | Ratio | Solar Radiation (W/m ²) | Ratio |
| A and B horizon depths | Ratio | Aspect (degrees) | Ratio |
| | | Profile Curvature | Ratio |
| | | Plan Curvature | Ratio |
| | | Curvature | Ratio |
| | | Specific Catchment Area | Ratio |
| | | Contributing Area | Ratio |
| | | Landform Position | Categorical |
| | | Slope Position | Categorical |
| | | Surficial Geology | Categorical |
| | | Bedrock Geology | Categorical |
| | | Land Cover Type | Categorical |
| | | Vegetation Density | Categorical |
| | | Vegetation Type | Categorical |
| | | Vegetation Height | Categorical |

Texture analyses for both A- and B-horizons were carried out for 104 sites. Descriptive statistics reveal a mean percent sand of 67%, a minimum of 32 % and a maximum of 98%. These results suggest that the soils are predominantly sandy soils. The mean percentage of silt was determined to be 28%, ranging between 0% and 62%. Further, mean percent clay was 4.4%, with a minimum of 0% and a maximum of 18%. These data all lead to the conclusion that the A horizon soils of the watershed are generally high in sand content and low in clay content, as is supported by the soil textural triangle (Figure 7.2).

Soil samples taken from the B horizon were processed for 69 sites. Mean percent sand in the B horizon is 69%, with a minimum 29% and a maximum of 97%. The mean silt content was 27%, ranging from 0% to 63%. The mean clay content was 4.4%, with a minimum of 0%, a maximum of 18% percent. As the A-horizon, the B-horizon is also very sandy, with similar overall descriptive statistics results.

Assessing the descriptive statistics for organic matter results it is evident that there is more organic matter in the A horizon than in the B horizon. This suggests that the water holding capacity is increased due to the organic matter content, thereby slightly offsetting the poor capacity of the sandiness of the soils. The high organic matter content is a result of large areas of mixed forests that accumulate much

litter from deciduous and coniferous leafy and woody deposits. This material is deposited each fall and decomposes into more stable soil organic carbon elements. Arguably a large influence on the OM is the highly productive forest vegetation which is supported by a complex root system leading to large amounts of root break down into OM additions (Fisher et al., 2000).

Table 7.2: Descriptive statistics of field and laboratory property measurements.

| Variable | N | Mean | SD | Min | Max | 25 th Percentile | 50 th Percentile | 75 th percentile |
|----------------------------|-----|-------|-------|------|--------|-----------------------------|-----------------------------|-----------------------------|
| Soil restrictive depth (m) | 118 | 74.00 | 41.25 | 3.0 | 182 | 41.9 | 63.4 | 104.4 |
| %Sand A | 104 | 67.3 | 14.8 | 31.5 | 97.6 | 56.0 | 67.6 | 77.4 |
| %Sand B | 69 | 68.6 | 16.04 | 28.9 | 97.1 | 58.8 | 68.1 | 81.2 |
| %Silt A | 104 | 28.28 | 13.49 | 0.00 | 62.32 | 20.14 | 26.64 | 38.21 |
| %Silt B | 69 | 26.56 | 14.50 | 0.00 | 66.27 | 16.20 | 26.35 | 37.13 |
| %Clay A | 104 | 4.40 | 3.28 | 0.00 | 17.86 | 2.33 | 3.74 | 6.08 |
| %Clay B | 69 | 4.85 | 3.74 | 0.00 | 22.80 | 2.39 | 4.75 | 5.92 |
| %OM A | 104 | 9.34 | 5.55 | 2.41 | 47.08 | 5.90 | 8.21 | 11.35 |
| %OM B | 69 | 6.17 | 3.60 | 2.12 | 23.39 | 3.78 | 5.97 | 6.88 |
| Solum Depth (cm) | 131 | 36.15 | 29.91 | 0.00 | 115.66 | 10.50 | 30.00 | 54.90 |

