Northumbria Research Link

Citation: Meloche, Julien, Langlois, Alexandre, Rutter, Nick, McLennan, Donald, Royer, Alain, Billecocq, Paul and Ponomarenko, Serguei (2022) High-resolution snow depth prediction using Random Forest algorithm with topographic parameters: A case study in the Greiner watershed, Nunavut. Hydrological Processes, 36 (3). e14546. ISSN 0885-6087

Published by: Wiley-Blackwell

URL: https://doi.org/10.1002/hyp.14546 <https://doi.org/10.1002/hyp.14546>

This version was downloaded from Northumbria Research Link: https://nrl.northumbria.ac.uk/id/eprint/48846/

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: http://nrl.northumbria.ac.uk/policies.html

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)

Northumbria University NEWCASTLE

University Library

High-resolution snow depth prediction using Random Forest algorithm with topographic parameters: a case study in the Greiner Watershed, Nunavut

Julien Meloche[1,2], Alexandre Langlois[1,2], Nick Rutter[3], Donald McLennan[4], Alain Royer[1,2], Paul Billecocq[1,2], Serguei Ponomarenko[5]

[1]Département de géomatique appliquée, Université de Sherbrooke, Sherbrooke, J1K 2R1, Canada
[2]Centre d'études nordiques, Québec, Canada
[3]Department of Geography and Environmental Sciences, Northumbria University, Newcastle upon Tyne, NE1 8QH, UK
[4]Arctic Research Foundation, Ontario, N2L 6C2, Canada
[5]Polar Knowledge Canada, Canadian High Arctic Research Station, 1 Uvajuq Road, Cambridge Bay, NU, X0BoCo, Canada

Correspondence to: Julien Meloche (julien.meloche@usherbrooke.ca)

Abstract.

Increased surface temperatures (0.7°C per decade) in the Arctic affects polar ecosystems by reducing the extent and duration of annual snow cover. Monitoring of these important ecosystems needs detailed information on snow cover properties at resolutions (< 100 m) that influence ecological habitats and permafrost thaw. A machine learning method using topographic parameters with the Random Forest (RF) algorithm previously developed in alpine environments was applied over an arctic landscape for the first time. The topographic parameters used in the RF algorithm were Topographic Position Index (TPI) and up-wind slope index (Sx), which were estimated from the freely available Arctic DEM at 2 m resolution. Addition of an ecotype parameter (proxy for vegetation height) showed minimal predictive improvement. Using RF, snow depth distributions were predicted from topographic parameters with a root mean square error = 8 cm (23%) (R^2 = 0.79) at 10 m resolution for an arctic watershed (1 500 km^2) in western Nunavut, Canada.

1 Introduction

Arctic environments are experiencing warmer air temperatures (Notz and Stroeve, 2016; Richter-Menge *et al.*, 2017) where the rate of warming over the last 20 years is 2-3 times higher than elsewhere on the planet (Serreze and Barry, 2011; Cohen *et al.*, 2020). This accelerated warming, also referred as Arctic amplification, affects surface state variables such as snow and permafrost (Serreze and Barry, 2011; Bokhorst *et al.*, 2016; Derksen *et al.*, 2019; Meredith *et al.*, 2019) which in turn modifies soil properties, microbiological activities, and moisture availability for plant communities, all potentially driving shrub expansion (Marsh *et al.*, 2010; Domine *et al.*, 2016). The arctic shrub expansion in turn induces changes in snowpack properties (height, density and microstructure) with implications for ground-thermal regime (Myers-Smith *et al.*, 2020; Royer *et al.*, 2021). Changing snow conditions will also influence foraging conditions for arctic ungulate species (Bilodeau *et al.*, 2013; Dolant *et al.*, 2018; Gautier *et al.*, 2021) given the observed densification of the snowpack arising from an increased occurrence of winter extreme events such as rain-on-snow and winter storms (Dolant *et al.*, 2016; Langlois *et al.*, 2017).

The influence of snow cover on tundra ecosystems is strong and influences soil moisture, habitat suitability (McLennan *et al.*, 2018) and protection from desiccating winter winds and freezing temperatures (Sturm *et al.*, 2001). Furthermore, the low thermal conductivity of snow insulates ground surface temperature to modify freeze-thaw cycles, deepen the active layer and alter soil moisture and growing conditions for vegetation (Myers-Smith *et al.*, 2011). To better quantify and predict the upcoming changes governed by all these processes, detailed and spatially explicit information on snow conditions is needed (Ponomarenko *et al.*, 2019; Levasseur *et al.*, 2021)and yet high resolution products for snow are not widely available for the Arctic. Snow cover properties are influenced by vegetation, soil moisture, temperature gradient and strong wind (Essery and Pomeroy, 2004; Davesne *et al.*, 2021). In turn, these snow properties (density, microstructure and thermal conductivity) govern heat exchange between the atmosphere and the soil (Royer *et al.*, 2021). Consequently, snow cover is crucial in our understanding of local processes in permafrost, hydrology and ecosystem modelling.

Previous studies have measured snow depth distribution in the Arctic (Liston and Sturm, 1998; Sturm *et al.*, 2008; Sturm and Wagner, 2010) and improved our empirical understanding of the processes governing the spatial variability of snow properties. Two dominant snow processes controlling snow depth in arctic regions are wind redistribution (Parr *et al.*, 2020) and vegetation trapping (Sturm *et al.*, 2001; Domine *et al.*, 2016; Busseau *et al.*, 2017). Vegetation increases ground surface roughness and reduces wind speed to facilitate accumulation of snow, while topography and wind create accumulation and erosion zones depending on wind direction and speed, slope and micro-topography (Winstral *et al.*, 2013). Although these processes are conceptually well understood, simulation of snow redistribution at meter scales remains challenging. Statistical modelling of snow depth using linear regression with terrain parameters as predictors has been developed in alpine regions (Grünewald *et al.*, 2013; Pulwicki *et al.*, 2018) and proven to be a simple and effective method estimating snow depth. However, the high degree of correlation between topographic parameters restricts the prediction capabilities when using a linear regression approach. Revuelto et al. (2020) used machine learning to predict snow depth at 1m resolution in an alpine area using the Random Forest algorithm, with terrain parameters as predictors. The method showed good performance on model validation (R^2= 0.82-0.94). Random Forest works well with nonlinear relationships and interaction among predictors (Kibtia et al., 2020).

The main objective of this paper is to predict spatially distributed snow depth at high resolution (10 m) by testing the RF methodology of Revuelto et al. (2020) for arctic environments using the Arctic DEM available for the entire pan arctic. Since vegetation plays a key role in arctic snow redistribution, ecosystem type was evaluated as a feature using an ecotype map (Ponomarenko et al., 2019) to incorporate a proxy for vegetation height (snow trapping). First, snow-vegetation interactions were evaluated by comparing snow depth measurements from 2015-2019 within ecotypes. Second, the influence of topography and vegetation governing snow depth distribution in the Arctic was evaluated. Finally, snow depth predictions using RF at 10 m resolution were evaluated.

