

Article

Regional Quantitative Cover Mapping of Tundra Plant Functional Types in Arctic Alaska

Matthew J. Macander ^{1,*} , Gerald V. Frost ¹, Peter R. Nelson ² and Christopher S. Swingley ¹

¹ ABR, Inc.—Environmental Research & Services, P.O. Box 80410, Fairbanks, AK 99708, USA; jfrost@abrinc.com (G.V.F.); cswingley@abrinc.com (C.S.S.)

² University of Maine—Fort Kent, 23 University Dr., Ft. Kent, ME 04743, USA; peter.nelson@maine.edu

* Correspondence: mmacander@abrinc.com; Tel.: +1-907-455-6777 (ext. 112)

Received: 17 August 2017; Accepted: 29 September 2017; Published: 4 October 2017

Abstract: Ecosystem maps are foundational tools that support multi-disciplinary study design and applications including wildlife habitat assessment, monitoring and Earth-system modeling. Here, we present continuous-field cover maps for tundra plant functional types (PFTs) across ~125,000 km² of Alaska's North Slope at 30-m resolution. To develop maps, we collected a field-based training dataset using a point-intercept sampling method at 225 plots spanning bioclimatic and geomorphic gradients. We stratified vegetation by nine PFTs (e.g., low deciduous shrub, dwarf evergreen shrub, sedge, lichen) and summarized measurements of the PFTs, open water, bare ground and litter using the cover metrics total cover (areal cover including the understory) and top cover (uppermost canopy or ground cover). We then developed 73 spectral predictors derived from Landsat satellite observations (surface reflectance composites for ~15-day periods from May–August) and five gridded environmental predictors (e.g., summer temperature, climatological snow-free date) to model cover of PFTs using the random forest data-mining algorithm. Model performance tended to be best for canopy-forming PFTs, particularly deciduous shrubs. Our assessment of predictor importance indicated that models for low-statured PFTs were improved through the use of seasonal composites from early and late in the growing season, particularly when similar PFTs were aggregated together (e.g., total deciduous shrub, herbaceous). Continuous-field maps have many advantages over traditional thematic maps, and the methods described here are well-suited to support periodic map updates in tandem with future field and Landsat observations.

Keywords: plant functional types; Arctic tundra; vegetation mapping; random forest; phenology; reflectance composites; Landsat; North Slope; Alaska

1. Introduction

Quantitative measurements of vegetation and environmental covariates support the classification and mapping of terrestrial ecosystems. Ecosystem map products, in turn, are foundational tools for diverse applications including ecosystem monitoring, landscape-change detection, wildlife habitat assessment and simulation modeling, and they serve as a basis for the design of spatially-stratified studies. Most Arctic ecosystem mapping efforts to date have applied categorical classifications, such as vegetation or land cover types, using raster analysis [1–5] or visual photo-interpretation [6,7]. Thematic approaches emphasize the central tendencies of map classes, but variability within classes is difficult to effectively retain in the map products. For example, categorical vegetation maps have limitations for long-term monitoring because changes in vegetation properties over time are more likely to involve shifts within the range of variability of a given map class, rather than a transition from one class to another. A more precise approach is to quantitatively map the cover of vegetation within biophysically-meaningful strata, such as plant functional types (PFTs) [8,9]; to date, quantitative or “continuous field” vegetation mapping efforts have mainly focused on forest ecosystems [10].

The dramatic loss of summer sea-ice and a warming climate are driving rapid changes to terrestrial ecosystems across much of the Arctic, where vegetation productivity and permafrost conditions are strongly influenced by summer temperature [11–13]. Quantitative ecosystem mapping approaches are therefore urgently needed to better document and monitor the composition and structure of tundra vegetation, as well as ecosystem properties that are strongly mediated by vegetation such as surface energy balance [14], albedo [15,16] and permafrost thermal regime [17,18]. For example, quantitative map products are well suited for the parameterization of spatially-explicit Earth-system models [19–21], scaling field measurements of vegetation cover and biomass [22] and to serve as inputs to resource selection functions for wildlife [23]. Yet, the development of robust field datasets over large spatial extents is logistically challenging in the Arctic, and remote sensing approaches must overcome the short growing season, persistent cloudiness, low stature and strong spectral similarities of tundra PFTs.

Shrub expansion is the most conspicuous form of tundra vegetation change being observed in Arctic Alaska [24] and is of particular interest because the development of an overtopping shrub canopy has consequences for surface energy balance, snow regime and many other ecosystem properties and processes in Arctic tundra [25]. For example, erect shrub canopies can strongly affect the thermal regime of permafrost by shading the ground in summer and trapping insulative snow in winter [26]. Shrub expansion often results in a reduction in local species diversity and the abundance of other PFTs, such as ground lichens [27]. Thus, spatially-explicit estimates of modern shrub cover are desirable for ecosystem modeling and to serve as a baseline for long-term monitoring using Earth-observing satellites [28,29]. Previous efforts to map Arctic shrub cover, however, have estimated shrub cover at large spatial scales with little field data, estimated shrub abundance within cover categories or aggregated shrub PFTs into a single category. Here, we seek to improve the precision, accuracy and usefulness of tundra shrub mapping by partitioning shrub PFTs by stature (tall, low and dwarf) and leaf habit (deciduous and evergreen), as well as using detailed field measurements of shrub cover to train our models.

Mosses and lichens have generally been underrepresented in studies of Arctic vegetation dynamics. Yet, mosses strongly influence Arctic soils and permafrost because of their insulative and moisture-holding properties [30], and they play a critical role in soil paludification and acidification [31]. Lichens constitute most of the winter diet of caribou (*Rangifer tarandus*) and are useful bioindicators of local environmental conditions such as soil moisture and disturbance regime [32–34]. Efforts to predict nonvascular plant cover using remote sensing are greatly complicated, however, by their low stature. Remote sensing approaches in the Arctic have tended to rely on observations from midsummer, when the spectral signature of low-growing PFTs is greatly obscured by overtopping vegetation. We sought to overcome these challenges by developing a seasonal archive of Landsat observations representing most of the snow-free season (~late May–August).

Here, we present quantitative spatial models of tundra PFTs, plant litter and nonvegetated cover types at 30-m resolution for a ~125,000 km² study area spanning most of Alaska's western and central North Slope (Figure 1). The study area encompasses strong climatic, geomorphic and topographic gradients that influence the structure and composition of tundra vegetation from regional to landscape scales. The large size and heterogeneity of the study area necessitated the development of a multi-temporal archive of Landsat data with minimal cloud contamination and scan-line artifacts and compilation of ancillary spatial data that capture important environmental gradients and ecosystem covariates.