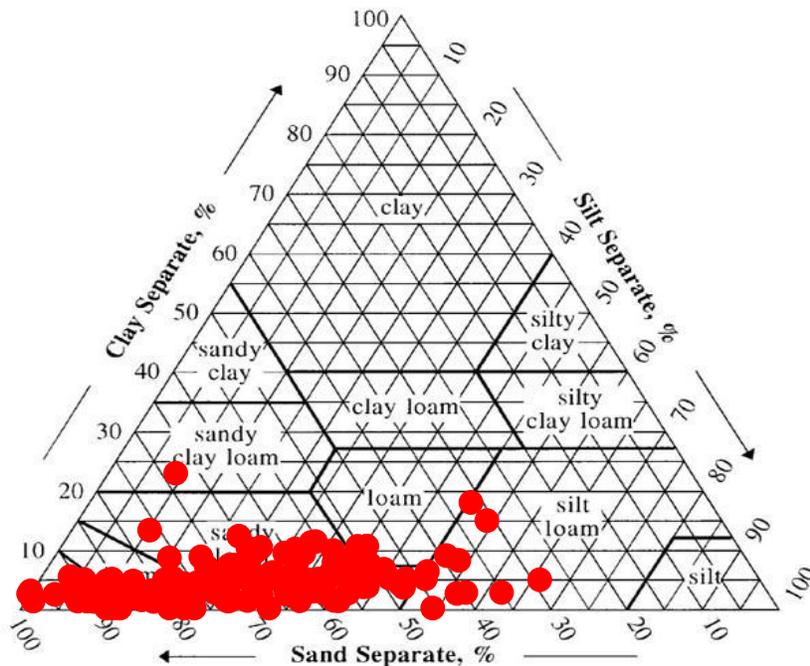


Figure 7.2: Soil textural triangle of A and B horizon texture classes.

An example of the stepwise regression model is given in Table 7.3 for soil restrictive depth (Deering, 2019).

Before a number of stepwise regression models were assembled, the data were tested for spatial autocorrelation. Investigation of Moran's I spatial autocorrelation statistic revealed that several soil variables were significantly clustered, therefore, further investigation of autocorrelation was needed. Semi-variogram analysis was conducted and it was found that all variables revealed weak spatial autocorrelation.

A soil restrictive depth model based on 17 variables was selected as the best model. Statistical performance of the training and validation data sets (Deering, 2019) reveals that the r^2 value of the validation data, using the training data formula, was very low ($r^2 = 0.006$). Also, the MAE is extremely large (42.63 cm), twice as large as the training data (20.12 cm). This poor modelling result is likely the result of a low sampling site when one regards the wide range of land cover, terrain and climate conditions found in a Rocky Mountain watershed. Further, soil restrictive depth is expected to change over micro-scales, rather than the sampled meso-scale.

The same poor regression statistics were found to be true for all three soil textures (sand, silt, clay) models, as well as the organic matter model. Low coefficients of determination, combined with a large root mean square errors, indicate that the application of a spatial model to map the spatial distributions of these soil properties across a highly variable terrain, while possible, is not very meaningful.

Table 7.3: Example of regression model: Model summary of 17 variable Soil restrictive depth regression model.