2. Method

2.1. Study site and ecosystems

The Greiner Lake watershed (69°13' N, 104°53' W) is located on eastern Victoria Island in western Nunavut, Canada, and is adjacent to the community of Cambridge Bay (CB). The calcareous, low relief landscape is characterized by dwarf shrub tundra on well drained upland areas, with productive sedge fens in depressions. A predominately northwest wind direction was measured by meteorological station data from Environment and Climate Change Canada (Station ID: CB-53512) at the Cambridge Bay airport located about 10 km to the south. A high resolution (10 m) ecosystem map of the study site, shown in Figure 1 (Ponomarenko et al., 2019), contains 17 ecosite types, which were identified using an unsupervised classification of Sentinel-2 multi-spectral imagery to delineate ecosystem units. Ecosystem units were verified and classified based on a standardized ecosystem classification unit (McLennan et al., 2018). The ecosystem classification, based on the Canadian Arctic-Subarctic Biogeoclimatic Ecosystem Classification (CASBEC), assesses plant community, soil moisture and nutrient regimes using a standardized approach organized by soil moisture and nutrient characteristics, and incorporates key abiotic processes of each ecosystem unit (McLennan et al., 2018). Soil moisture regimes used in Figure 2 were associated to ecotypes based on CASBEC. The map from Ponomarenko et al. (2019) was used to incorporate ecotype information as a proxy for vegetation height. The area is characterized by tussock sedge, dwarf-shrub, moss and graminoid tundra (Figure 1), based on the Circumpolar Arctic Vegetation Map (CAVM) project (Raynolds *et al.*, 2019). The snowpack corresponds to arctic tundra type (Royer et al., 2021) with an average depth between 30-40 cm, a basal layer of depth hoar and wind slab on top. The study site well represents arctic tundra processes, some areas of thicker vegetation for sub-arctic processes and some drier areas for polar desert.

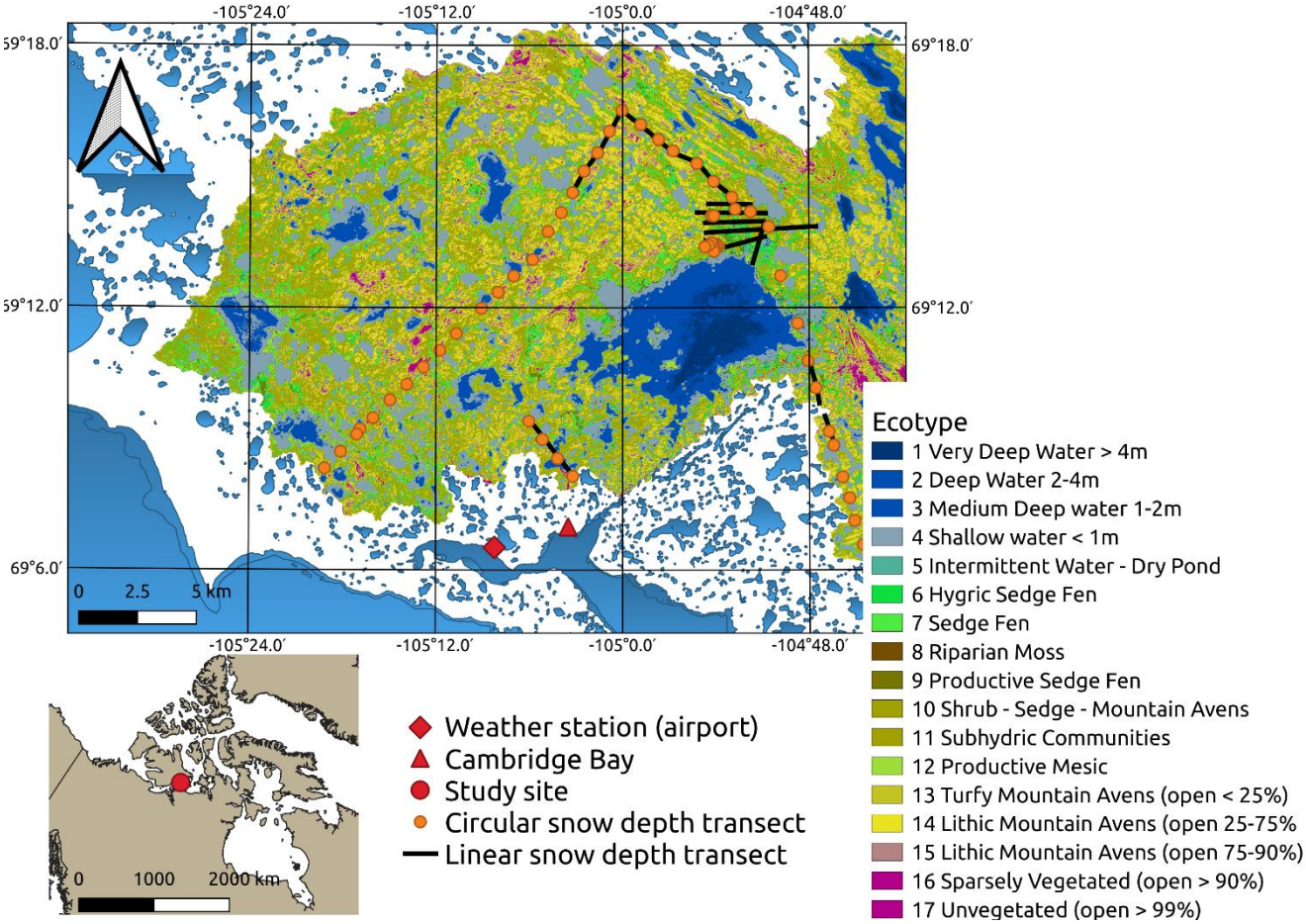

Figure 1: Study site of the Greiner watershed, Victoria Island, Nunavut with ecomap from Ponomarenko et al. (2019). Snow depth data (circular and linear transect) selected randomly across the entire dataset for training and validation. Grid of WGS84 lat/long is shown.

2.2. Data

2.2.1. Snow depth measurements

Geolocated snow depth measurements (Table 1) were made using a magnaprobe (Sturm and Holmgren, 2018), which is equipped with a standard GPS unit (accuracy 2.5 m). The snow depth accuracy is assumed to be 5 cm for this study because of soft vegetation; it can vary from 0 to > 5 cm depending on the substrate (Sturm and Holmgren, 2018). The device was used to conduct spatial surveys from linear and circular

transects (Figure 1), made near maximum annual snow accumulation. Circular transects of at least 30 snow depth measurements (approximately 30 - 50 m in diameter) provide spatially representative values. Linear snow depth transects between circular transects were added when possible. All ecotypes contained at least n > 225, and the number of samples per ecotype was a representation of the spatial extent in the field. Most of the dominant ecotypes (Mountain Avens, Sedge fen and Shrub) had n ≈ 2000 - 3000 samples. Each year, a different region of the Greiner watershed was measured to build up an overall assessment of spatial variability of snow depth because the interannual spatial pattern within a tundra watershed is consistent (Sturm and Wagner, 2010; Ménard *et al.*, 2014; Marsh *et al.*, 2020)

Table 1: Summary of magnaprobe snow depth survey near maximum annual snow accumulation used for training the RF algorithm. The mean (μ_{sd}) and standard deviation (σ_{sd}) of snow depth (m) from each year are presented. Temperature of accumulation period is from October to January and ablation period from May to July. NA indicates that at least one month of temperature data was unavailable.

