Machine learning analyses of remote sensing measurements establish strong relationships between vegetation and snow depth in the boreal forest of Interior Alaska

Thomas A Douglas and Caiyun Zhang
1 U.S. Army Cold Regions Research and Engineering Laboratory, Fort Wainwright, AK, United States of America
2 Department of Geosciences, Florida Atlantic University, Boca Raton, FL, United States of America

E-mail: thomas.a.douglas@usace.army.mil

Abstract

The seasonal snowpack plays a critical role in Arctic and boreal hydrologic and ecologic processes. Though snow depth can be markedly different from one season to another there are strong repeated relationships between ecotype and snowpack depth. In the diverse vegetative cover of the boreal forest of Interior Alaska, a warming climate has shortened the winter season. Alterations to the seasonal snowpack, which plays a critical role in regulating wintertime soil thermal conditions, have major ramifications for near-surface permafrost. Therefore, relationships between vegetation and snowpack depth are critical for identifying how present and projected future changes in winter season processes or land cover will affect permafrost. Vegetation and snow cover areal extent can be assessed rapidly over large spatial scales with remote sensing methods, however, measuring snow depth remotely has proven difficult. This makes snow depth–vegetation relationships a potential means of assessing snowpack characteristics. In this study, we combined airborne hyperspectral and LiDAR data with machine learning methods to characterize relationships between ecotype and the end of winter snowpack depth. More than 26 000 snow depth measurements were collected between 2014 and 2019 at three field sites representing common boreal ecoregion land cover types. Our results show hyperspectral measurements account for two thirds or more of the variance in the relationship between ecotype and snow depth. Of the three modeling approaches we used, support vector machine yields slightly stronger statistical correlations between snowpack depth and ecotype for most winters. An ensemble analysis of model outputs using hyperspectral and LiDAR measurements yields the strongest relationships between ecotype and snow depth. Our results can be applied across the boreal biome to model the coupling effects between vegetation and snowpack depth.

1. Introduction

Across the Arctic and sub-Arctic, the seasonal snowpack blankets the landscape for up to three quarters of the year. Snow cover plays roles in the soil thermal regime, water balance, land use decisions, winter recreation and mobility, and provides a critical ecological habitat (Kumpula and Colpaert 2007, Euskirchen et al 2010, Pozzanghera et al 2016, Olnes and Kielland 2017, Boelman et al 2019, Chapin et al 2000). Many areas are warming and the potential for permafrost thaw could alter ecosystem structure (Euskirchen et al 2007) and liberate large carbon stores that are currently frozen (Yi et al 2015). The snowpack plays a critical role in the soil thermal regime (Loranty et al 2018). The snowpack accumulates from fall through the following spring with few melt or runoff events. Spring snowmelt is typically the largest yearly hydrologic event and is associated with the transfer of a variety of nutrients from the snowpack and land to streams and eventually to the ocean (Judd and Kling 2002, Dittmar and Kattner 2003).
The taiga biome, which includes the northern boreal ecoregion, is the largest biome on Earth. Much of the boreal ecoregion in Russia, Canada, and the U.S. (Alaska) is underlain by permafrost. Local vegetative cover plays a major role as biogeophysical ‘ecosystem protection’ for this permafrost (Shur and Jorgenson 2007, Euskirchen et al 2010, 2016, Loranty et al 2014, Kropp et al 2020), largely due to snow-vegetation interactions (Ling and Zhang 2003, Sturm et al 2005, Jafarov et al 2014, 2018, Brown et al 2015). Changes in vegetation cover that lead to increased snow depths are a positive feedback for permafrost thaw, predominantly due to providing a thermal buffer against winter cold (Zhang 2005). As such, the timing of snow-on and snow-off, the snowpack depth, and the end of winter snow water equivalent (SWE) of the seasonal snowpack are major controls on the permafrost thermal regime (Sturm 1992, Zhang 2005, Jafarov et al 2018, Grünberg et al 2020).

Snow depth and SWE are variable from one winter to another due to weather variations but, more critically, snow depth can be markedly different across ecotypes under the same climatic conditions. This is especially true in boreal regions where land cover is a mosaic of dense tall-stature spruce forests, deciduous forests, shrublands, and low-lying herbaceous wetlands (Chapin et al 2000). Land cover spatial heterogeneity yields a corresponding variability in snow capture by canopies (Kozii et al 2017).

In Interior Alaska, three major climate change drivers are affecting vegetation and snowpacks. First, fire and post-fire succession can redistribute the fraction of vegetation cover types over time and this has favored the transition from conifer to mixed or deciduous forests (Johnstone et al 2010, Mekonnen et al 2019, Holloway et al 2020). Second, permafrost is undergoing top-down thaw, leading to changes in soil and vegetation composition (Jorgenson et al 2020, Douglas et al 2021) associated with conversion of some forests to wetlands (Jorgenson et al 2001, Lara et al 2016). Third, the number of days per year in which snow covers the ground has been decreasing and this is projected to continue (Lader et al 2017, 2020, Littell et al 2018, Mudryk et al 2020). These ongoing and projected future changes in the fraction of vegetative cover types and length of the winter season will have major ramifications for snow-vegetation interactions and, thus, the surface energy balance and soil thermal regimes of permafrost and the seasonally thawed ‘active’ surface layer. Despite the potentially critical role the snowpack plays in ecosystem protection of boreal permafrost, knowledge of how, where, and when permafrost thaw will result from these vegetation changes is limited due to sparse records of snowpack depth.

Optical hyperspectral sensors (Zhang et al 2020) and repeat light detection and ranging (LiDAR) measurements (Douglas et al 2016) have proven effective in mapping vegetation and permafrost characteristics. Snow cover areal extent can be measured with Moderate Resolution Imaging Spectroradiometer (MODIS; Bennett et al 2019) and passive microwave sensors (Semmens and Ramage 2013). However, remote sensing tools for directly estimating snowpack depth are limited. Snow depth has been estimated at some sites in Alaska by combining satellite microwave measurements with ground-based observations (Wang et al 2020). Increasingly, machine learning analyses have been used to estimate SWE in remote areas from limited field measurements (Bair et al 2018, Broxton et al 2019).