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Change Statistics | | | | Sig. F Change | Durbin-Watson |
|-------|-------------------|----------|-------------------|----------------------------|-------------------|----------|-----|-----|---------------|---------------|
| | | | | | R Square Change | F Change | df1 | df2 | | |
| 1 | .231 ^a | .053 | .044 | 40.769 | .053 | 5.478 | 1 | 97 | .021 | |
| 2 | .329 ^b | .108 | .090 | 39.773 | .055 | 5.921 | 1 | 96 | .017 | |
| 3 | .409 ^c | .167 | .141 | 38.644 | .059 | 6.689 | 1 | 95 | .011 | |
| 4 | .459 ^d | .211 | .177 | 37.822 | .043 | 5.174 | 1 | 94 | .025 | |
| 5 | .508 | .258 | .218 | 36.859 | .048 | 5.977 | 1 | 93 | .016 | |
| 6 | .540 | .292 | .246 | 36.202 | .034 | 4.407 | 1 | 92 | .039 | |
| 7 | .590 | .348 | .298 | 34.941 | .056 | 7.760 | 1 | 91 | .007 | |
| 8 | .630 | .397 | .343 | 33.789 | .049 | 7.309 | 1 | 90 | .008 | |
| 9 | .663 | .440 | .384 | 32.731 | .043 | 6.914 | 1 | 89 | .010 | |
| 10 | .693 | .480 | .421 | 31.714 | .040 | 6.799 | 1 | 88 | .011 | |
| 11 | .713 | .508 | .446 | 31.024 | .028 | 4.956 | 1 | 87 | .029 | |
| 12 | .727 | .528 | .462 | 30.581 | .019 | 3.544 | 1 | 86 | .063 | |
| 13 | .739 | .546 | .476 | 30.174 | .018 | 3.332 | 1 | 85 | .071 | |
| 14 | .760 | .577 | .507 | 29.280 | .032 | 6.272 | 1 | 84 | .014 | |
| 15 | .771 | .594 | .521 | 28.867 | .017 | 3.418 | 1 | 83 | .068 | |
| 16 | .780 | .608 | .532 | 28.531 | .014 | 2.966 | 1 | 82 | .089 | |
| 17 | .788 ^q | .621 | .542 | 28.220 | .013 | 2.817 | 1 | 81 | .097 | 2.016 |

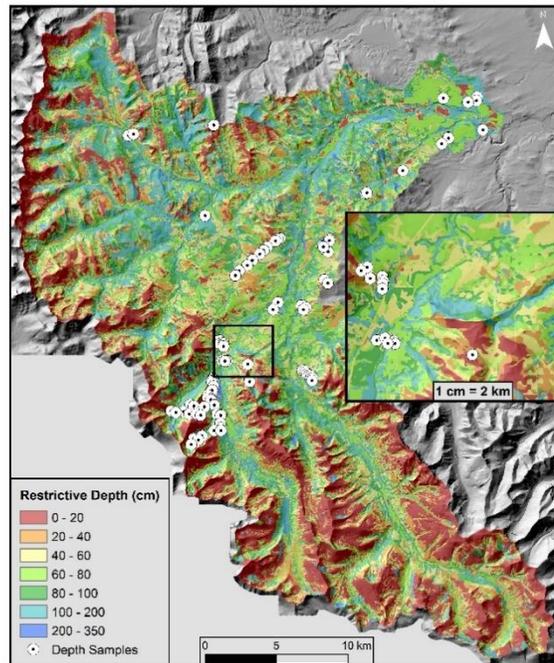


Figure 7.2: Example of mapped results: Restrictive soil depth based on regression model.

7.3 Conclusions

Soil information was gathered through random stratified sampling of 131 sites in the West Castle watershed. Soil lab results were collected through implementation of accepted lab methods, closely following methods in “Soil Sampling and Methods of Analysis” by Carter and Gregorich (2008). Soil texture and soil organic matter were compared against professional soil labs for insuring lab methods were conducted adequately. This information has greatly improved our understanding of soils in the study area. From early preliminary field scouting activities, it was hypothesized that the soils of the watershed would be highly sandy, and this hypothesis is supported by this data. We can conclude that the sandy soils are well drained and have a relatively low water holding capacity.

Implementation of stepwise linear regression allowed for the calculation of large-scale soil property predictions throughout the entire watershed. These predictions enable the extrapolation of the measured soil properties to the scale of the watershed so that the data can be used at a larger scale. Also, of great importance, predictions allow for the assessment of soil property spatial distribution for logical spatial relationships that are established and explainable by regression models.