					Accumulation period temperature (°C)			Ablation period temperature (°C)		
Year	**Date of measurement**	**Points**	μ_{sd}	σ_{sd}	**max**	**min**	**mean**	**max**	**min**	**mean**
2019	May 8	981	0.34	0.18	-7.5	-35.5	-23.1	11.7	-12.3	1
2018	04-May	577	0.27	0.16	-7.8	-30.9	-21.3	NA	NA	NA
2017	May 1-14,	3407	0.41	0.15	NA	NA	NA	13.2	-9.9	2.6
2016	April 2-10,	2916	0.34	0.18	-7.1	-31.7	-21.3	14.4	-9.6	3
2015	April 12-24,	9602	0.42	0.17	-6.4	-34	-22.1	12.7	-11.2	1.9
Total		**17 483**	**0.35**	**0.17**						

2.2.2 Random Forest (RF) algorithm

The Random Forest (RF) regressor algorithm builds a collection of decision trees and outputs the arithmetic average of each tree's prediction. Each tree is trained on a random subsample of the overall training set and with a random subsample of the predictors drawn among: Sx, TPI, ecotype, yearly mean snow depth and slope angle. As a result, RF can be trained with correlated variables, even though correlation can lower the importance of one of the predictors (Strobl *et al.*, 2008). The data set was first divided into 50% training and 50% validation to evaluate RF performance, and then evaluated over the

range of 10% to 90% of the data dedicated to training the model. The algorithm is trained with snow depth and corresponding predictors (continuous or discrete values) to make prediction based on weighted decision trees ensemble. Each snow depth measurement was associated with corresponding topographic parameters and ecotype using a spatial buffer of 10 m. For the final raster-based prediction, each raster of predictor had the same spatial extent and pixel size. Snow depth prediction was done by applying the RF algorithm pixel by pixel across the area. The spatial predictors and snow depth output can be represented as a data cube and a single snow depth prediction with all predictors within a column. RF was chosen over Support Vector Machine (SVM) and Neural Network (NN) because it performs well with correlated variables and automatically control nonlinear relationship and interaction among predictor (Kibtia et al., 2020).

RF expresses how predictions are made based on feature importance, which ranges from 0 to 1 (Strobl et al., 2007). There are two main algorithms for computing feature importance: Gini importance and permutation importance. Gini importance estimates feature importance based on the impurity reduction and the Gini gain splitting criterion (Strobl *et al.*, 2007). The permutation method randomly permutates a feature (predictor) which has the same overall effect of removing a feature to evaluate the performance without the feature (Strobl *et al.*, 2008). Both methods have a reported bias with correlated variables. Two variables will split their importance with each other so that one of the variables will have a lower feature importance. The Gini importance performs better with correlated variables but has issues with variation in number of categories of the feature (Revuelto et al., 2020; Strobl et al., 2008, 2007). Since we have both categorical variables and correlated variables, the permutation method was used in this paper because it is the most advanced way of estimating the importance (Strobl *et al.*, 2007). However, we noticed little difference between both methods with our data.

2.2.3. Predictors for snow depth

To obtain predictors (features) for the RF algorithm, topographic variables were derived from the ArcticDEM, which is a digital surface model with a 2 m-resolution (Porter *et al.*, 2021) created from stereo high-resolution optical imagery over the entire arctic domain. Using ArcticDEM, standard parameters such as elevation and slope were extracted for the feature analysis. Additional parameters were also calculated and include: 1) the Topographic Position Index (TPI) (Weiss, 2000) to evaluate pixel elevation with respect to surrounding pixels within a defined radius, and 2) the up-wind slope index (Sx) (Winstral *et al.*, 2002) to estimate the predisposition of a pixel to erosion or accumulation of snow depending on wind direction and slope. TPI accounts for wind redistribution over flat areas, and Sx for steeper areas (alpine regime). A TPI using a 150 m radius was used to analyze local effects, and a second TPI at 5 km was used to compensate for large elevational differences, and to act as a normalized elevation across the watershed. S_x was used with a radius of 100 m and a direction of 315° (dominant wind). In their study, Revuelto et al., (2020) used the snow depth measured at a local weather station to account for temporal variation of the mean, thereby linking mean snow depths at the point scale to the spatial distribution of snow depth across an alpine area. We applied a similar idea by adding the mean of the corresponding survey (Table 1) as a feature to each point i.e., all points collected in 2015 labeled with the fixed mean of 2015, repeated yearly. To account for vegetation interaction in RF, ecotypes from Figure 1 were used as another feature.

3. Results

3.1. Ecotypes and snow depth

In Figure 2, the average snow depth for each ecotype was correlated with the soil moisture regime from the edatopic grid presented in Figure 3 of McLennan et al. (2018) comparing snow cover and soil moisture regimes used in the CASBEC classification. Vegetated ecotypes with taller vegetation (30 cm height - Sedge Fen and Shrub-Sedge-Mountain avens) were correlated with higher snow depth averages given the trapping effect (Domine *et al.*, 2016) and which may also reduce sublimation (Liston *et al.*, 2002). Hygric

and hydric soil moisture regimes, associated with higher vegetation and moisture had deeper snow depths than the mean.

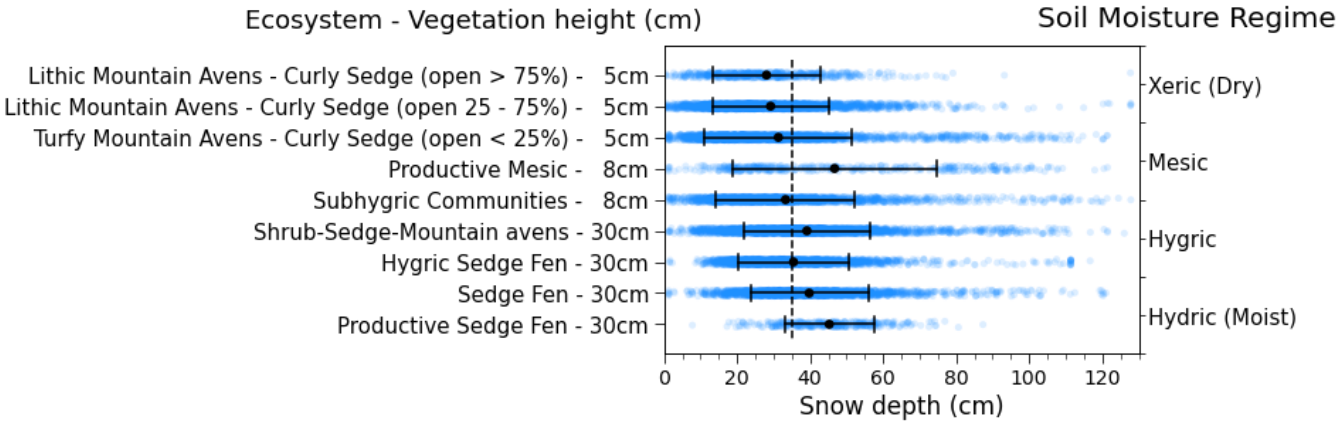

Figure 2: Mean and standard deviation of snow depth with principal ecotypes (Ponomarenko et al. 2019), vegetation height and soil moisture regime. The overall average (across all classes) snow depth is shown by the black dotted line. Vegetation height was estimated based on the type of vegetation in the ecotype. Soil moisture regime is based on the CASBEC classification (McLennan et al. 2018).