In this study, we used machine learning methods to establish relationships between ecotype and the Interior Alaska seasonal snowpack during high and low snowpack winter seasons. From 2014 to 2019, we made more than 26 000 snow depth measurements at three field sites. The sites encompass common boreal ecoregion land cover classes (evergreen and deciduous forest, mixed forest, shrubland, grassland, wetland, barren, built-up, and water), cumulatively representing 74% of the boreal and taiga (Latifovic et al 2017). The National Aeronautics and Space Administration (NASA)’s Arctic-Boreal Vulnerability Experiment (ABoVE; Fisher et al 2018) and our aerial campaigns collected hyperspectral imagery and LiDAR data at the sites. A rich dataset and the sophistication of machine learning techniques provide an opportunity to quantitatively analyze the coupling effects of vegetation and snowpack over permafrost terrains. The established relationships can be used to project how and where a warmer climate, changing snowpack characteristics, and an altered vegetation cover will affect the thermal regime of near-surface permafrost.

2. Materials and methods

2.1. Study sites and local climatology

We made our measurements at three long-term sites located in Interior Alaska near Fairbanks: the Farmer’s Loop permafrost experimental station (hereafter refer to as Farmer’s Loop), the Creamer’s Field Migratory Wildfowl Refuge (Creamer’s Field), and the Alaska Peatland Experiment (APEX; figure 1; Douglas et al 2019). Interior Alaska is characterized by a cold continental climate with dry winters. The mean annual air temperature is −3.3 °C and extremes range from 38 °C to −51 °C (Jorgenson et al 2001). Mean annual precipitation is 280 mm (Wendler and Shulski 2009). The mean annual snowfall of 1.7 m (Jorgenson et al 2001) represents 40%–45% of annual precipitation (Liston and Hiemstra 2011). The area is underlain by discontinuous permafrost up to 60 m thick that is primarily located along north-facing slopes, in lowlands, and where vegetative cover or soils offer thermal protection (Jorgenson et al 2008,
Douglas et al 2014). The Pleistocene permafrost at our field sites consists of syngenetic ice-rich wind-blown loess and organic matter representing high carbon ‘yedoma’ permafrost that is highly vulnerable to thaw following changes in the soil thermal regime that include an altered seasonal snowpack (Shur and Jorgenson 2007, Loranty et al 2018).

The Fairbanks area is characterized by a boreal forest snowpack (Sturm et al 1995, Sturm and Liston 1997) that covers the ground roughly October to May (Domine et al 2006). The historical observed snow onset date from the Global Historical Climatology Network daily database (GHCN-D) for Fairbanks is 16 October (Lader et al 2020). Numerous (20 or more) precipitation events of 1–5 cm of snow occur throughout the winter (Jacobi et al 2010). Strong temperature gradients from the ground to the atmosphere over winter form depth-hoar which is associated with an increase in grain sizes and air space and a decrease in the number of snow grains (Sturm and Benson 1997). The snowpack is typically ~40–100 cm deep. Since the Fairbanks area is not windy at any time of the year, redistribution of snow from major wind events is usually minor (Taillandier et al 2006).

In forested areas, ~40% of the snow is intercepted by vegetation where it can then be sublimated (Hedstrom and Pomeroy 1998). In unforested areas, such as wetlands and tussock tundra, minor wind redistribution occurs (Sturm and Benson 2004). When wind events do occur, they form wind slabs interspersed with layers of softer lightly sintered grains. This is most common in treeless open areas with greater fetch than in forested locations. Above freezing temperatures are uncommon between November and early March so melt layers and percolation columns are not typically formed until late March. Recent research projects an increase in rain on snow events in a warmer future (Bieniek et al 2018). The snow-ground interface remains frozen from late October onward and temperature gradients in the snowpack range from 50 °C m⁻¹ to 200 °C m⁻¹ (Domine et al 2006).

2.2. Snow and ecotype surveys
We made 26 902 end of season snowpack thickness measurements at a 1–2 m spacing between 2014 and 2019 at our field sites (figure 1) using a GPS-enabled...
magnaprobe’ (SnowHydro, Fairbanks Alaska). The magnaprobe consists of a 1.3 m long ~1 cm diameter rod with a moveable basket that uses a GPS and a magneto restrictive material inside the rod to record snow depth and location in seconds (Sturm and Holmgren 2018). GPS location accuracy is ±3 m in open areas but may decrease to 10–15 m in dense forest. We excavated snow pits at ~100 m intervals along our transects and snow pit depths and Magnaprobe depths were typically within 5–10 cm of one another which is in agreement with published results (Sturm and Holmgren 2018). Ecotype characteristics were recorded along the transects in the summer of 2015.

Our measurements were made at three different sites within 35 km of one another. At Farmer’s Loop, we made measurements along two 400 m transects starting at (64.877 N, 147.674 W) and (64.874 N, 147.677 W), respectively (figure 1(c)). The transects start in a mixed deciduous forest and then, sequentially, cross wetland, tussock tundra, and a moss-black spruce forest. Some grass-dominated trails are also crossed and are evident in the natural color image (figure 1). At nearby Creamer’s Field, measurements were conducted along a 500 m transect starting at (64.868° N, 147.738° W; figure 1(e)). This transect starts in a mixed deciduous forest that abruptly transitions into tussock tundra. Two trails are crossed. In the mixed forest large degrading ice wedge polygons roughly 5–10 m across are evident in summer and winter but are not visible in the natural color image shown in figure 1. An additional 1891 snow depths were measured along a separate 2 km transect in 2017 in support of NASA’s Operation IceBridge campaign. At APEX, measurements were made along three transects: transect 1 starting at (64.704 N, 148.325 W) and headed southeast for ~1.2 km; transect 2 starting at (64.697 N, 148.332 W) and headed southeast for ~1.5 km long; and transect 3 starting at (64.695 N, 148.334 W) following a 1.2 km tear-dropped shape (figure 1(d)). The northern two transects start in mixed forest and transition abruptly into moss-black spruce forest and then a large open fen wetland. The southernmost APEX transect is located predominantly in moss-black spruce forest, crosses a bog wetland, and returns into moss-black spruce forest (Anderson et al 2019, McPartland et al 2019). We also measured snow depth at 3517 locations along a separate 2 km transect at this site in 2017 as part NASA’s Operation IceBridge campaign. The spruce forest cover at APEX is highly variable. Toward the western portion of the transects there is a dense forest canopy. Further east, on either side of the open bogs and fens, the spruce trees are up to 5 m apart and the increase in the fraction of open exposed snowpack leads to lower snow interception rates and higher potential for wind redistribution.