Despite the failure to spatially map the observed soil properties with confidence, this research has provided detailed soil analyses results in diverse hydro-climatological and ecological locations. This provides new insight into soils in the Rocky Mountains. However, it is concluded that only several thousand additional soil samples are required to build a spatial model of soil properties that is representative.

Soil properties researched through this research project can aid in modelling hydrological responses resulting from climate changes, specifically where physical models use soil data. The ACRU agro-hydrological modelling system is one such model that uses soil information. Soil information is highly

important in the ACRU model due to the multi-layer soil water budget module (Aduah et al., 2017). Having data that was sampled and measured in the area can help to accept or reject the soil data produced through model default data and educated guesses. Kienzle et al. (2012) and Bonifacio (2016) suggest that improved soil information will increase the credibility of ACRU model outputs, strengthening confidence that the soil data derived from this research will aid in hydrological modelling efforts in the West Castle watershed.

Based on more refined and better calibrated spatial model, further land cover change research is facilitated. Climate induced land cover change research benefits from soil information because soil is an important factor in determining suitable locations for vegetation to grow (Zolkos et al., 2015). Soil moisture, soil temperature, and soil nutrients are important properties that aid tree growth (Butler et al, 2007). Geomorphic properties (Macias-Fauria and Johnson, 2013), on their own or coupled with soil properties, that influence these soil properties are also important to understanding forest habitat changes. Detailed soil texture and depth properties will aid in understanding soil moisture and temperature potential that influences vegetation movement/position.

Three main lessons of soil property research in mountainous regions were learned: 1) the soil sampling network did not represent the environmental variables governing soil genesis; 2) the sampling focus should be on vegetation type, height, and density; and 3) a more patterned sampling approach could be adopted for the effective implementation of geostatistical and other modelling techniques. Vegetation properties should be considered for stratifying a sampling regime due to the findings that suggest vegetation properties are highly important proxies for soil property occurrence in the watershed. Also, landform position, surface geology, and bedrock geology were important in models and should be used for stratifying soil mapping units.

7.4 References

Aduah MS, Jewitt GPW and Toucher MLW 2017: Assessing suitability of the ACRU hydrological model in a rainforest catchment in Ghana, West Africa. *Water Science*, 31(2), 198-214.

Bonifacio CMT 2016: Developing ACRU utilities for modelling future water availability: a case study of the Oldman Reservoir Watershed, Alberta. Lethbridge, Alta: University of Lethbridge, Dept. of Geography.

Butler DR, Malanson GP, Walsh SJ and Fagre DB 2007: Influences of Geomorphology and Geology on Alpine Treeline in the American West—More Important than Climatic Influences? *Physical Geography*, 28(5), 434-450.

Carter MR and Gregorich EG 2008: *Soil sampling and methods of analysis* (2nd ed.). Boca Raton, Fla: CRC Press.

Deering T 2019: Mapping fundamental soil properties in the West Castle Watershed, Alberta. MSc thesis, University of Lethbridge.

Fischer G, Tubiello FN, van Velthuis H and Wiberg DA 2007: Climate change impacts on irrigation water requirements: Effects of mitigation, 1990–2080. *Technological Forecasting and Social Change*, 74(7), 1083-1107.

Fisher RF, Binkley D and Pritchett WL 2000: Ecology and management of forest soils (Vol. 3rd --). New York, Toronto: John Wiley.

Forkuor G, Hounkpatin OKL, Welp G and Thiel M 2017: High Resolution Mapping of Soil Properties Using Remote Sensing Variables in South-Western Burkina Faso: A Comparison of Machine Learning and Multiple Linear Regression Models. PLoS ONE, 12(1), e0170478.

Grunwald S 2009: Multi-criteria characterization of recent digital soil mapping and modeling approaches. Geoderma, 152(3), 195-207.

Hitziger M and Lieb M 2014: Comparison of Three Supervised Learning Methods for Digital Soil Mapping: Application to a Complex Terrain in the Ecuadorian Andes. Applied and Environmental Soil Science, 2014, 12.