Mean snow depth in each ecotype was evaluated using a one-way Analysis of Variance (ANOVA) test which estimated that at least one of the ecotypes mean snow depth is significantly different than others (p-value < 0.05). Furthermore, a pairwise Tukey test highlighted which ecotypes were different from one another where ecotypes with similar means caused rejection of the null hypothesis. The null hypothesis states that snow depth means are significantly different. Results suggest that most ecotypes pairs (32 of 36) had significantly different means and both statistical tests indicated correlations between ecotypes and snow depth. The wide range of observations for each ecotype are probably the result of misclassification errors where the 10m resolution of the mapping will often incorporate more than one ecotype.

3.2. RF algorithm evaluation

The RF algorithm predicted snow depth with an RMSE of 8 cm (23%) and an absolute error of 6 cm (Figure 3). Both the explained variance and R^2 score were 0.79 yielding high accuracies, similar to what was found by Revuelto et al. (2020). Figure 3(b) suggests that the model performs well over a wide range of Arctic snow depths given that residuals are uniformly scattered over the depth range of 0 – 0.8 m. Figure 3(c) shows that good precision ($R^2 > 0.7$) and accuracy RMSE < 10 cm are reached even with a low percentage (10%) training sample.

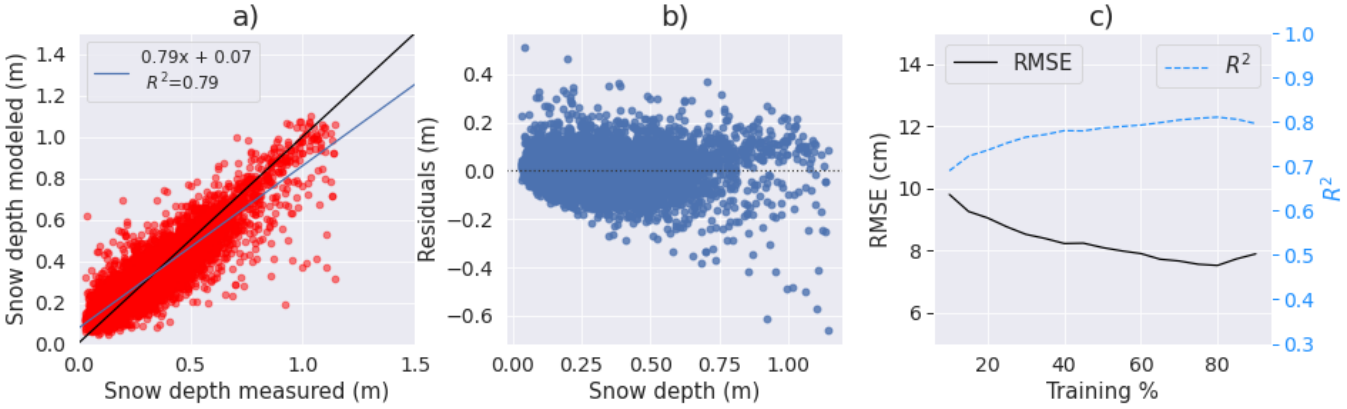

Figure 3: a) Measured and modelled snow depth from RF predictions. b) Residuals with measured values from the RF prediction.

Using the RF algorithm with snow depths and features defined above, variables derived from permutation importance for RF prediction can be estimated. The most important features for snow depth prediction (Figure 4 a) were wind-topographic parameters (TPI, Sx). As the Greiner watershed has a relatively low relief environment, TPI (150m) was the most significant variable for prediction. The annual mean snow depth had a high importance (0.3) since precipitation amount differed from year to year.

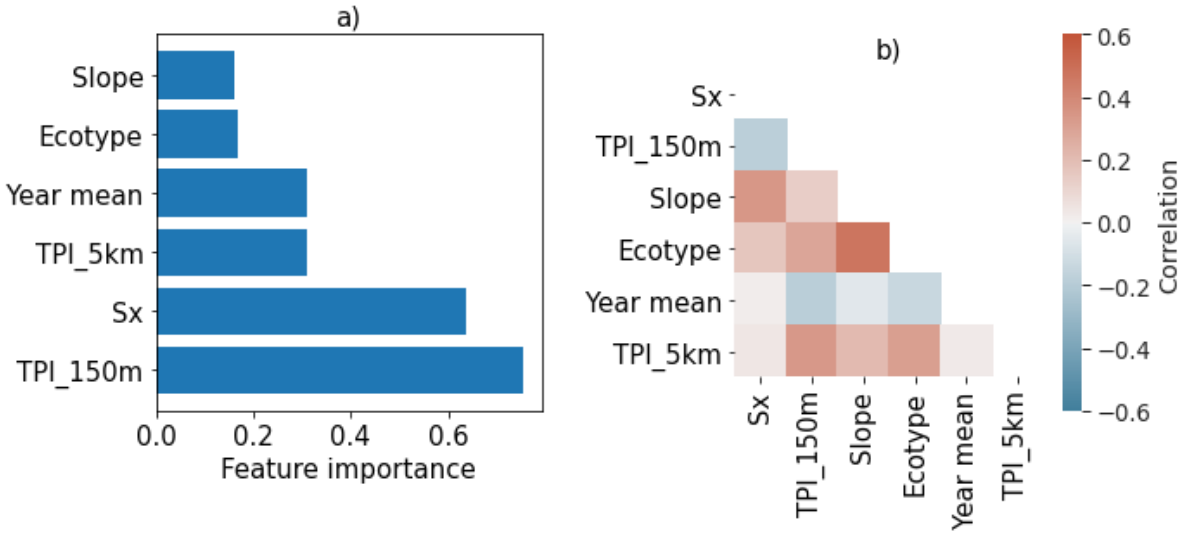

Figure 4: a) Feature importance in RF snow depth prediction and b) Correlation matrix between features calculated with the Pearson coefficient. "Year mean" corresponds to the mean winter snow depth.

Both TPI (150m and 5km) variables have an important positive correlation (0.3) with each other (Figure 4 b). Also, slope and Sx also had an expected positive correlation (0.3) given that Sx is a directional up-wind slope index. Even if RF performs well with correlated variables, two correlated variables will split their importance with each other so that one of the variables will have a lower feature importance. Consequently, ecotype is well correlated with all topographic parameters (section 3.1 and > 0.3, Figure 4 b) and explains why a low feature importance was found. Annual mean had little correlation (≈ 0, Figure 4 b) with other parameters.

Ecotypes were removed from the predicting variables, which accounted for a reduction of 0.01 in R^2 and 0.2 cm in RMSE. Ecotypes were found to have a small weight in feature importance of the RF algorithm compared to topographic parameters because they were correlated with each other (Figure 4). Topographic information (Sx and TPI) from the Arctic DEM can then predict snow depth alone. Since TPI and ecotype were highly correlated, the influence of ecotype on prediction was minimal. In Figure 5, predictions were tested by removing both TPI parameters. The performance was reduced to $R^2 = 0.65$ and

RMSE = 10 cm but the ecotype variable becomes much more important in the prediction (from 0.1 to 0.7).