### 2.3 Aerial campaigns

Airborne hyperspectral imagery was collected in July 2014 at Farmer’s Loop and APEX using the ProSpecTIR-VS hyperspectral imager (SpecTIR, LLC) that collects through 360 channels covering visible, near-infrared (400 nm–970 nm) with a spectral resolution of 2.9 nm and shortwave infrared (970 nm–2450 nm) with a spectral resolution of 8.5 nm. The hyperspectral data were radiometrically and geometrically corrected by the vendor and available for this study in 1 m reflectance products. LiDAR data were also collected for these sites in May 2014, processed by the vendor Quantum Spatial, and provided as 1 m digital terrain model (DTM) and 1 m canopy height model (CHM) products. A Leica ALS70 system was used for this LiDAR data collection with an absolute vertical accuracy of 0.084 m at the 95% confidence level in open terrain. The DTM and CHM products are spatially continuous by using a hydro-flattening procedure that eliminates artifacts caused by increased variability in ranges or dropouts in laser returns due to low reflectivity of water. For Creamer’s Field, hyperspectral data were acquired by NASA’s Airborne Visible Infrared Imaging Spectrometer—Next Generation (AVIRIS-NG) on 23 July 2018 (Miller et al 2019) and data were available for this study at 5.2 m reflectance products.

### 2.4 Machine learning analyses

We applied an object-based machine learning ensemble approach to upscale snow depth measurements using the remote sensing datasets. This approach has proven effective for upscaling marsh biomass (Zhang et al 2018) and soil properties (Zhang et al 2019) in wetlands. First, we reduced the high dimensionality of hyperspectral imagery using the minimum noise fraction (MNF) algorithm. Here we applied the standard MNF (Green et al 1988) which orders the output bands along axes of minimum signal-to-noise-ratio. The derived MNF eigenimages show steadily decreasing quality and corresponding eigenvalues. We selected the first 18 MNF eigenimages based on visual inspection and evaluation of the corresponding eigenvalues, and stacked eigenimages into one image to be used for object generation. We applied the object-based image analysis (OBIA) technique that segments imagery into relatively homogeneous objects for analysis. OBIA is extensively applied for fine spatial resolution image processing to improve analytical results (Blaschke 2010). To apply OBIA in snow depth modeling, the MNF imagery was segmented into spatial objects with different shapes and sizes using the multiresolution segmentation algorithm. The segmentation results for our three sites are provided in supplementary figure 1 (available online at stacks.iop.org/ERL/16/065014/mmedia) with spatial object size varying...
from 3 m² to 1106 m². If LiDAR was available at a site, LiDAR statistical descriptors of DTM and CHM including mean and standard deviation were determined for each object and merged with spectral features, leading to a fused dataset. We spatially matched field measurements to image objects leading to a reference dataset for upscaling model calibration and validation. If multiple measurements were within an object, an average snow depth was determined for this object. Only field data collected in 2016 and 2018 were matched with remote sensing data to represent the lowest (2016) and highest (2018) end of winter snow depths at our sites. In total, we had 101, 32, and 281 reference objects for farmer’s loop, Creamer, and APEX, respectively.

Initial analyses of multiple machine learning algorithms revealed two commonly used machine learning algorithms, support vector machine (SVM; Vapnik 1995) and random forest (RF; Breiman 2001), were valuable for snow depth prediction. We thus selected SVM and RF to develop statistical/empirical models for snow depth estimation with field measurements as the dependent variable and remote sensing derived data as independent variables. Multiple linear regression (MLR) was also applied. We applied an ensemble approach to enhance the estimation by combining the predictions of multiple models using a weighted scheme (Zhang et al 2018, 2019) based on the correlation coefficient ($r$) to assign weight to each model in the ensemble prediction. A model with a higher $r$ would get a higher weight and the sum of weights is 1.0. The accuracy of the estimation was evaluated using $r$, mean absolute error (MAE), and root mean squared error (RMSE). K-fold cross validation was used to avoid over-fitting issues if the same dataset was used for both validation and calibration. We set $k$ to 5 for Farmer’s Loop and Apex datasets, and three to Creamer’s Field dataset due to its less matched reference objects.

We also conducted a cross transect validation for Creamer’s Field and APEX using the independent snow depth dataset collected in 2017. The ecotype map for each site was also produced by classifying the remote sensing dataset using the RF classifier, field data as training dataset, and OBIA techniques. The classification procedure was detailed in Zhang et al (2020). We achieved an overall accuracy of larger than 90% in ecotype mapping. We also conducted experimental analyses to examine the use of LiDAR data for snow depth modeling by including/excluding LiDAR features in the model as predictors.

3. Results

3.1. Field snow and ecotype surveys and climatology

Between 2014 and 2019, the Fairbanks climate was slightly warmer and wetter in winter and summer than the long-term (96 year) average. Total snowfall from November through March for 2014–2019 ranked as the #34, 68, 50, 15, 17, and 54th out of 90 winters, in order (Jorgenson et al 2020). Of particular note are the winters of 2015–2016 and 2017–2018 because the machine learning analyses were applied toward the end of winter snowpack in those two winters. The winter of 2015–2016 was the warmest in the record and had a slightly below average end of winter snowfall. The 2017–2018 winter was slightly colder than the mean with a total snowpack more than a standard deviation higher than the long-term mean.

Snow depth measurements, including the mean and upper and lower 75% quartiles for the depth measurements at each site and ecotype (Douglas et al 2019), are summarized in figure 2 and table 1. Results from a student’s $t$-test analysis identifies ecotypes with statistically significantly deeper or shallower snow for the different ecotypes. At all sites and across all years the tussock tundra snowpack is statistically significantly deeper than the mixed forest snowpack. For all years and sites except Farmer’s Loop in 2015 the moss spruce forest also has a statistically significantly deeper snow than the mixed forest. The moss spruce forest ecotype has the deepest snow at the APEX site but the means are not always statistically significantly different from the other ecotypes. The end of season snowpack in the disturbed ecotypes is typically lower than the tussock tundra and higher than the mixed forest but disturbed ecotype snow depths are not statistically significantly different from the forested ecotypes. The end of winter snowpack in 2018 was roughly double what was measured in 2016 for all ecotypes.