In order to quantify vegetation properties involving relationships between a large set of co-varying predictors, we employed the random forest data-mining algorithm [35] to relate the training data to spectral and environmental predictors (Figure 2). Our specific objectives were to develop spatial models for the cover of nine tundra PFTs, plant litter, open water and bare ground (Table 1). The nine PFTs were defined by growth form and leaf habit: tall deciduous shrub (>1.5 m height), low deciduous shrub (0.2–1.5 m), dwarf deciduous shrub (<0.2 m), dwarf evergreen shrub, sedge, grass, forb, moss

and lichen; we also mapped five aggregate PFTs in which similar PFTs were combined (e.g., total shrubs, total herbaceous).

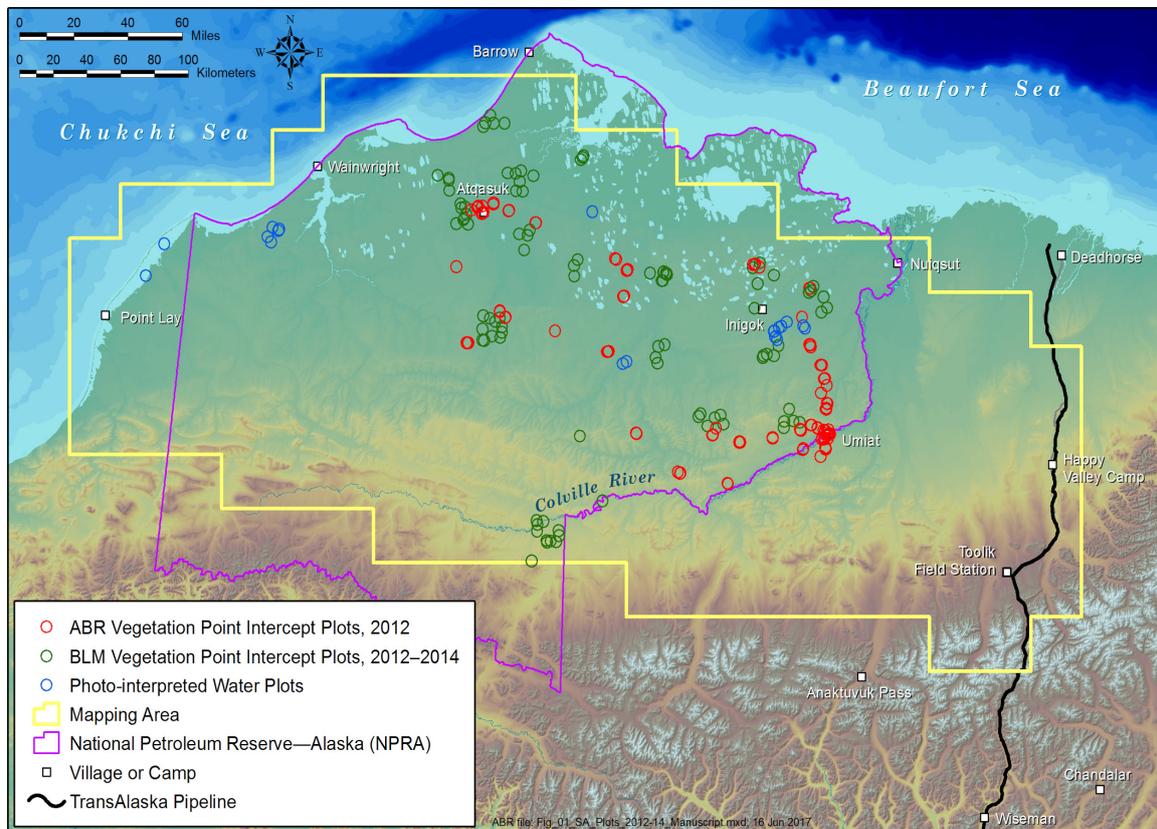


Figure 1. Map of quantitative-cover mapping area showing the distribution of field plots, North Slope, Alaska. ABR, ABR, Inc.—Environmental Research & Services. BLM, Bureau of Land Management.

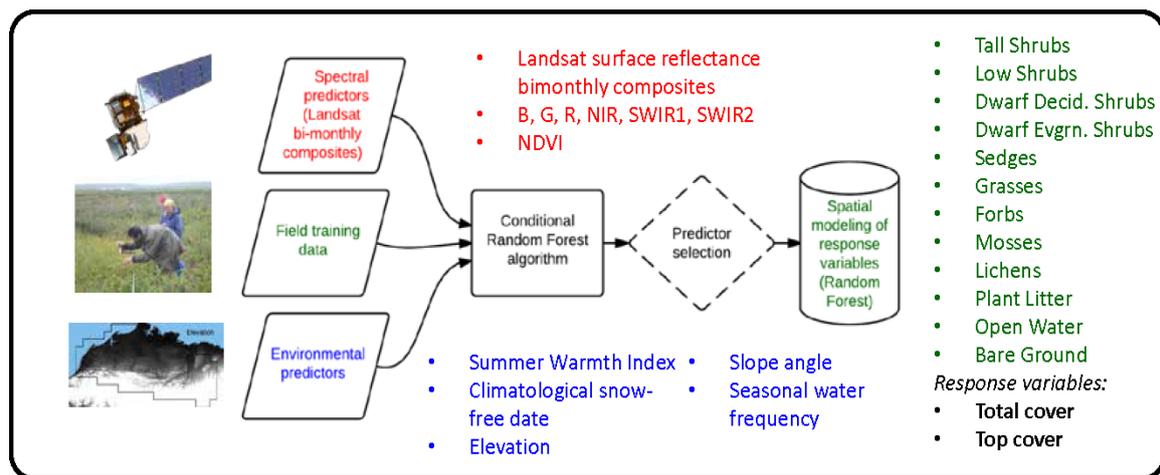


Figure 2. Overview of random forest modeling process to predict cover of plant functional types (PFTs), litter and nonvegetated cover types, North Slope, Alaska.

Table 1. Tundra PFTs and other cover types used for cover modeling, North Slope, Alaska.