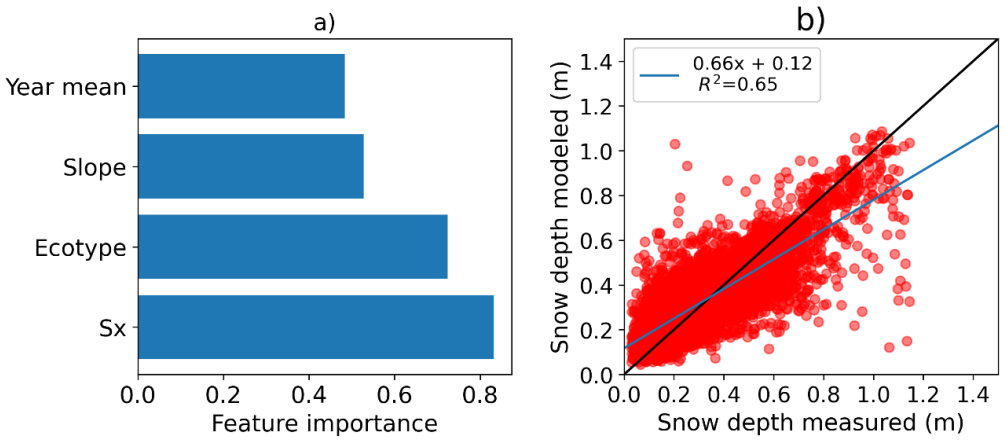

Figure 5: Feature importance of prediction without the variable TPI. The RMSE of snow depth is 10 cm.

3.3. Spatial prediction of snow depth across Greiner Watershed

Snow depth was predicted for the entire watershed. Water pixels were masked out since the RF algorithm was trained on land only using topographic variables and ecotypes. The predicted snow depth raster was generated pixel by pixel using the RF algorithm (no interpolation applied to final raster). Rasters of each predictor variable with the same spatial extent and resolution (10 m) were needed so the output can be predicted for each pixel. The influence of the directional Sx parameter at 315° (northwest wind) is clearly visible on the southeast aspect (Figure 6), where a snow depth > 0.75 m was predicted. Snow depth patterns appear highly dependent on local topography measured by TPI at 150 m according to local valleys and ridges.

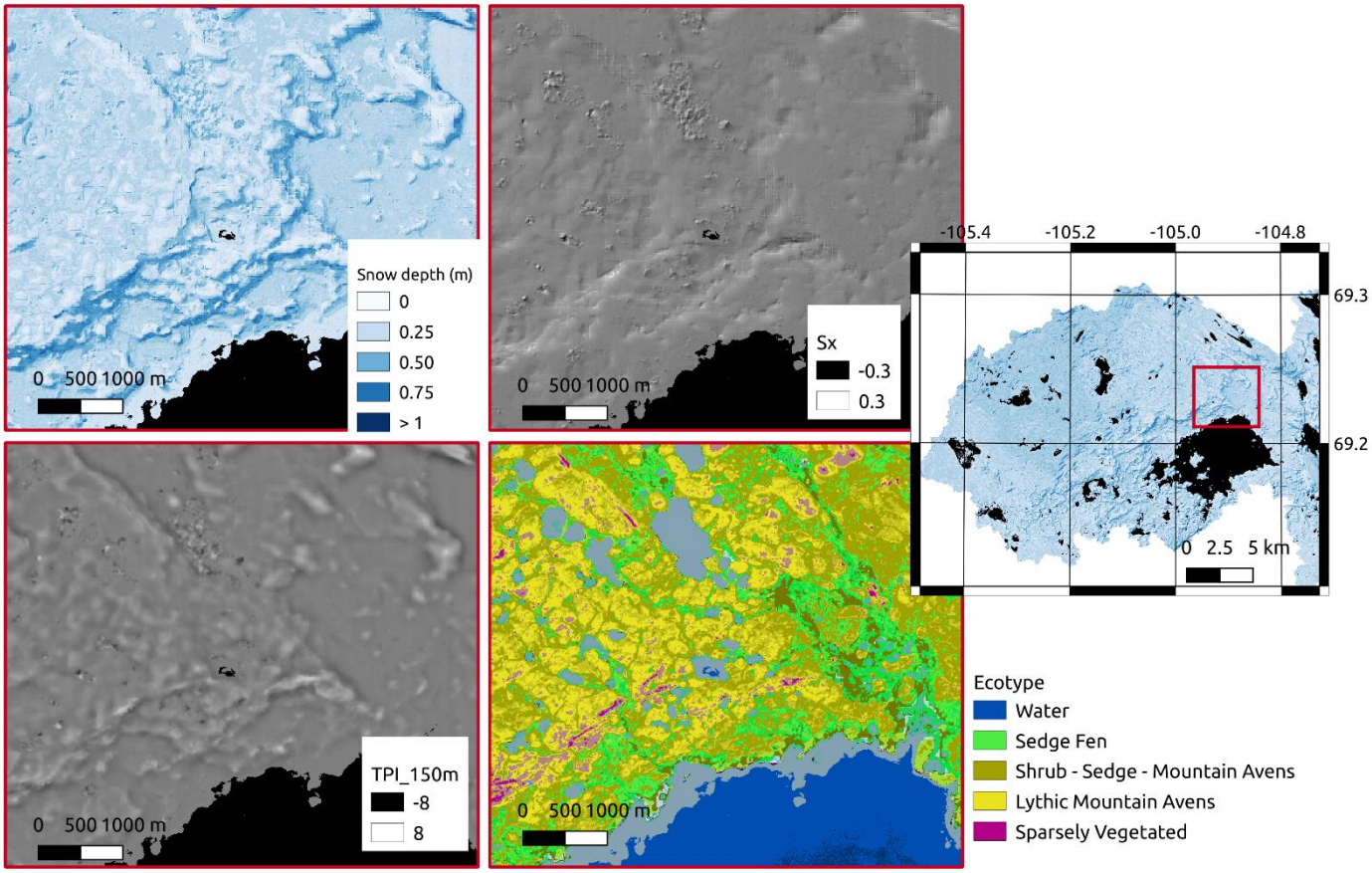

Figure 6: Snow depth prediction using the main variables of the RF algorithm (Sx, TPI and ecotype). Only a small subset of ecotypes is shown in the legend although all ecotypes (see Figure 1) are displayed on the map and were used in the prediction.

4. Discussion and Conclusion

This study demonstrates the applicability of RF to provide snow predictions using a high-resolution ecosystem map and topographic information. The RF prediction had high accuracy (RMSE = 8 cm, R^2 = 0.79) and was useful to learn more about tundra snow distribution processes. The algorithm showed that topographic parameters such as TPI at 150 m and Sx (R =100 m and Wind = N-W) were the dominant features in the prediction, which was also suggested by Revuelto et al. (2020). The mean snow depth per

year was used to normalize each year and allow the algorithm to be used for future predictions in the watershed with a mean depth fixed by a weather station or a physical snow model like SNOWPACK (Bartelt and Lehning, 2002). Spatial variation in feature importance across the watershed was not evaluated but we expect features to be uniform since processes for tundra environment will still be mostly represented by TPI (Wainwright *et al.*, 2017; Davesne *et al.*, 2021) and the Sx parameter with a potential small change in wind direction or search radius. Due to a lack of data, the algorithm could not be evaluated on an independent dataset in the area. The trained algorithm presented in this article should not be used outside the Greiner watershed since it is sensitive to trained data and cannot be used outside the range of response variables (Revuelto *et al.*, 2020). Deep NN could help make the prediction more general but would requires an intensive training dataset. Future work could involve training the algorithm on multiple Arctic sites by incorporating community-based science of snow depth measurements as well as large scale snow depth maps from RPAS or airborne observations (Nolan *et al.*, 2015; Walker *et al.*, 2020).