3.2. Machine learning analyses

Table 2 shows performance of the object-based machine learning models and MLR for snow depth upscaling for 2016 (the lowest snowpack) and 2018 (the deepest snowpack). For the Farmer’s Loop and APEX sites we also included the model comparison with/without LiDAR data. The results show inclusion of LiDAR in the model increases the accuracy of the snow depth estimation for these two sites by more than 10% for Farmer’s Loop and 3% for APEX based on $r$. However, for Creamer’s Field, the contribution from LiDAR is marginal and thus LiDAR was not used for snow depth mapping (results were not shown). In general, the models produce encouraging results with correlation coefficient ($r$) values varying from 0.60 to 0.84, MAE in the range of 2.2–9.1 cm, and the RMSE varying from 2.8 cm to 11.5 cm. The Farmer’s Loop and Creamer’s Field sites generate consistent results with a higher $r$ for 2018 than 2016, although higher MAE and RMSE of 2018 are generated due to the deeper snow in 2018. This is expected because these two sites are geographically close. In contrast, for the APEX site, there is a higher accuracy for 2016 than 2018. Across the algorithms, SVM
achieved the best estimation for farmer’s loop; MLR produced the best estimation for Creamer’s Field in 2016 and RF achieved the highest accuracy in 2018; for APEX, SVM produced the best prediction in 2015 while RF generated the highest accuracy in 2018. These three algorithms had a comparable accuracy in terms of $r$, MAE and RMSE, but different predictions. Therefore, an ensemble analysis was conducted to make the snowpack predictions more robust. The accuracy of the ensemble analysis for each site is also provided in table 2. Ensemble analysis slightly increases the accuracy compared to each individual model. The cross transect validation result was also encouraging with a high $R^2$ for Creamer’s Field and APEX sites where additional snow depth measurements were made in 2017. The scatterplot of measured and predicted snow depth and the corresponding regression line are displayed in figure 3 for the cross transect validation. The model was calibrated by our measurements and applied to predict snow depth.
Table 1. A summary of mean snow depths (in cm) and standard deviation values (in parentheses) for the field sites and ecotypes in this study.

|                  | Creamer’s field | Farmer’s Loop | APEX          |
|------------------|-----------------|---------------|---------------|
|                  | Disturbed       | Mixed forest  | Wetland       | Tusock tundra | Moss spruce forest |
| 2014             | 59.2 (7.4)      | 42.3 (5.8)    | 54.1 (9.2)    | 59.7 (11.5)   | 55.4 (14.1)        |
| 2015             | 56.6 (7.0)      | 37.4 (6.2)    | 51.6 (7.3)    | 57.3 (9.7)    | 49.4 (8.6)         |
| 2016             | 45.8 (8.8)      | 32.8 (5.6)    | 38.4 (8.2)    | 41.0 (8.8)    | 42.9 (8.4)         |
| 2017             | 80.9 (6.2)      | 66.1 (7.0)    | 72.9 (5.6)    | 81.3 (8.3)    | 75.1 (10.1)        |
| 2018             | 84.0 (8.0)      | 66.3 (7.7)    | 81.0 (8.8)    | 88.5 (9.0)    | 81.9 (9.7)         |
| 2019             | 62.2 (8.2)      | 45.0 (6.4)    | 62.6 (8.3)    | 68.3 (9.8)    | 67.9 (10.4)        |
| 2020             | 85.7 (7.7)      | 62.6 (8.1)    | 80.1 (9.0)    | 82.8 (9.2)    | 72.3 (15.5)        |

Maps of ecotype, snow depth measured along the transects in 2016 and 2018, and snow depths estimated from the ensemble analysis of three model outputs (SVM, RF, and MLR) in 2016 and 2018 for our three field sites are displayed in figures 4–6. Snow depth was estimated using the inputs of MNF transformed hyperspectral reflectance and LiDAR features at Farmer’s Loop and APEX, and only MNF transformed hyperspectral reflectance for Creamer’s Field. The spatial heterogeneity of snow depth is well delineated responding to the spatial heterogeneity of vegetation composition with shallow snowpack in forests and a deeper snowpack in open areas. Maps of estimated snow depth, by ecotype, agree well with field measurements. The change between modeled 2016 and 2018 snow depths (figures 4(e), 5(e), and 6(e)) shows open areas with little forest canopy (tusock tundra at farmer’s loop and Creamer’s Field and wetland at APEX) have proportionally deeper snow in a high snowpack year than the mixed forest and moss spruce forest do. The coverage of wetland and disturbed areas at Farmer’s Loop are small (~100 m on a side). The disturbed areas include power line and other clearings and trails and the wetland feature is a small ephemeral stream. The wetland is associated with far deeper snow in 2018. The ecotype-snowpack depth relationships are far more evident in the deeper snow year (2018) than the shallower snow year (2016) based on maps for all three sites.

4. Discussion, implications, and conclusions

This study shows repeated and predictable relationships between end of winter snowpack depth and ecotype across three Interior Alaska study sites that represent almost three quarters of the ecotypes in the boreal and taiga biome. When vegetation type, canopy density, and/or structure are changed from climate warming, wildfire, or other disturbances ramifications for ecosystem protection of permafrost can be projected through thermal models (Loranty et al 2018, Kropp et al 2020). A 90 year meteorologic record for the Fairbanks area indicates the length of the summer growing season has increased by 45% (Wendler and Shulski 2009). This equates to a shortening of the winter season (days at or below the 0 °C freezing point) by 38 days, predominantly through an earlier onset of springtime. The study also found winter precipitation has decreased by 20% since 1916. The snow season has decreased by approximately 2.5 d decade$^{-1}$ (Euskirchen et al 2006, Euskirchen et al 2007). The warming and shortening of the winter season are expected to continue (Lader et al 2017, 2020, Littell et al 2018). The projected shallower snowpack could allow more winter cooling to reach the ground surface, however, warmer winter temperatures could reduce the rates of ground freezing. Though our study does not focus on detailed relationships between winter climatology (weather data, particularly winds) it does show that changes in vegetation cover also have a marked effect on snow depth.
Table 2. Model performance for estimating snow depth in 2016 and 2018 using hyperspectral imagery and comparison with/without LiDAR data. SVM: support vector machine; RF: random forest; MLR: multiple least regression; MAE (in cm): mean absolute error; RMSE (in cm): root mean square error. Results from an Ensemble Analysis (EA) of the three models with LiDAR data are also listed.