PFT or Other Cover Type	Description	Cover Metric(s) Modeled
Tall Deciduous Shrubs	deciduous shrubs ≥ 1.5 m height; mainly <i>Salix alaxensis</i> , <i>S. arbusculoides</i> and <i>Alnus fruticosa</i>	Total cover
Low Deciduous Shrubs	deciduous shrubs 0.2–1.5 m height; e.g., <i>Betula nana</i> , <i>Salix pulchra</i> and <i>S. glauca</i>	Total cover
Dwarf Deciduous Shrubs	deciduous shrubs ≤ 0.2 m height; e.g., <i>Salix rotundifolia</i> , <i>Arctous</i> spp. and <i>Vaccinium uliginosum</i> .	Total cover
Dwarf Evergreen Shrubs	evergreen shrubs ≤ 0.2 -m height; e.g., <i>Dryas integrifolia</i> , <i>Cassiope tetragona</i> , <i>Vaccinium vitis-idaea</i> and <i>Ledum decumbens</i>	Total cover
Sedges	herbaceous plants of <i>Cyperaceae</i> and <i>Juncaceae</i> ; e.g., <i>Carex</i> and <i>Eriophorum</i> spp.	Total cover
Grasses	herbaceous plants of <i>Poaceae</i> ; e.g., <i>Arctagrostis</i> , <i>Arctophila</i> , <i>Deschampsia</i>	Total cover
Forbs	non-graminoid herbaceous plants; e.g., legumes and composites	Total cover
Mosses	total of all mosses and liverworts	Total cover
Lichens	total of all ground lichens	Total cover
Total Herbaceous	total of sedges, grasses, and forbs	Total cover
Total Shrubs	total of all shrub PFTs	Total cover
Low & Tall Deciduous Shrubs	deciduous shrubs ≥ 0.2 -m height	Total cover
Vascular Plants	total of all vascular PFTs	Total and top cover
Nonvascular Plants	total of all nonvascular PFTs including mosses, terricolous lichens, liverworts and algae	Total and top cover
Litter	dead plant matter	Top cover
Open Water	open water	Top cover
Bare Ground	bare soil or rock	Top cover

2. Materials and Methods

2.1. Study Area

The 125,000-km² mapping area extends from the Chukchi Sea near Point Lay eastward to the Dalton Highway and encompasses most of National Petroleum Reserve—Alaska (NPRA) (Figure 1). Strong climatic and topographic gradients produce substantial regional differences in vegetation structure and composition [22]. The northern half of the mapping area is part of the Arctic Coastal Plain physiographic province [36], characterized by flat topography, poorly drained soils, abundant waterbodies and widespread periglacial landforms such as ice-wedge polygons. Most of the Arctic Coastal Plain corresponds to Bioclimate Subzone D of the Circumpolar Arctic Vegetation Map [37]; shrubs are common, but they seldom exceed 1-m height. The southern half of the mapping area encompasses the Arctic Foothills physiographic province and is characterized by widespread uplands dissected by well-defined drainage networks. The Arctic Foothills are situated in Bioclimate Subzone E, the warmest tundra subzone; shrubs are widespread, and tall thickets (>1.5-m height) are common on floodplains and adjacent slopes, particularly along the Colville River. Hereafter, we refer to the two major physiographic provinces as “coastal plain” and “foothills.” The entire mapping area lies in the zone of continuous permafrost [38].

2.2. Model Inputs

2.2.1. Field Data Collection

We developed a detailed field dataset from an expansive network of plots that were established to support baseline environmental assessment and long-term ecological monitoring in and near the National Petroleum Reserve—Alaska (NPRO) (Figure 1). We sampled 106 field plots in and near NPRO during July–August 2012 and pooled our dataset with 119 plots sampled by the U.S. Bureau of Land Management (BLM) in 2012–2014 as part of BLM’s Assessment, Inventory and Monitoring program. All field efforts utilized a point-intercept sampling method tailored for long-term monitoring of vegetation [39]; this protocol also facilitated sampling at scales appropriate for the analysis of 30-m resolution Landsat data. Plots sampled by us consisted of three 50-m lines; each line began 5 m from the plot center to avoid vegetation that was trampled during plot setup (Figure 3). The azimuth of the first line was selected randomly, and the others were offset from the first by 120 degrees. We recorded vegetation “hits” at 1-m intervals using a rod-mounted laser pointer (51 sampling points per line), except at a few plots where logistical constraints necessitated quicker sampling at 2.5-m spacing (21 points per line). At each point, we identified vegetation hits by species by sequentially moving upper canopy leaves aside to allow the laser to hit underlying layers and finally the ground surface. For the ground surface, we recorded a single hit consisting either of live prostrate vegetation, litter, water or bare ground. Although we recorded multiple hits in the field, only the first live hit for each species was retained for this analysis. The BLM data were collected using the same protocol except that lines were 25 m long and point spacing was 0.5 m (51 points per line).

BLM allocated plots according to a stratified random sampling design based on the 2013 North Slope Science Initiative (NSSI) Land Cover map [5]; we subjectively selected plot locations in representative vegetation types based on photo-signatures evident in 2.5-m resolution aerial imagery. Waterbodies are abundant in the mapping area, but they are impractical to sample using the point-intercept method. Therefore, we defined 20 water-only plots that we photo-interpreted from satellite imagery within representative areas of clear and turbid water. Water-only plots comprised 9% of the total plot count, which matches the relative extent of open water portrayed on the NSSI map.

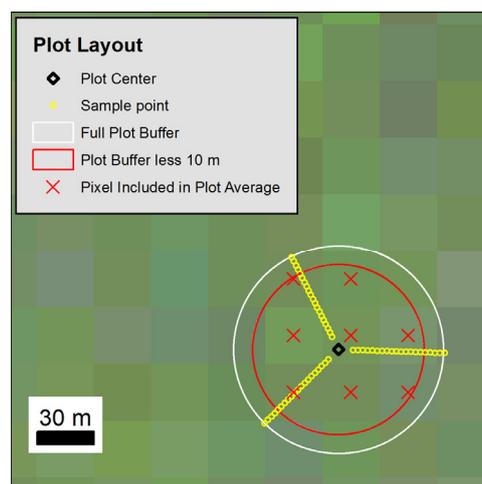


Figure 3. Layout of point-intercept field plot, North Slope, Alaska. Lines sampled by ABR were 50 m in length; lines sampled by BLM were 25 m. Pixels from a midsummer Landsat composite (red/green/blue) are in the background. The pixels with centers inside a circular plot buffer less than 10 m were included in the calculation of average predictor values for each plot.

For analysis, we summarized the cover data by species, then aggregated to PFTs using two cover metrics: (1) total cover, the percent of sample points at which species belonging to a PFT occurred,

summed for all species in a PFT; and (2) top cover, the percent of sample points at which a PFT was the first hit. Total cover values could exceed 100%, but top cover values could not exceed 100%. Importantly, because each species “hit” at a point was recorded and multiple species (including different species of the same PFT) could co-occur at different levels in the canopy, the sum of “total cover” values for all PFTs generally exceeded 100% and occasionally exceeded 100% for a single PFT; for example, in a very dense alder-willow shrubland with >50% cover of both species.