The strong correlation with small scale topography (TPI) (10 – 1000 m) and snow depth in flat areas of the Arctic was known (Wainwright *et al.*, 2017; Davesne *et al.*, 2021). This parameter located landscape depressions favorable for snow accumulation without any wind direction which is particularly relevant for flat areas where only wind speed and surface roughness affect transported snow (Liston and Sturm, 1998). The Sx parameter also had a high importance, this parameter was designed for sloping areas (Winstral *et al.*, 2002) where wind direction creates favorable conditions for up wind transport depending on slope and aspect. Snow accumulation on south-eastern aspects was common in the modeled area confirming the use of Sx at 315°, due to the usual northwest wind from winter storms. Topographic variables (TPI_{150m}, TPI_{5km} and slope) used in RF were correlated (> 0.3) with ecotypes (Figure 4 b) and soil moisture since more productive ecotypes adapted to higher moisture regimes were located in valleys, ephemeral ponds and shallow water. Drier ecosystems were mostly located on ridges (Ponomarenko et al., 2019).

It was found (Figure 2) that the soil moisture regime was linked to snow depth because depth was significantly deeper for ecotypes in hygric and hydric soil moisture regimes. This relationship has

hydrological implications; snow melt and water will accumulate at the surface because soil with high moisture content reduces infiltration due to frozen ice in the ground (Granger *et al.*, 1984). This relates to the basal layer of the arctic snowpack (depth hoar) since soil moisture has been shown to promote depth hoar formation (Davesne et al., 2021).

Additionally, based on a one-way ANOVA and Tukey pairwise test, most of the ecotypes showed significantly different mean snow depths indicating that ecotypes were correlated with snow depth. Vegetation trapping from ecotypes with taller vegetation was reflected in deeper snow depths (Figure 2), similar to other arctic regions (Sturm *et al.*, 2001; Domine *et al.*, 2016; Royer *et al.*, 2021). The importance of ecotype was reduced due to the correlation between vegetation and topography. Removing TPI parameters made ecotype relevant in the prediction. Vegetation buried by the weight of snow (Ménard *et al.*, 2014) will reduce the capacity for redistributed snow to accumulate in these areas and may additionally reduce the impact of vegetation on snow redistribution from early winter. It is important to distinguish the snow trapping effects of taller vegetation (Sturm *et al.*, 2001; Marsh *et al.*, 2010) from site soil moisture regime that is linked to the topographic parameter at a local scale. Water provided by spring snowmelt can be the main runoff event of the year (Marsh and Pomeroy, 1996; Pohl and Marsh, 2006).

The main result from this study is that topographic parameters, Sx and TPI, have a strong correlation with snow depth, which can be applied to other models or studies where snow depth is needed. One example could be predicting the coefficient of variation of snow depth for climate modelling (Nitta *et al.*, 2014) or microwave remote sensing of SWE (Meloche *et al.*, 2022). One disadvantage of a RF model such as presented in this study is that training data can be restrictive and rely on empirical relations. Machine learning models will never replace generalist-physical models, but they can effectively map the computation in parameter space (topography) of certain effects (in this case snow accumulation) and could be used to replicate patterns at a certain scale. One possible avenue to make the RF model more general is to use remote sensing products, such as Sentinel-1 in order to incorporate C band data at 10 m resolution as another feature in the RF modelling. Microwave active sensors with similar resolution (C or X band) are sensitive to vegetation (Duguay *et al.*, 2015) and snow depth (Lievens *et al.*, 2022).

Snow depth was predicted at 10 m resolution for the entire Greiner watershed using the freely available Arctic DEM. This method fulfilled part of the need for high resolution snow depth information and can be repeated easily with magnaprobe data or potentially Remotely Piloted Aircraft System (RPAS) snow depth cartography. The snow depth prediction in section 3.3 can be useful for other applications such as permafrost modelling in Cambridge Bay, which requires the high resolution (less than 100m) spatial information on snow depth that this map can provide. The algorithm can be trained and applied to other areas that require snow depth input at high resolution. SWE products at these larger scales are already benefiting from machine learning approaches (Xue and Forman, 2015; Bair *et al.*, 2018; Xiao *et al.*, 2018) to predict snow over the same area, and this method can do the same using the Arctic DEM anywhere in the Arctic. The RF method provides important insight on snow processes and snow-ecosystem interactions (trapping and metamorphism) and showed useful results that can provide the snow information necessary to conduct studies in related fields in the Arctic like permafrost, hydrology, or ecosystem modelling.

Acknowledgements

The authors would like to thank the entire Groupe de Recherche Interdisciplinaire en milieu Polaire research team from Université de Sherbrooke and CHARS staff team for field work assistance from 2015 to 2019 in Cambridge Bay.

Financial support

This research was made possible thanks to the financial support of the Natural Sciences and Engineering Research Council of Canada (NSERC), Polar Knowledge Canada, the Canadian Foundation for Innovation (CFI), Environment and Climate Change Canada (ECCC), Fonds de recherche du Québec – Nature et technologies (FRQNT), Northern Scientific Training Program (NSTP) and research funding from Northumbria University, UK.

Data availability

The snow depth map presented in section 3.3 available in raster (.tiff) format at 10 m resolution will be available: https://grimp.ca/data/

Code (Jupyter Notebook) and training data are also available:

https://github.com/JulienMeloche/RF_snowdepth

References

Bair EH, Abreu Calfa A, Rittger K, Dozier J. 2018. Using machine learning for real-time estimates of snow water equivalent in the watersheds of Afghanistan. *The Cryosphere* **12** (5): 1579–1594 DOI: 10.5194/tc-12-1579-2018

Bartelt P, Lehning M. 2002. A physical SNOWPACK model for the Swiss avalanche warning. *Cold Regions Science and Technology* **35** (3): 123–145 DOI: 10.1016/S0165-232X(02)00074-5

Bilodeau F, Gauthier G, Berteaux D. 2013. Effect of snow cover on the vulnerability of lemmings to mammalian predators in the Canadian Arctic. *Journal of Mammalogy* **94** (4): 813–819 DOI: 10.1644/12-MAMM-A-260.1

Bokhorst S, Pedersen SH, Brucker L, Anisimov O, Bjerke JW, Brown RD, Ehrich D, Essery RLH, Heilig A, Ingvander S, et al. 2016. Changing Arctic snow cover: A review of recent developments and assessment of future needs for observations, modelling, and impacts. *Ambio* **45** (5): 516–537 DOI: 10.1007/s13280-016-0770-0

Busseau BC, Royer A, Roy A, Langlois A, Domine F. 2017. Analysis of snow-vegetation interactions in the low Arctic-Subarctic transition zone (northeastern Canada). *Physical Geography* **38** (2): 159–175 DOI: 10.1080/02723646.2017.1283477