| Site            | 2016  | 2018  |
|-----------------|-------|-------|
|                 | SVM   | SVM   | RF    | RF    | MLR   | MLR   | EA    | EA    |
|                 | With LiDAR | Without LiDAR | With LiDAR | Without LiDAR | With LiDAR | Without LiDAR | | |
| Farmer’s Loop   |       |       |       |       |       |       |       | |
| 2016            |       |       |       |       |       |       |       | |
| r               | 0.64  | 0.56  | 0.62  | 0.51  | 0.6   | 0.53  | 0.66  | |
| MAE             | 5.5   | 5.6   | 5.3   | 5.8   | 6     | 6.3   | 6.2   | |
| RMSE            | 6.9   | 7.3   | 6.8   | 7.4   | 7.2   | 7.8   | 6.4   | |
| 2018            |       |       |       |       |       |       |       | |
| r               | 0.75  | 0.69  | 0.71  | 0.65  | 0.71  | 0.68  | 0.75  | |
| MAE             | 5.9   | 6.6   | 6.5   | 6.8   | 7.3   | 7.2   | 6     | |
| RMSE            | 8.1   | 9.1   | 8.7   | 9.3   | 9.1   | 9.3   | 8.1   | |
| Creamer’s Field | 2016  |       | 2018  |       |       |       |       | |
| 2016            |       |       |       |       |       |       |       | |
| R               | 0.67  | 0.66  | 0.71  | 0.73  |       |       |       | |
| MAE             | 2.4   | 2.3   | 2.4   | 2.2   | 2.8   | 2.8   | 2.8   | |
| RMSE            | 3.1   | 3     | 3     | 2.8   | 3.1   | 3     | 3.1   | |
| 2018            |       |       |       |       |       |       |       | |
| R               | 0.85  | 0.87  | 0.79  | 0.86  |       |       |       | |
| MAE             | 4.4   | 4.3   | 5     | 4.1   | 5     | 5.1   | 5     | |
| RMSE            | 5.3   | 5.2   | 6     | 5     | 5.3   | 5.2   | 5     | |
| APEX            | 2016  |       | 2018  |       |       |       |       | |
| 2016            |       |       |       |       |       |       |       | |
| R               | 0.84  | 0.82  | 0.83  | 0.81  | 0.75  | 0.7   | 0.85  | |
| MAE             | 4.5   | 4.8   | 4.6   | 4.9   | 5.7   | 6.3   | 4.3   | |
| RMSE            | 6.1   | 6.5   | 6.2   | 6.6   | 7.4   | 8     | 6.1   | |
| 2018            |       |       |       |       |       |       |       | |
| R               | 0.69  | 0.69  | 0.73  | 0.72  | 0.63  | 0.59  | 0.72  | |
| MAE             | 8.1   | 8.1   | 7.7   | 8     | 9     | 9.1   | 7.9   | |
| RMSE            | 10.5  | 10.5  | 9.8   | 9.9   | 11.2  | 11.5  | 10    | |

and thus the soil thermal regime. However, to support how, where, and when future climate projections will affect the snowpack and, from that, the snowpack-ecotype-soil thermal regime a greater focus on wintertime meteorology is warranted.

Tussock tundra and moss spruce forest were associated with the deepest snowpacks while mixed forests typically had the shallowest snow. From the perspective of water resources and snowmelt runoff hydrologic processes we assume that in the boreal biome deeper snowpacks have larger SWE (Sturm et al 2010). The deeper snow cover above the tussocks would be expected to lead to warmer winter surface soil temperatures by insulating the ground from winter cold. However, protrusions of large graminoid tussocks through the snowpack, particularly at the onset of winter, lead to earlier freeze-back of the active layer and cooler permafrost temperatures (Kholodov et al 2012, Loranty et al 2018). Because of this structural characteristic, tussocks are associated with the coldest snow-ground surface temperatures among the

Figure 3. Scatterplot and regression between predicted snow depths from ensemble analysis of three model outputs (SVM, RF and MLR) and measured snow depths for the Creamer’s Field site using hyperspectral data only, and the APEX site using hyperspectral and LiDAR data for cross transect validation in 2017.
Figure 4. Farmer's Loop site snow depth (cm) measured in (a) 2016, (b) 2018 overlaid on the ecotype map; estimated from ensemble analysis of three model outputs (SVM, RF and MLR) using hyperspectral and LiDAR data in (c) 2016 and (d) 2018, and (e) change between 2016 and 2018.

Figure 5. Creamer's Field site snow depth (cm) measured in (a) 2016, (b) 2018 overlaid on the ecotype map; estimated from ensemble analysis of three model outputs (SVM, RF and MLR) using hyperspectral data in (c) 2016 and (d) 2018, and (e) change between 2016 and 2018.
Figure 6. APEX site snow depth (cm) measured in (a) 2016, (b) 2018 overlaid on the ecotype map; estimated from ensemble analysis of three model outputs (SVM, RF and MLR) using hyperspectral and LiDAR data in (c) 2016 and (d) 2018, and (e) change between 2016 and 2018.

variety of ecotypes in our study area (Sturm et al 2001, Douglas et al 2021).

The variable canopy density of conifer trees in the moss spruce forests likely allow more snowfall to reach the ground in some areas than others. At the APEX site, areas with dense spruce cover had a deeper snowpack minimally affected by wind redistribution. However, at locations with smaller spruce trees that were spaced further apart the snowpack was shallower and more wind affected due to greater exposure to wind. Conifer density has been shown to dictate snowpack depth in other studies and wind was determined to be a major driver (Hedstrom and Pomeroy 1998, Pugh and Small 2013). We did not account for conifer density in this study, however, applications of airborne LiDAR to quantify canopy height and density could support more focused modeling efforts to link conifer density with snowpack characteristics.

The mixed forest ecotype has the shallowest snow of our study sites. The canopy is dense and multi-layered with birch and aspen typically found in the upper layer, willows and alder present at intermediate heights, and shrubs, tall grasses, and deadfall on the ground. This multi-layer canopy structure is effective at blocking snow from falling to the ground and thus largely reduces snow depth. The greater rates of interception and subsequent sublimation in mixed forests have been associated with shallow snowpack depths (Pomeroy et al 1999, 2003, 2012).