2.2.2. Spectral Predictors: Landsat Seasonal Composites

We compiled calibrated, atmospherically corrected and precision terrain corrected (L1T) Landsat Surface Reflectance High Level Data Products [40] derived from Landsat 4–5 Thematic Mapper (TM), Landsat 7 Enhanced Thematic Mapper (ETM+) and Landsat 8 Operational Land Imager (OLI) observations between 15 May and 31 August 1999–2015. This date range captures the spring, summer and early fall seasons on the North Slope. To facilitate the creation of Landsat image stacks from overlapping orbit paths, we developed a 30×30 -km tiling scheme in the Alaska Albers (North American Datum 1983) coordinate system.

Automated cloud- and cloud-shadow-masking algorithms such as the Automated Cloud-Cover Assessment [41] and Fmask [42] have limitations in Arctic regions. For example, snow and lake-ice are often misidentified as clouds, while surface water is often misidentified as cloud shadow, and thin clouds are frequently not masked. Therefore, we only used the Fmask results to reject tiles with >80% cloud contamination and visually selected cloud-free tiles and tiles for which it was feasible to mask clouds manually.

We developed six seasonal Landsat composites to represent ground conditions for key phenological windows from snowmelt to fall senescence (Figure 4), using all available Landsat data for specific date and year ranges. For each composite, we included six reflective bands as predictors: blue, green, red, near-infrared (NIR) and two short-wave infrared (SWIR) bands (Table 2). We attempted to construct composites for seven bimonthly windows of ~15 days (mid-May–August) using all available, recent (2010–2015) Landsat observations. This was possible for the late May, early June, late June and early July compositing windows, but cloud cover becomes more frequent on the North Slope after mid-July. To include more potential observations, we used a one-month seasonal window (16 July–15 August) to represent “midsummer”, the period of peak vegetation productivity. For the latest window (late August), it was not possible to lengthen the seasonal window because plant senescence is well advanced, and low Sun angles produce excessive shadow by early September. Instead, we used a longer year range (1999–2015) to construct the late August seasonal composite. To capture typical phenological conditions and minimize the influence of abnormally early or late seasonal changes in the final composites, we selected the observation with the median NIR reflectance value for each pixel for all seasonal windows except early June. Snow is usually widespread at that time, and both snow and green tundra have high NIR reflectance, so we selected the observation with the 20th percentile of NIR reflectance instead of the median.

Finally, we developed a spring snow-free reflectance composite to capture the temporally-variable period after snowmelt, but before green-up. We reasoned that a composite of snow-free, “brown” tundra would represent a brief, but important phenological window for predicting the cover of evergreen and/or low-growing PFTs that become increasingly overtopped later in summer. For the spring snow-free composite, we first constructed four bimonthly composites (late May–early July) selecting for the 20th percentile of NIR reflectance. Since snow and green tundra both tend to have high NIR reflectance, selecting for the 20th percentile tended to avoid both snow-covered and greened-up observations. From these four values, we selected the observation with the lowest NIR reflectance for each pixel to produce the spring snow-free composite.

For each seasonal composite, we calculated six spectral indices: Normalized Difference Vegetation Index (NDVI), Normalized Difference Moisture Index (NDMI), Normalized Difference Wetness Index (NDWI), Normalized Difference Snow Index (NDSI), Normalized Burn Ratio (NBR) and Enhanced

Vegetation Index-2 (EVI2) (Table 3). These commonly-used indices maximize the information content available for characterizing vegetation, hydrology and geomorphology; although there is some redundancy among the indices, their use increased the dynamic range available for characterizing ground conditions. For example, EVI2 can provide useful information for plant communities with high biomass (e.g., riparian shrublands) because EVI2 is less subject to saturation than NDVI. For NDVI, we also produced an “NDVI change” composite representing the difference between midsummer and spring snow-free NDVI; we reasoned that the magnitude of NDVI increase from snowmelt to the peak of the growing season would help to distinguish PFTs in highly productive communities.