Cohen J, Zhang X, Francis J, Jung T, Kwok R, Overland J, Ballinger TJ, Bhatt US, Chen HW, Coumou D, et al. 2020. Divergent consensuses on Arctic amplification influence on midlatitude severe winter weather. *Nature Climate Change* **10** (1): 20–29 DOI: 10.1038/s41558-019-0662-y

Davesne G, Domine F, Fortier D. 2021. Effects of meteorology and soil moisture on the spatio-temporal evolution of the depth hoar layer in the polar desert snowpack. *Journal of Glaciology*: 1–16 DOI: 10.1017/jog.2021.105

Derksen C, Burgess D, Duguay C, Howell S, Mudryk L, Smith S, Thackeray C, Kirchmeier-Young M. 2019. Changes in Snow, Ice, and Permafrost Across Canada. *Canada's Changing Climate Report*: 194–260 Available at: https://changingclimate.ca/site/assets/uploads/sites/2/2018/11/CCCR-Chapter5-ChangesInSnowIcePermafrostAcrossCanada.pdf

Dolant C, Langlois A, Montpetit B, Brucker L, Roy A, Royer A. 2016. Development of a rain-on-snow detection algorithm using passive microwave radiometry. *Hydrological Processes* **30** (18): 3184–3196 DOI: 10.1002/hyp.10828

Dolant C, Montpetit B, Langlois A, Brucker L, Zolina O, Johnson CA, Royer A, Smith P. 2018. Assessment of the Barren Ground Caribou Die-off During Winter 2015–2016 Using Passive Microwave Observations. *Geophysical Research Letters* **45** (10): 4908–4916 DOI: 10.1029/2017GL076752

Domine F, Barrere M, Morin S. 2016. The growth of shrubs on high Arctic tundra at Bylot Island: Impact on snow physical properties and permafrost thermal regime. *Biogeosciences* **13** (23): 6471–6486 DOI: 10.5194/bg-13-6471-2016

Duguay Y, Bernier M, Lévesque E, Tremblay B. 2015. Potential of C and X band SAR for shrub growth monitoring in sub-arctic environments. *Remote Sensing* **7** (7): 9410–9430 DOI: 10.3390/rs70709410

Essery R, Pomeroy J. 2004. Vegetation and topographic control of wind-blown snow distributions in distributed and aggregated simulations for an arctic tundra basin. *Journal of Hydrometeorology* **5** (5): 735–744 DOI: 10.1175/1525-7541(2004)005<0735:VATCOW>2.0.CO;2

Gautier C, Langlois A, Sasseville V, Neave E, Johnson CA. 2021. Evaluation of Peary caribou (Rangifer tarandus pearyi) habitat conditions and migratory patterns using remote sensing, snow modelling and traditional knowledge. *Polar Research (submitted)*

Granger RJ, Gray DM, Dyck GE. 1984. Snowmelt infiltration to frozen Prairie soils. *Canadian Journal of Earth Sciences* **21** (6): 669–677 DOI: 10.1139/e84-073

Grünewald T, Stötter J, Pomeroy JW, Dadic R, Moreno Baños I, Marturià J, Spross M, Hopkinson C, Burlando P, Lehning M. 2013. Statistical modelling of the snow depth distribution in open alpine terrain. *Hydrology and Earth System Sciences* **17** (8): 3005–3021 DOI: 10.5194/hess-17-3005-2013

Kibtia H, Abdullah S, Bustamam A. 2020. Comparison of random forest and support vector machine for prediction of cognitive impairment in Parkinson's disease. In *AIP Conference Proceedings*020093. DOI: 10.1063/5.0030332

Langlois A, Johnson C-A, Montpetit B, Royer A, Blukacz-Richards EA, Neave E, Dolant C, Roy A, Arhonditsis G, Kim D-K, et al. 2017. Detection of rain-on-snow (ROS) events and ice layer formation using passive microwave radiometry: A context for Peary caribou habitat in the Canadian Arctic. *Remote Sensing of Environment* **189**: 84–95 DOI: 10.1016/j.rse.2016.11.006

Levasseur S, Brown K, Langlois A, McLennan D. 2021. Measurement of Snow Physical Properties and Stable Isotope Variations in the Canadian Sub-Arctic and Arctic Snowpack. *Atmosphere-Ocean* **59** (3): 137–151 DOI: 10.1080/07055900.2021.1962240

Lievens H, Brangers I, Marshall H-P, Jonas T, Olefs M, de Lannoy G. 2022. Sentinel-1 snow depth retrieval at sub-kilometer resolution over the European Alps. *The Cryosphere* **16** (1): 159–177 DOI: 10.5194/tc-16-159-2022

Liston GE, Sturm M. 1998. A snow-transport model for complex terrain. *Journal of Glaciology* **44** (148): 498–516 DOI: 10.1017/S0022143000002021

Liston GE, Mcfadden JP, Sturm M, Pielke RA. 2002. Modelled changes in arctic tundra snow, energy and moisture fluxes due to increased shrubs. *Global Change Biology* **8** (1): 17–32 DOI: 10.1046/j.1354-1013.2001.00416.x

Marsh CB, Pomeroy JW, Spiteri RJ, Wheater HS. 2020. A Finite Volume Blowing Snow Model for Use With Variable Resolution Meshes. *Water Resources Research* **56** (2) DOI: 10.1029/2019WR025307

Marsh P, Pomeroy JW. 1996. Meltwater fluxes at an arctic forest-tundra site. *Hydrological Processes* **10** (10): 1383–1400 DOI: 10.1002/(sici)1099-1085(199610)10:10<1383::aid-hyp468>3.0.co;2-w

Marsh P, Bartlett P, MacKay M, Pohl S, Lantz T. 2010. Snowmelt energetics at a shrub tundra site in the western Canadian Arctic. *Hydrological Processes* **24** (25): 3603–3620 DOI: 10.1002/hyp.7786

McLennan DS, MacKenzie WH, Meidinger D, Wagner J, Arko C. 2018. A Standardized Ecosystem Classification for the Coordination and Design of Long-term Terrestrial Ecosystem Monitoring in Arctic-Subarctic Biomes. *Arctic* **71** (5): 1–15 DOI: 10.14430/arctic4621

Meloche J, Langlois A, Rutter N, Royer A, King J, Walker B, Marsh P, Wilcox EJ. 2022. Characterizing tundra snow sub-pixel variability to improve brightness temperature estimation in satellite SWE retrievals. *The Cryosphere* **16** (1): 87–101 DOI: 10.5194/tc-16-87-2022

Ménard CB, Essery R, Pomeroy J. 2014. Modelled sensitivity of the snow regime to topography, shrub fraction and shrub height. *Hydrology and Earth System Sciences* **18** (6): 2375–2392 DOI: 10.5194/hess-18-2375-2014

Meredith M, Sommerkorn M, Cassotta S, Derksen C, Ekaykin A, Hollowed A, Kofinas G, Mackintosh A, Melbourne-Thomas J, Muelbert MMC, et al. 2019. Polar regions. In H.-O Pörtner, et al. (Eds.)