Our field and airborne measurements and machine learning modeling results reveal general relationships between ecotype and the bend of winter snowpack depth. An ensemble analysis of three models presented $r$ values between 0.66 and 0.86, MAE between 2.2 cm and 7.9 cm, and RMSE varying from 2.8 cm to 10 cm across the sites for the estimation of snow depth in 2016 (the lowest) and 2018 (the deepest) using hyperspectral and LiDAR data. We feel confident that these relationships and analyses could be applied to broader areas without snowpack measurements but where ecotype information is available. Hyperspectral measurements alone show great utility in calculating snow depth, however, hyperspectral measurements combined with LiDAR provide a slightly stronger tool for modeling snowpack depth. A direct application of hyperspectral reflectance and LiDAR data in snow depth mapping not only revealed the strong relationship between ecotype and snow depth, but also illustrated the snow depth variation within an ecotype caused by difference in vegetation structure (e.g. tree density/height). Application of hyperspectral imagery or multispectral satellite sensors such as WorldView-2/3 or some optical sensors with a similar spatial resolution may produce a similar ecotype map, even during summertime.
Climate warming and an increase in fire severity and extent in Interior Alaska have led to permafrost degradation and ecotype transitions. In lowland areas mixed forests that undergo permafrost thaw and subsidence tend to shift to wetlands (Lara et al. 2016). Some black spruce forests and the permafrost that underlies them are resilient to climate warming (Chapin et al. 2010). However, following high severity fires the conifer forests are increasingly transitioning to mixed forests (Beck et al. 2011). Based on our study a projected increase in wetland cover type from mixed forests should not be expected to be associated with a major change in snowpack depth, water equivalent, or thermal characteristics. However, a change from spruce to either birch or mixed forest will lead to a marked decrease in end of winter snowpack depths and water equivalent.

Three main results have emerged from this study. First, we established the use of optical sensors to upscale snow depth across a variety of ecotypes from limited field measurements. Second, we show that airborne hyperspectral and LiDAR measurements and machine learning techniques can be used to model snow depth to locations with sparse or no field data. Encouraging results were achieved by integrating multi-year field measurements with single-date remote sensing data and the application of ensemble analysis for documenting the spatial and temporal variation of snowpack depth. This means future changes in ecotype cover and their commensurate drivers on the thermal state of permafrost can be projected into the future. Finally, our study expands the application of AVIRIS-NG measurements for snowpack characterization. A machine learning based ensemble approach can be used to upscale snow depth measurements at regional to global scales by integrating optical satellite data with field measurements or other snow products.

**Data availability statement**

The data that support the findings of this study are available upon reasonable request from the authors.

**Acknowledgments**

This research was funded by the U.S. Army Engineer Research and Development Center Army Basic Research Program under PE 0601102/AB2 (Protection, Maneuver, Geospatial, Natural Sciences) and the Department of Defense’s Strategic Environmental Research and Development Program (Project RC18-1170). We thank Chris Hiemstra for providing data and insights into this work. Amanda Barker, Marc Beede, Art Gelvin, and Stephanie Saari for help with field measurements and Merritt Turetsky and Jamie Hollingsworth for coordinating access to the APEX site.

**ORCID iD**

Thomas A Douglas [https://orcid.org/0000-0003-1314-1905](https://orcid.org/0000-0003-1314-1905)

**References**

Anderson J E, Douglas T A, Barbato R A, Saari S, Edwards J D and Jones R M 2019 Vegetation mapping and seasonal thaw estimates in interior Alaska permafrost Remote Sens. Environ. 233 111363

Bair E H, Abreu Calfa A, Rittger K and Dozier J 2018 Using machine learning for real-time estimates of snow water equivalent in the watersheds of Afghanistan Cryosphere 12 1579–94

Beck P S, Goetz S J, Mack M C, Alexander H D, Jin Y, Randerson J T and Loranty M M 2011 The impacts and implications of an intensifying fire regime on Alaskan boreal forest composition and albedo Glob. Change Biol. 17 2853–66

Bennett K E, Cherry J E, Balk B and Lindsey S 2019 Using MODIS estimates of fractional snow cover area to improve streamflow forecasts in interior Alaska Hydrol. Earth Syst. Sci. 23 2439–59

Bieniek P A, Bhatt U S, Walsh J E, Lader R, Griffith B, Roach J K and Thomann R I. 2018 Assessment of Alaska rain-on-snow events using dynamical downscaling J. Appl. Meteorol. Climatol. 57 1847–63

Blaschke T 2010 Object based image analysis for remote sensing ISPRS J. Photogramm. Remote Sens. 65 2–16

Boelman N T et al 2019 Integrating snow science and wildlife ecology in Arctic-boreal North America Environ. Res. Lett. 14 010401

Breiman L 2001 Random forests Mach. Learn. 45 5–32

Brown D R, Jorgenson M T, Douglas T A, Romanovsky V E, Kielland K, Hiemstra C, Euskirchen E S and Ruess R W 2015 Interactive effects of wildfire and climate on permafrost degradation in Alaskan lowland forests J. Geophys. Res.: Biogeosci. 120 1619–37

Broxton P D, van Leeuwen W J and Biederman J A 2019 Improving snow water equivalent maps with machine learning of snow survey and lidar measurements Water Resour. Res. 55 3739–57

Chapin F S et al 2000 Arctic and boreal ecosystems of western North America as components of the climate system Glob. Change Biol. 6 211–23

Chapin F S et al 2010 Resilience of Alaska’s boreal forest to climatic change Can. J. For. Res. 40 1360–70

Dittmar T and Kattner G 2003 The biogeochemistry of the river and shelf ecosystem of the Arctic Ocean: a review Mar. Chem. 83 103–20

Domine F, Taillandier A, Houdier S, Parrenin F, Simpson W R and Douglas T A 2006 Interactions between snow metamorphism and climate: physical and chemical aspects Spec. Publ.-R. Soc. Chem. 1 311–27

Douglas T A et al 2021 Recent degradation of Interior Alaska permafrost mapped with ground surveys, geophysics, deep drilling, and repeat airborne LiDAR Cryosphere [https://doi.org/10.5194/tc-2021-47](https://doi.org/10.5194/tc-2021-47)