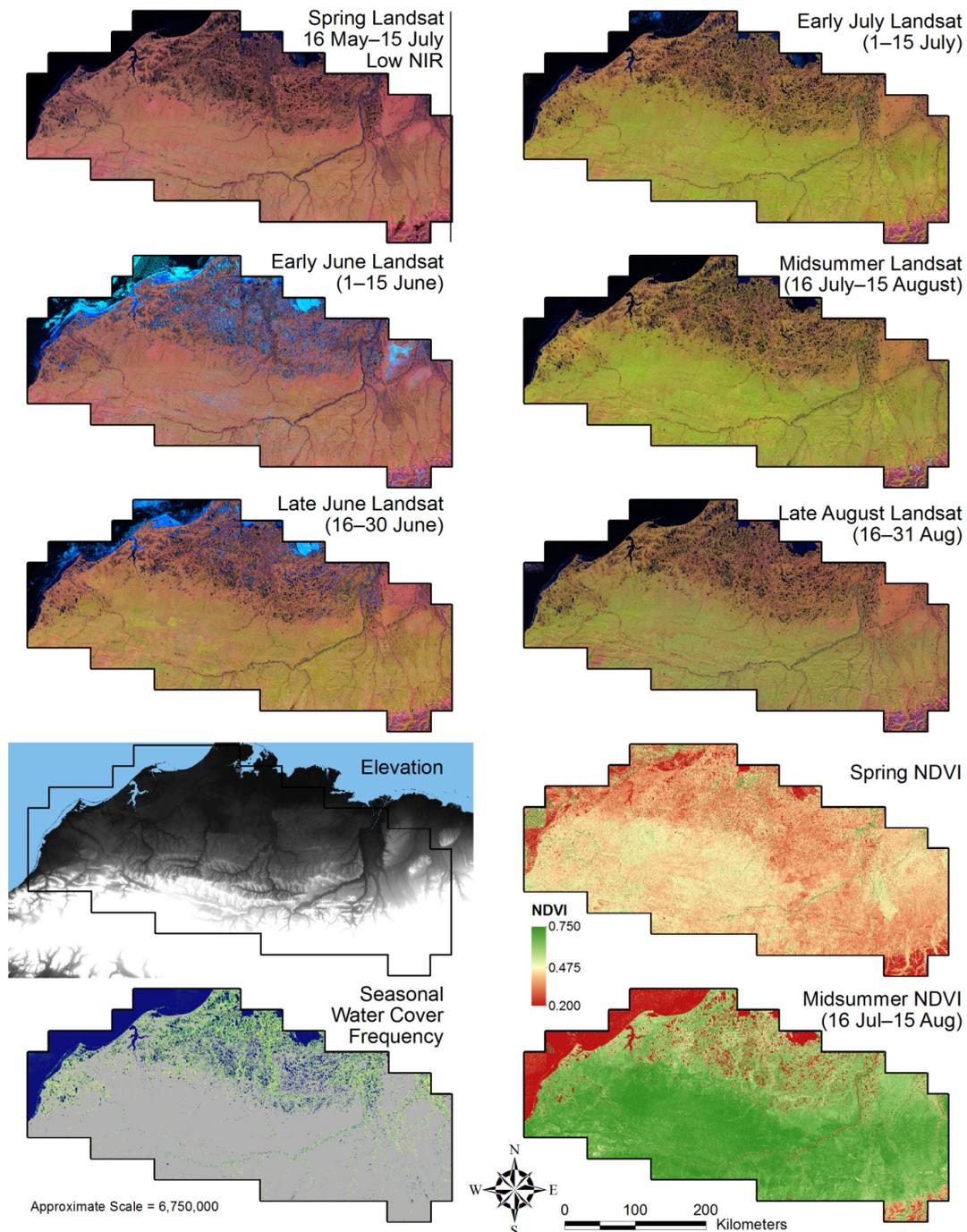


Figure 4. Selected spectral and ancillary predictors used to predict cover of PFTs and other cover types, North Slope, Alaska.

2.2.3. Environmental Predictors

We supplemented the spectral predictors with 5 gridded datasets representing climatic, phenological and topographic properties (Figure 4, Table 2). A climatological Summer Warmth Index (SWI; the sum of mean monthly temperatures $>0^{\circ}\text{C}$) was based on Advanced Very-High Resolution Radiometer (AVHRR) observations of land surface temperature for 1982–2003 [43]. Although this dataset is dated, it captured the pronounced summer temperature gradients of the North Slope. The climatological snow-free date was based on Landsat data for 1985–2011 [44] and represented regional- and landscape-scale gradients of snowmelt timing. While there are likely trends in summer warmth and snow persistence since these datasets were developed, we do not expect that the overall structure of these gradients has changed much; and it was not feasible to recalculate these metrics using more current imagery. Elevation (m) and slope (degrees) were calculated from a 30-m resolution digital terrain model (DTM) for the study area. The frequency of surface water inundation during the growing season was also calculated based on a Landsat time-series analysis; see Appendix A for details.

2.3. Quantitative Cover Modeling

To associate the training data with the spectral and environmental predictors, we overlaid the plot center on the predictors in a Geographic Information System (GIS). Each plot center was buffered by the plot diameter minus 10 m, which provided a representative sample of 2–12 Landsat pixels for each plot while avoiding edge pixels that were mainly influenced by properties outside of the plot. We converted each buffer to a 30-m raster and calculated the mean value of spectral and environmental predictors for each plot.

Table 2. Spectral and environmental predictors for random forest models of tundra PFTs and other cover types, North Slope, Alaska. SR is surface reflectance.

Predictor	Spectral Bands	Spectral Indices	<i>n</i>	Data Source or Reference
Spring snow-free SR: lowest NIR of four periods, 16 May–15 July 2010–2015	R, G, B, NIR, SWIR1/2	NDVI, NBR, NDMI, NDWI, NDSI, EVI2	12	Landsat TM/ETM+/OLI
Early June 20th percentile NIR SR: 1–15 June 2010–2015	R, G, B, NIR, SWIR1/2	NDVI, NBR, NDMI, NDWI, NDSI, EVI2	12	Landsat TM/ETM+/OLI
Late June median NIR SR: 16–30 June 2010–2015	R, G, B, NIR, SWIR1/2	NDVI, NBR, NDMI, NDWI, NDSI, EVI2	12	Landsat TM/ETM+/OLI
Early July median NIR SR: 1–15 July 2010–2015	R, G, B, NIR, SWIR1/2	NDVI, NBR, NDMI, NDWI, NDSI, EVI2	12	Landsat TM/ETM+/OLI
Midsummer median NIR SR: 16 July–15 August 2010–2015	R, G, B, NIR, SWIR1/2	NDVI, NBR, NDMI, NDWI, NDSI, EVI2	12	Landsat TM/ETM+/OLI
Late August median NIR SR: 16–31 August 1999–2015	R, G, B, NIR, SWIR1/2	NDVI, NBR, NDMI, NDWI, NDSI, EVI2	12	Landsat TM/ETM+/OLI
NDVI change, spring snow-free to midsummer	n/a	NDVI	1	Landsat TM/ETM+/OLI
Median snow-free date	n/a	n/a	1	[44]
Elevation	n/a	n/a	1	IFSAR ¹ DEM
Slope angle	n/a	n/a	1	IFSAR ¹ DEM
Summer Warmth Index	n/a	n/a	1	[43]
Seasonal water frequency (%)	n/a	n/a	1	Landsat TM/ETM+/OLI

1. Interferometric Synthetic Aperture Radar.

Table 3. Summary of spectral indices used for cover modeling, North Slope, Alaska.

Index	Formula	Reference
Normalized Difference Vegetation Index (NDVI)	$(\text{NIR} - \text{Red})/(\text{NIR} + \text{Red})$	[45,46]
Enhanced Vegetation Index-2 (EVI2)	$(\text{Red} - \text{Green})/(\text{Red} + [2.4 \times \text{Green}] + 1)$	[47]
Normalized Difference Water Index (NDWI)	$(\text{Green} - \text{NIR})/(\text{Green} + \text{NIR})$	[48]
Normalized Difference Moisture Index (NDMI)	$(\text{NIR} - \text{SWIR1})/(\text{NIR} + \text{SWIR1})$	[49]
Normalized Difference Snow Index (NDSI)	$(\text{Green} - \text{SWIR1})/(\text{Green} + \text{SWIR1})$	[50]
Normalized Burn Ratio (NBR)	$(\text{NIR} - \text{SWIR2})/(\text{NIR} + \text{SWIR2})$	[51]

2.3.1. Cover Modeling

We trained regression models with the randomForest algorithm in R [52] using all 78 predictors and mapped the cover of 18 response variables: 15 PFTs (including 6 aggregate PFTs), litter, open water and bare soil. Each of the response variables was modeled independently. We chose the random forest data-mining technique because it can use a large pool of correlated predictors where a priori relationships between predictor and response variables are unknown and interactions between predictors are likely to be nonlinear [35]. Random forest accomplishes this by repeating a bootstrapping process for each response variable in which a random sample of training plots is recursively split into smaller and smaller groups based on the spectral and environmental predictors. The predictors used at each split are also randomly selected. Random forest also provides internal cross-validation metrics for model performance.