Myers-Smith IH, Forbes BC, Wilmking M, Hallinger M, Lantz T, Blok D, Tape KD, Macias-Fauria M, Sass-Klaassen U, Lévesque E, et al. 2011. Shrub expansion in tundra ecosystems: dynamics, impacts and research priorities. *Environmental Research Letters* **6** (4): 045509 DOI: 10.1088/1748-9326/6/4/045509

Myers-Smith IH, Kerby JT, Phoenix GK, Bjerke JW, Epstein HE, Assmann JJ, John C, Andreu-Hayles L, Angers-Blondin S, Beck PSA, et al. 2020. Complexity revealed in the greening of the Arctic. *Nature Climate Change* **10** (2): 106–117 DOI: 10.1038/s41558-019-0688-1

Nitta T, Yoshimura K, Takata K, O'ishi R, Sueyoshi T, Kanae S, Oki T, Abe-Ouchi A, Liston GE. 2014. Representing variability in subgrid snow cover and snow depth in a global land model: Offline validation. *Journal of Climate* **27** (9): 3318–3330 DOI: 10.1175/JCLI-D-13-00310.1

Nolan M, Larsen C, Sturm M. 2015. Mapping snow depth from manned aircraft on landscape scales at centimeter resolution using structure-from-motion photogrammetry. *Cryosphere* **9** (4): 1445–1463 DOI: 10.5194/tc-9-1445-2015

Notz D, Stroeve J. 2016. Observed Arctic sea-ice loss directly follows anthropogenic CO2 emission. *Science* **354** (6313): 747–750 DOI: 10.1126/science.aag2345

Parr C, Sturm M, Larsen C. 2020. Snowdrift Landscape Patterns: An Arctic Investigation. *Water Resources Research* **56** (12) DOI: 10.1029/2020WR027823

Pohl S, Marsh P. 2006. Modelling the spatial-temporary variability of spring snowmelt in an arctic catchment. *Hydrological Processes* **20** (8): 1773–1792 DOI: 10.1002/hyp.5955

Ponomarenko S, McLennan D, Pouliot D, Wagner J. 2019. High Resolution Mapping of Tundra Ecosystems on Victoria Island, Nunavut – Application of a Standardized Terrestrial Ecosystem Classification. *Canadian Journal of Remote Sensing* **45** (5): 551–571 DOI: 10.1080/07038992.2019.1682980

Porter C, Morin P, Howat I, Noh M-J, Bates B, Peterman K, Keesey S, Schlenk M, Gardiner J, Tomko K, et al. 2021. ArcticDEM – Polar Geospatial Center. *Harvard Dataverse*: V1 DOI: 10.7910/DVN/OHHUKH

Pulwicki A, Flowers GE, Radic V, Bingham D. 2018. Estimating winter balance and its uncertainty from direct measurements of snow depth and density on alpine glaciers. *Journal of Glaciology* **64** (247): 781–795 DOI: 10.1017/jog.2018.68

Raynolds MK, Walker DA, Balser A, Bay C, Campbell M, Cherosov MM, Daniëls FJA, Eidesen PB, Ermokhina KA, Frost G v., et al. 2019. A raster version of the Circumpolar Arctic Vegetation Map (CAVM). *Remote Sensing of Environment* **232**: 111297 DOI: 10.1016/j.rse.2019.111297

Revuelto J, Billecocq P, Tuzet F, Cluzet B, Lamare M, Larue F, Dumont M. 2020. Random forests as a tool to understand the snow depth distribution and its evolution in mountain areas. *Hydrological Processes* **34** (26): 5384–5401 DOI: 10.1002/hyp.13951

Richter-Menge J, Overland JE, Mathis JT, Osborne E. 2017. Arctic Report Card 2017 Available at: http://www.arctic.noaa.gov/Report-Card

Royer A, Domine F, Roy A, Langlois A, Marchand N, Davesne G. 2021. New northern snowpack classification linked to vegetation cover on a latitudinal mega-transect across northeastern Canada. *Écoscience* **28** (3–4): 225–242 DOI: 10.1080/11956860.2021.1898775

Serreze MC, Barry RG. 2011. Processes and impacts of Arctic amplification: A research synthesis. *Global and Planetary Change* **77** (1–2): 85–96 DOI: 10.1016/j.gloplacha.2011.03.004

Strobl C, Boulesteix A-L, Kneib T, Augustin T, Zeileis A. 2008. Conditional variable importance for random forests. *BMC Bioinformatics* **9** (1): 307 DOI: 10.1186/1471-2105-9-307

Strobl C, Boulesteix AL, Zeileis A, Hothorn T. 2007. Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics* **8** DOI: 10.1186/1471-2105-8-25

Sturm M, Holmgren J. 2018. An Automatic Snow Depth Probe for Field Validation Campaigns. *Water Resources Research* **54** (11): 9695–9701 DOI: 10.1029/2018WR023559

Sturm M, Wagner AM. 2010. Using repeated patterns in snow distribution modeling: An Arctic example. *Water Resources Research* **46** (12): 1–15 DOI: 10.1029/2010WR009434

Sturm M, Derksen C, Liston G, Silis A, Solie D, Holmgren J, Huntington H. 2008. A Reconnaissance Snow Survey across Northwest Territories and Nunavut, Canada, April 2007 Available at: http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA476959

Sturm M, Holmgren J, McFadden JP, Liston GE, Chapin FS, Racine CH. 2001. Snow–Shrub Interactions in Arctic Tundra: A Hypothesis with Climatic Implications. *Journal of Climate* **14** (3): 336–344 DOI: 10.1175/1520-0442(2001)014<0336:SSIIAT>2.0.CO;2

Wainwright HM, Liljedahl AK, Dafflon B, Ulrich C, Peterson JE, Gusmeroli A, Hubbard SS. 2017. Mapping snow depth within a tundra ecosystem using multiscale observations and Bayesian methods. *The Cryosphere* **11** (2): 857–875 DOI: 10.5194/tc-11-857-2017

Walker B, Wilcox EJ, Marsh P. 2020. Accuracy assessment of late winter snow depth mapping for tundra environments using Structure-from-Motion photogrammetry. *Arctic Science* **17** (November): 1–17 DOI: 10.1139/as-2020-0006

Weiss AD. 2000. Topographic Position and Landforms analysis. *Poster presentation, ESRI User Conference, San Diego, CA* **64**: 227–245 Available at: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Topographic+Position+and+Landforms+Analysis#0

Winstral A, Elder K, Davis RE. 2002. Spatial Snow Modeling of Wind-Redistributed Snow Using Terrain-Based Parameters. *Journal of Hydrometeorology* **3** (5): 524–538 DOI: 10.1175/1525-7541(2002)003<0524:SSMOWR>2.0.CO;2

Winstral A, Marks D, Gurney R. 2013. Simulating wind-affected snow accumulations at catchment to basin scales. *Advances in Water Resources* **55**: 64–79 DOI: 10.1016/j.advwatres.2012.08.011

Xiao X, Zhang T, Zhong X, Shao W, Li X. 2018. Support vector regression snow-depth retrieval algorithm using passive microwave remote sensing data Remote Sensing of Environment Support vector regression snow-depth retrieval algorithm using passive microwave remote sensing data. *Remote Sensing of Environment* **210** (February 2019): 48–64 DOI: 10.1016/j.rse.2018.03.008

Xue Y, Forman BA. 2015. Comparison of passive microwave brightness temperature prediction sensitivities over snow-covered land in North America using machine learning algorithms and the Advanced Microwave Scanning Radiometer. *Remote Sensing of Environment* DOI: 10.1016/j.rse.2015.09.009