Douglas T A, Hiemstra C A and Barker A J 2019 ABoVE: end of season snow depth at CRREL sites near Fairbanks, Alaska,
2014–2019 (Oak Ridge, TN: ORNL DAAC) (https://doi.org/10.3334/ORNLDAAC/1702)

Douglas T A, Jones M C, Hiemstra C A and Arnold J 2014 Sources and sinks of carbon in boreal ecosystems of Interior Alaska: current and future perspectives for land managers Ecol. Sci. Anth. 2 000032

Douglas T A, Jorgenson M T, Brown D R, Campbell S W, Hiemstra C A, Saari S P, Bjella K and Liljedahl A K 2016 Degrading permafrost mapped with electrical resistivity tomography, airborne imagery and LiDAR, and seasonal thaw measurements Geophys. 81 WA71–85

Euskirchen E S et al 2006 Importance of recent shifts in soil thermal dynamics on growing season length, productivity, and carbon sequestration in terrestrial high-latitude ecosystems Global Change Biol. 12 731–50

Euskirchen E S, Bennett A P, Breen A L, Genet H, Lindgren M A, Kurkowski T A, McGuire A D and Rupp T S 2016 Consequences of changes in vegetation and snow cover for climate feedbacks in Alaska and northwest Canada Environ. Res. Lett. 11 105003

Euskirchen E S, McGuire A D, Chapin F S and Rupp T S 2010 The changing effects of Alaska’s boreal forests on the climate system Can. J. For. Res. 40 1536–46

Euskirchen E S, McGuire A D and Chapin III F S 2007 Energy feedbacks of northern high-latitude ecosystems to the climate system due to reduced snow cover during 20th century warming Glob. Change Biol. 13 2425–38

Fisher J B et al 2018 Missing pieces to modeling the arctic-boreal puzzle Environ. Res. Lett. 13 020202

Green A A, Berman M, Switzer P and Craig M D 1988 A transformation for ordering multispectral data in terms of image quality with implications for noise removal IEEE Trans. Geosci. Remote Sens. 26 65–74

Grünewald J, Wilcox E J, Zwieback S, Marsh P and Boike J 2020 Linking tundra vegetation, snow, soil temperature, and permafrost Biogeosciences 17 4261–79

Hedstrom N R and Pomeroy J W 1998 Measurements and modelling of snow interception in the boreal forest Hydrol. Process. 12 1611–25

Holloway J E, Lewkowicz A G, Douglas T A, Li X, Turetsky M R, Baltzer J L and Jin H 2020 Impact of wildfire on permafrost landscapes: a review of recent advances and future prospects Permafrost Periglac. Process. 31 371–82

Jacobi H W, Domine F, Simpson W R, Douglas T A, Jones M C, Hiemstra C A and Arnold J 2014 Sources and sinks of carbon in boreal ecosystems of Interior Alaska: current and future perspectives for land managers Ecol. Sci. Anth. 2 000032

Jafarov E E, Coen E T, Hargrave B S, Wilson C J, Painter S L, Atchley A J and Romanovsky V E 2018 Modeling the role of preferential snow accumulation in throughfall development and hillslope groundwater flow in a transitioning permafrost landscape Environ. Res. Lett. 13 105006

Jafarov E E, Nicolaus D J, Romanovsky V E, Walsh J E, Panda S K and Serreze M C 2014 The effect of snow: how to better model ground surface temperatures Cold Reg. Sci. Technol. 102 63–77

Johnstone J F, Chapin F S, Hollingsworth T N, Mack M C, Romanovsky V and Turetsky M 2010 Fire, climate change, and forest resilience in interior Alaska Can. J. For. Res. 40 1302–12

Jorgenson M T et al 2008 Permafrost characteristics of Alaska (Proc. Ninth Int. Conf. on Permafrost)

Jorgenson M T, Douglas T A, Liljedahl A K, Roth J E, Catter T C, Davis W A, Frost G V, Miller P F and Racine C H 2020 The roles of climate extremes, ecological succession, and hydrology in repeated permafrost aggradation and degradation in fens on the tanaana flats, Alaska J. Geophys. Res.: Biogeosci. 125

Jorgenson M T, Racine C H, Walters J C and Osterkamp T E 2001 Permafrost degradation and ecological changes associated with a warming climate in central Alaska Clim. Change 48 551–79

Judd K E and Kling G W 2002 Production and export of dissolved C in arctic tundra mesocosms: the roles of vegetation and water flow Biogeochemistry 60 213–34

Khodakov A et al 2012 Regional and local variability of modern natural changes in permafrost temperature in the Yakutian coastal lowlands, Northeastern Siberia Proc. Tenth Int. Conf. Permafrost, 2012

Koci N, Laudon H, Ottosson-Löfvenius M and Hasselquist N J 2017 Increasing water losses from snow captured in the canopy of boreal forests: a case study using a 30 year data set Hydrol. Process. 31 3558–67

Kropp H et al 2020 Shallow soils are warmer under trees and tall shrubs across Arctic and Boreal ecosystems Environ. Res. Lett. 16 015001

Kumpula J and Colpaert A 2007 Snow conditions and usability value of pastureland for semi-domesticated reindeer (Rangifer tarandus tarandus) in northern boreal forest area Rangifer 27 16

Lader R, Walsh J E, Bhatt U S and Brienick P A 2017 Projections of twenty-first-century climate extremes for Alaska via dynamical downscaling and quantile mapping J. Appl. Meteorol. Climatol. 56 2393–409

Lader R, Walsh J E, Bhatt U S and Briebenick P A 2020 Anticipated changes to the snow season in Alaska: elevation dependency, timing and extremes Int. J. Climatol. 40 169–87

Lara M J, Genet H, McGuire A D, Euskirchen E S, Zhang Y, Brown D R N, Jorgenson M T, Romanovsky V, Breen A and Bolton W R 2016 Thermokarst rates intensify due to climate change and forest fragmentation in an Alaskan boreal forest lowland Glob. Change Biol. 22 816–29

Latirovic R, Pouliot D and Olhoff I 2017 Circa 2010 land cover of canada: local optimization methodology and product development Remote Sens. 9 1098