We modeled the total cover for all 15 PFTs, and the top cover for vascular plants, nonvascular plants, litter, water and bare soil. For each PFT and cover metric, we produced four separate models that used a different number of predictors: (1) a full model that used all 78 predictors; (2) a manually-trimmed model that only included the most important predictors identified following a subjective predictor-importance assessment (see below); (3) an automatically-trimmed (auto-trimmed) model that excluded predictor(s) with importance value(s) <0 (i.e., predictors that caused model performance to decrease); and (4) a midsummer model in which all spectral predictors came from the midsummer seasonal window (16 July–15 August) that is traditionally used in remote sensing of Arctic vegetation. We supplemented the full models (i.e., using all predictors) with the manually and auto-trimmed models because although the full models are simplest to produce without requiring an additional predictor-importance assessment, the supplemental models provide insights into the key predictors and biophysical mechanisms by which specific PFTs can be distinguished using remote sensing. They also provide information on whether model performance is improved in our application by including a parsimonious set of important predictors (with some correlated or low information variables removed), as it has improved performance in another study [53].

2.3.2. Assessment of Predictor Importance

We performed a predictor-importance assessment for each PFT using conditional random forest (cforest), an independent data-mining technique related to random forest [54–56]. Predictor-importance measures provide a context for understanding the biophysical basis by which the data-mining process distinguished PFTs using a large set of predictors and provide a quantitative basis for pruning models to include fewer, more important predictors. Although the random forest algorithm in R produces its own set of predictor-importance metrics, these have several known limitations; for example, random forest tends to be biased towards correlated predictors. Conditional random forests provide an alternative with unbiased estimates of variable importance. We implemented cforest in the “party” package of R statistical software [50]. The cforest implementation of variable importance is superior to that in the “randomForest” package [49] because it better accounts for intercorrelated variables, over-fitting and bias towards selecting predictors with higher number of states [55,56]. The cforest regression models were trained using the same set of 78 predictors and 18 response variables. These models were not suitable for the production of map models, however, because they were ~100-times more

computationally intensive and unfeasibly slow for our 125,000-km² mapping area. For each PFT, we ranked and plotted the 78 predictors by their importance and produced two reduced sets of predictors for the manually-trimmed and auto-trimmed models. Each manually-trimmed model was produced by excluding variables below a breakpoint where predictor importance greatly decreased, and auto-trimmed models were produced by excluding variables with negative variable importance values (i.e., variables that decreased model performance, based on the “mean decrease in accuracy” importance scores).

2.3.3. Model Validation

We evaluated all cover models (1) using internal cross-validation from the training dataset (i.e., based on “out-of-bag” predictions); and (2) by comparing the cross-validation results to an “independent” validation dataset developed by reserving 20% of the field plots as validation samples; samples reserved for validation were randomly selected for each randomForest model run. Since the results from the “independent” validation varied for each model run, we ran 100 models for each PFT using randomForest (with a different 80%/20% partitioning in each “independent” run) and calculated 3 model performance metrics for each run: R^2 , root mean squared error (RMSE) and mean absolute error (MAE). To assess performance and uncertainty for each cover model, we calculated the ensemble mean and standard deviation of each performance metric across all 100 model runs for each PFT.

2.3.4. Comparison with Other Vegetation Maps

We also compared the results of the PFT models to two existing vegetation maps with comparable 30-m resolution: (1) the NSSI Land Cover map [5]; and (2) a previous continuous-field map of shrub cover by Beck et al. (2011) [29]. The NSSI Land Cover map is a thematic map covering our entire mapping area; it classified Landsat pixels according to a land cover classification system that emphasizes the dominant PFT(s) within each class (e.g., birch ericaceous low shrub, wet sedge). For the NSSI map comparison, we calculated the mean and standard deviation of our predicted total cover values for select PFTs within each NSSI land cover class. The Beck et al. (2011) effort was more similar to ours, and presents continuous-field maps for two shrub classes that were also developed using random forest: (1) “total shrub”, the proportional cover of all shrubs; and (2) “tall shrub”, the proportional cover of shrubs >1-m height. The maps of Beck et al. (2011) overlap most of our mapping area; we compared the low and tall deciduous shrub PFT and the total shrub PFT (our study) to the “all shrubs” map of Beck et al. (2011).

3. Results

3.1. Cover Modeling

Final PFT cover maps in raster format are hosted in the North Slope Science Catalog at <http://catalog.northslope.org/catalog/entries/8595>. The field measurements and resultant PFT cover maps indicated moderate to high cover values for low deciduous shrubs, dwarf evergreen shrubs, sedges, bryophytes and lichens for most of the mapping area, while tall deciduous shrubs, dwarf deciduous shrubs, grasses and forbs typically occurred at very low cover values except in specific landscape settings (Table 4). For the presentation of map results, we focus here on the total cover values developed using the full set of 78 predictors for dwarf evergreen shrub, and the aggregate PFTs Total shrub, low and tall deciduous shrub, total herbaceous, and total nonvascular (Figures 5–9). This combination of individual and aggregate PFTs collectively portrays the distribution and abundance of all nine individual PFTs present in the mapping area. Full results for all other PFTs and cover types are provided as the Supplemental Materials (Supplemental Figures S1–S3).

Across the full mapping area, the predicted total shrub cover averaged 46.8% for land pixels (i.e., excluding pixels mapped as water or ice in the NSSI map [5]) (Figure 5). Total shrub cover was higher in the foothills (mean total cover 57.3%; SD 15.9) than in the coastal plain (30.4%; SD 21.4). Within both physiographic provinces, there was a general decline in total shrub cover with increasing proximity to the Chukchi Sea coast. At landscape scales, the 30-m resolution cover maps display a variety of strong contrasts in total shrub cover across the mapping area, with the highest cover values found along hillslope water tracks, streambanks and floodplains, as well as at a variety of disturbance features. The predicted total cover of low and tall deciduous shrubs—those deciduous species with height greater than 0.2 m—declined abruptly across the physiographic transition from foothills to coastal plain (Figure 6). This PFT occurred at moderate cover values throughout most of the foothills, with the highest values predicted on floodplains, river bluffs, and well-drained uplands in the southernmost part of the study area (mean total cover 22.2%; SD 10.6). Low and tall deciduous shrubs are generally scarce on the coastal plain except on the Colville River floodplain (mean total cover 9.5%; SD 7.4). The predicted total cover of dwarf evergreen shrub averaged 27.0% (Figure 7); as for the preceding shrub PFTs, cover tended to be higher in the foothills (32.9%; SD 12.7) than on the coastal plain (17.7%; SD 12.2). Dwarf evergreen shrubs were scarcest within the footprint of the 2007 Anaktuvuk River fire, as well as in wetter parts of the coastal plain. In contrast to shrub PFTs, herbaceous and nonvascular PFTs tended to be more evenly distributed throughout the mapping area. The total herbaceous aggregate PFT included sedges, grasses and forbs; total herbaceous cover averaged slightly higher in the foothills (mean cover 36.7%; SD 6.7), but was comparable in a variety of landscape positions on the coastal plain, such as drained lake basins and meandering floodplains (32.3%; SD 8.7) (Figure 8). The predicted total cover of nonvascular plants was very similar between the foothills (53.8%; SD 9.5) and the coastal plain (49.4%; SD 17.6) (Figure 9).