Kienzle SW, Nemeth MW, Byrne JM and MacDonald RJ 2012: Simulating the hydrological impacts of climate change in the upper North Saskatchewan River basin, Alberta, Canada. Journal of Hydrology, 412–413, 76-89.

Macias-Fauria M and Johnson EA 2013: Warming-induced upslope advance of subalpine forest is severely limited by geomorphic processes. Proceedings of the National Academy of Sciences, 201221278.

Shukla M 2014: Soil physics: an introduction. Boca Raton: CRC Press.

Zhang S, Huang Y, Shen C, Ye H and Du Y 2012: Spatial prediction of soil organic matter using terrain indices and categorical variables as auxiliary information. Geoderma, 171–172, 35-43.

Zolkos SG, Jantz P, Cormier T, Iverson LR, McKenney DW and Goetz SJ 2015: Projected tree species redistribution under climate change: Implications for ecosystem vulnerability across protected areas in the eastern United States. Ecosystems, 18(2), 202-220.

8 Other info

8.1 MSc Theses completed for this project:

Charmaine Bonifacio, MSc. (2017), Dept. of Geography, U. of Lethbridge, Developing ACRU utilities for modelling future water availability: A case study of the Oldman Reservoir Watershed, Alberta.

Trevor Deering, MSc. (2019), Dept. of Geography, U. of Lethbridge, Mapping fundamental soil properties in the West Castle Watershed, Alberta.

Three manuscripts for peer-review are in preparation and mostly written.

8.2 Public and conference presentations

Kienzle SW 2019: Analyzing Impacts of Climate Change: Approaches, Challenges and Potential Solutions. Invited public presentation, University of KwaZulu-Natal, Centre for Water Research, April 30, 2019.

Kienzle SW 2019: On the potential of mapping climate indices to delineate current and future weather extremes and ecosystem changes. 17th Savanna Science Network Meeting. Skukuza March 3-7, 2019.

Kienzle SW and Clark C 2018: Creation of an Interactive Website for Mapping Changes of 43 Climate Indices for Alberta, Canada, for the Period 1950 to 2010. Universities Council on Water Resources Conference, Pittsburgh, PA, USA, June 26-28, 2018.

Kienzle SW 2018: Mapping changes of 43 climate indices at high spatial resolution for Alberta, Canada, for the period 1950 to 2010. In: Climate Data Homogenization and Analysis of Climate Variability, Trends and Extremes. Poster Presentation, EGU General Assembly, 7–12 April 2019, Vienna, Austria.

Kienzle SW 2018: Water security and climate change in the Oldman River basin. Lethbridge College Science Café, Nov. 07, 2018.

Kienzle SW 2018: Alberta Water under conditions of climate change. What should we do? Public Lecture and Panel Discussion at Pat Carlson Speaker Series: Climate change and the future of energy. Medicine Hat, Jan. 25, 2018.

Kienzle SW 2018: Alberta Water under conditions of climate change. What should we do? Public Lecture and Panel Discussion at Pat Carlson Speaker Series: Climate change and the future of energy. Lethbridge, Jan. 24, 2018.

Kienzle SW 2017: Mapping trends of 43 climate indices at high spatial resolution for Alberta, Canada, for the period 1950 – 2010. Canadian Association of Geographers, May 29-June 02, 2017, York University, Toronto.

Kienzle SW 2016: Did it really get warmer in Alberta between 1950 and 2010? Better Climate Information for a Better Future. Public presentation at the Galt Museum. Nov. 16, 2016. Invited presentation.

Kienzle SW 2016: How did the climate change between 1950 and 2010 and what does it mean? Public lecture at the Lethbridge Library. Invited presentation.

Kienzle SW 2016: High-resolution mapping of climate indices and their trends across Alberta, Canada, for the period 1950 to 2010. The 13th International Meeting on Statistical Climatology. June 6-10, 2016, Canmore, Alberta.