Ling F and Zhang T 2003 Impact of the timing and duration of seasonal snow cover on the active layer and permafrost in the Alaskan Arctic Permafrost Periglacial Process. 14 141–50

Liston G E and Hiemstra C A 2011 The changing cryosphere: pan-arctic snow trends (1979–2009) J. Clim. 24 5691–712

Littell J S, McAfee S A and Hayward G D 2018 Alaska snowpack response to climate change: statewide snowfall equivalent and snowpack water scenarios Water 10 668

Loranty M M et al 2018 Changing ecosystem influences on soil thermal regimes in northern high-latitude permafrost regions Biogosciences 15 5287–313

Loranty M M, Berner L T, Goetz S J, Jin Y and Randerson J T 2014 Vegetation controls on northern high-latitude snow-albedo feedback: observations and CMIP5 model simulations Glob. Change Biol. 20 594–609

McPartland M Y et al 2019 Characterizing boreal peatland plant composition and species diversity with hyperspectral remote sensing Remote Sens. 11 1685

Mekonnen Z A, Santolaria-Otín M, Krinner G, Ménégoz M, Mudryk L, and Serreze M C 2014 The effect of snow: how to better model ground surface temperatures Cold Reg. Sci. Technol. 102 63–77

Miller C E et al 2013 ABoVE: hyperspectral imagery from AVIRIS-NG, Alaskan and Canadian Arctic, 2017–2018 (Oak Ridge, TN: ORNL DAAC) (https://doi.org/10.3334/ORNLDAAC/1569)

Mudryk L, Santolaria-Otín M, Krinner G, Ménégoz M, Derksen C, Brulet-Uvillet C, Brady M and Essery R 2020 Historical Northern Hemisphere snow cover trends and projected changes in the CMIP6 multi-model ensemble Cryosphere 14 595–614

Olness J and Kieland K 2017 Asynchronous recruitment dynamics of snowshoe hares and white spruce in a boreal forest For. Ecol. Manage. 384 83–91

Pomeroy J W and Essery R L 1999 Turbulent fluxes during blowing snow: field tests of model sublimation predictions Hydrol. Process. 13 2963–75
Pomeroy J W, Storck P, Parviainen J and Essery R 2003 Sublimation of snow from coniferous forests in a climate model J. Clim. 16 1855–64

Pomeroy J, Fang X and Ellis C 2012 Sensitivity of snowmelt hydrology in Marmot Creek, Alberta, to forest cover disturbance Hydrol. Process. 26 1891–904

Pozzanghera C B, Sivy K J, Lindberg M S and Prugh L R 2016 Variable effects of snow conditions across boreal mesocarnivore species Can. J. Zool. 94 697–705

Pugh E T and Small E F 2013 The impact of beetle-induced conifer death on stand-scale canopy snow interception Hydrol. Res. 44 644–57

Semmens K A and Ramage J M 2013 Recent changes in spring snowmelt timing in the Yukon River basin detected by passive microwave satellite data Cryosphere 7 905–16

Shur Y and Jorgenson M 2007 Patterns of permafrost formation and degradation in relation to climate and ecosystems Permafrost Periglacial Process. 18 7–19

Sturm M 1992 Snow distribution and heat flow in the taiga Arct. Alp. Res. 24 145–52

Sturm M and Benson C S 1997 Vapor transport, grain growth and depth-hoar development in the subarctic snow J. Glaciol. 43 42–59

Sturm M and Benson C 2004 Scales of spatial heterogeneity for perennial and seasonal snow layers Ann. Glaciol. 38 253–60

Sturm M, Douglas T A, Racine C and Liston G E 2005 Changing snow and shrub conditions affect albedo with global implications J. Geophys. Res.: Biogeosci. 110 G01004

Sturm M and Holmgren J 2018 An automatic snow depth probe for field validation campaigns Water Resour. Res. 54 9695–701

Sturm M, Holmgren J and Liston G E 1995 A seasonal snow cover classification system for local to global applications J. Clim. 8 1261–83

Sturm M, Holmgren J, McPadden J P, Liston G E, Chapin III F S and Racine C H 2001 Snow–shrub interactions in Arctic tundra: a hypothesis with climatic implications J. Clim. 14 336–44

Sturm M, Taras B, Liston G E, Derksen C, Jonas T and Lea J 2010 Estimating snow water equivalent using snow depth data and climate classes J. Hydrometeorol. 11 1380–94

Taillandier A S, Domine F, Simpson W R, Sturm M, Douglas T A and Severin K 2006 Evolution of the snow area index of the subarctic snowpack in central Alaska over a whole season. Consequences for the air to snow transfer of pollutants Environ. Sci. Technol. 40 7521–7

Vapnik V N 1995 The Nature of Statistical Learning Theory (Berlin: Springer)

Wang J, Yuan Q, Shen H, Liu T, Li T, Yue L, Shi X and Zhang L 2020 Estimating snow depth by combining satellite data and ground-based observations over Alaska: a deep learning approach J. Hydrol. 585 124828

Wendler G and Shulski M 2009 A century of climate change for Fairbanks, Alaska Arctic 62 295–300

Yi Y, Kimball J S, Rawlins M A, Moghaddam M and Euskirchen E S 2015 The role of snow cover affecting boreal-arctic soil freeze–thaw and carbon dynamics Biogosciences 12 5811–29

Zhang C, Denka S, Cooper H and Mishra D R 2018 Quantification of sawgrass marsh aboveground biomass in the coastal Everglades using object-based ensemble analysis and Landsat data Remote Sens. Environ.

Zhang C, Douglas T A and Anderson J 2020 Mapping vegetation and seasonal thaw depth in central Alaska using airborne hyperspectral and lidar data 2020 IEEE Int. Geoscience and Remote Sensing Symp. (IGARSS) (https://doi.org/10.1109/IGARSS59084.2020.9323660)

Zhang C, Douglas T A and Anderson J 2021 Modeling and mapping permafrost active layer thickness using 1 m airborne hyperspectral imagery Int. J. Appl. Earth Obs. Geoinf. (in review)

Zhang C, Mishra D R and Penning S 2019 Mapping salt marsh soil properties using imaging spectroscopy ISPRS J. Photogramm. Remote Sens. 148 221–34

Zhang T 2005 Influence of the seasonal snow cover on the ground thermal regime: an overview Rev. Geophys. 43 RG4002