Table 4. Model performance metrics for cover models using full set of 78 predictors, based on internal cross-validation, North Slope, Alaska.

Plant Functional Type	Cover Metric	Mean Cover at Plots (%)	R ²	RMSE (% Cover)	MAE (% Cover)
Tall deciduous shrub	Total	2.2	0.49	5.7	2.0
Low deciduous shrub	Total	11.6	0.66	10.1	6.4
Low & tall deciduous shrub	Total	13.8	0.75	9.6	6.3
Dwarf deciduous shrub	Total	2.5	0.25	3.6	2.2
Total deciduous shrub	Total	16.3	0.75	10.0	6.8
Dwarf evergreen shrub	Total	15.4	0.70	10.2	6.7
Total shrub	Total	31.7	0.80	13.3	9.6
Sedges	Total	19.6	0.40	14.4	10.2
Grasses	Total	2.6	0.13	8.1	3.4
Forbs	Total	5.9	0.42	7.1	4.5
Total herbaceous	Total	28.1	0.50	15.1	10.8
Total vascular plants	Total	59.8	0.81	17.4	13.0
Bryophytes	Total	32.5	0.51	18.1	13.7
Lichens	Total	6.7	0.55	6.8	4.3
Total nonvascular plants	Total	39.6	0.61	18.4	13.6
Total vascular plants	Top	39.6	0.75	12.0	9.2
Total nonvascular plants	Top	10.0	0.36	7.7	5.5
Litter	Top	24.9	0.55	11.7	8.9
Bare ground	Top	10.6	0.88	8.5	4.5
Water	Top	14.8	0.91	8.9	4.4



Figure 5. Map of total shrub cover based on the random forest model with 78 predictors, North Slope, Alaska. The inset map displays landscape-scale patterns in total shrub distribution in the upland terrain south of Umiat.

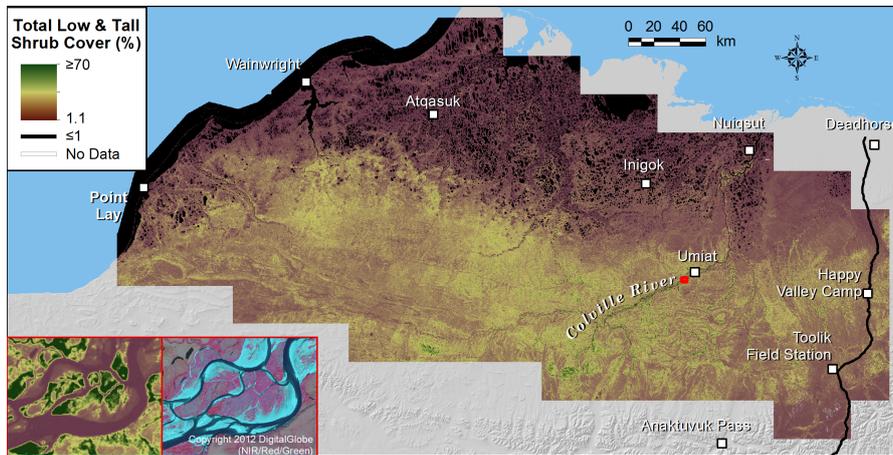


Figure 6. Map of total low and tall deciduous shrub cover based on the random forest model with 78 predictors, North Slope, Alaska. The inset map displays examples of highly productive shrublands on the active floodplain of the Colville River.

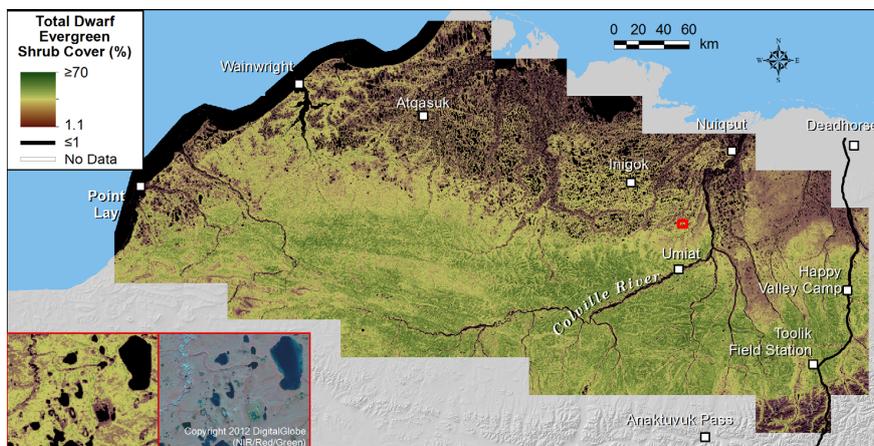


Figure 7. Map of dwarf evergreen shrub cover based on the random forest model with 78 predictors, North Slope, Alaska. The inset map shows an example of dwarf evergreen shrub distribution on the eolian sand sheet region of the coastal plain, where this PFT is common, except in wet areas.

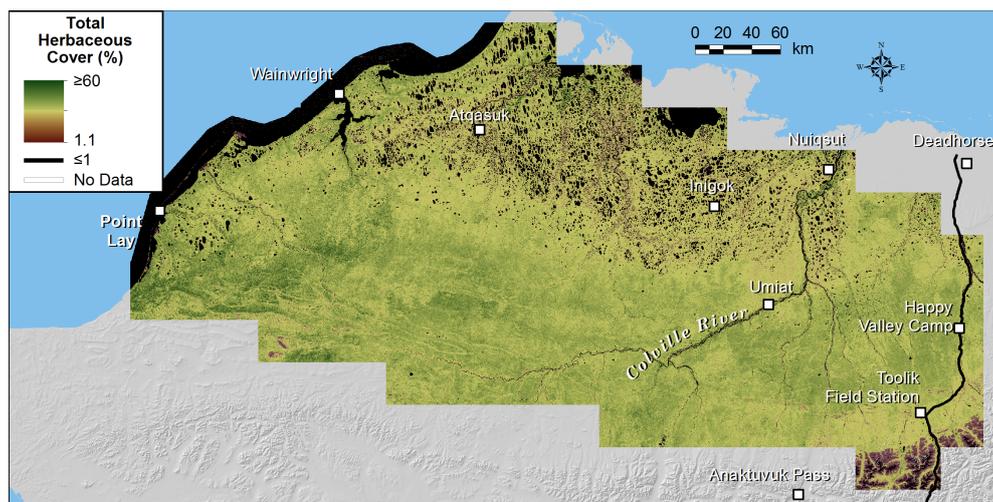


Figure 8. Map of total herbaceous cover based on the random forest model with 78 predictors, North Slope, Alaska.

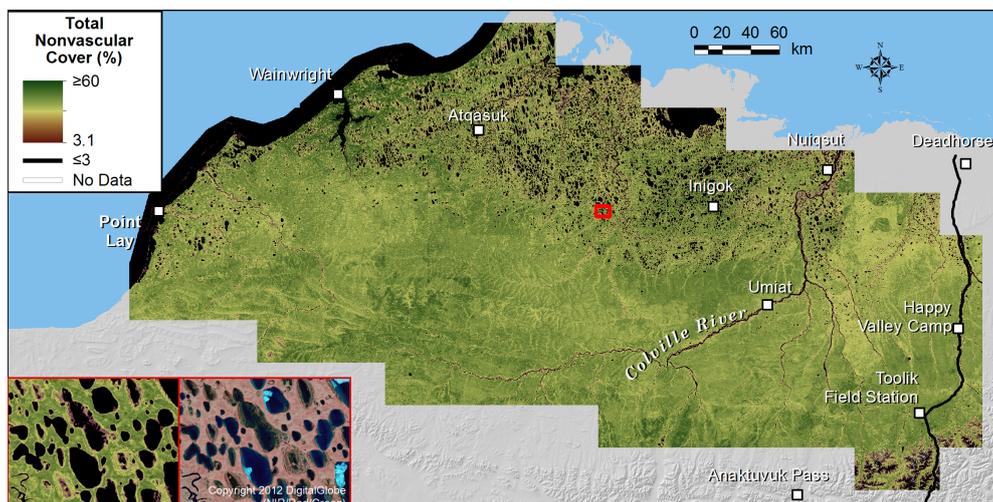


Figure 9. Map of total nonvascular cover based on the random forest model with 78 predictors, North Slope, Alaska. The inset map portrays a typical lake-rich portion of the coastal plain; total nonvascular cover is generally high except in flooded lake margins and drained lake basins.

3.2. Assessment of Predictor Importance

Most ranked predictor importance plots had an exponential shape, indicating that a small number of predictors had moderate to high importance, and many had low importance (Figure 10). For all PFTs and cover types except open water, the most important predictors were spectral predictors from the Landsat seasonal composites rather than environmental predictors (Supplementary Table S1). For deciduous shrub PFTs, most of the top predictors came from the early July seasonal composite. In contrast, the most important predictors for Dwarf evergreen shrubs, lichens and total nonvascular came from the spring snow-free (“brown tundra”) composite. The most important predictors for the herbaceous PFTs did not exhibit any strong patterns and came from various seasonal windows. Environmental predictors seldom had high importance values, except for open water for which seasonal water frequency was the most important predictor.

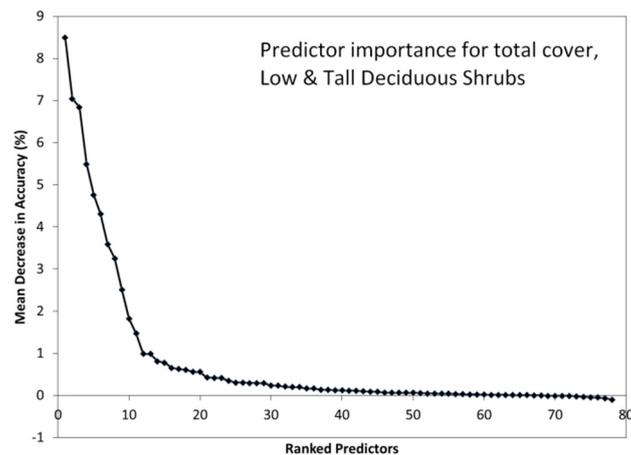


Figure 10. Plot of unbiased variable importance determined by conditional random forest, by predictor rank for total cover of low and tall shrubs, North Slope, Alaska.

The most important spectral predictors for most PFTs were indices rather than surface reflectance values for the six individual Landsat bands. EVI2 and NDVI were by far the most common indices with high importance values and were particularly important for shrub PFTs. NBR and NDSI were infrequently identified as top predictors; however, NDSI was useful for modeling lichens and litter. NDWI and NDMI were rarely selected as important predictors. Individual Landsat bands were seldom selected as important predictors except for bare ground.

3.3. Model Validation

We ran bootstrap simulations 100 times for each model and summarized the model performance metrics (Table 5). Mean model performance metrics derived from internal cross-validation (i.e., based on “out-of-bag” predictions) and “independent” validation data from randomly selected 80%/20% training/validation splits were generally similar (Table 5). However, the variability in the performance metrics was much higher for the independent runs compared to the internal cross-validation, with the standard deviation of the performance metrics at least three-times greater across 100 runs. The random variability in the 20% of training data selected for the validation of each independent model run is large enough that the results for the performance metrics can vary substantially depending on the composition of a particular reserved sample. This is not a novel finding, but it does demonstrate that internal cross-validation provides a more stable estimate of model performance than withholding 20% of training data. Therefore, we relied on the internal cross-validation metrics for model comparison and evaluation. For computational reasons, our actual PFT cover predictions for millions of pixels (rather than 225 plots) were generated from single random forest model runs, not 100 runs.

In all cases, the auto-trimmed or manually-trimmed models performed slightly better than full models that used all 78 predictors (Table 6). The auto-trimmed models had 60–73 predictors, compared to 9–15 predictors for the manually-trimmed models and 78 predictors for the full models. The auto-trimmed models performed best for dwarf evergreen shrub and total shrub, while manually-trimmed models performed best for the other three PFTs.

The improvements for manual and auto-trimmed models compared to the full models were very small on average, for example 0.008–0.009 for R^2 and 0.02–0.05% for MAE (Table 7). Averaged across all models, the manual trimmed model had the highest R^2 , while the auto-trimmed models had slightly lower MAE and RMSE. The differences among the full, manual and trimmed models were small enough that the additional work of developing parsimonious models does not appear to be worthwhile as a general practice. The midsummer models had substantially lower performance than the full and trimmed models, however, with R^2 0.08 lower, MAE 0.8% higher and RMSE 1.2% higher on average.

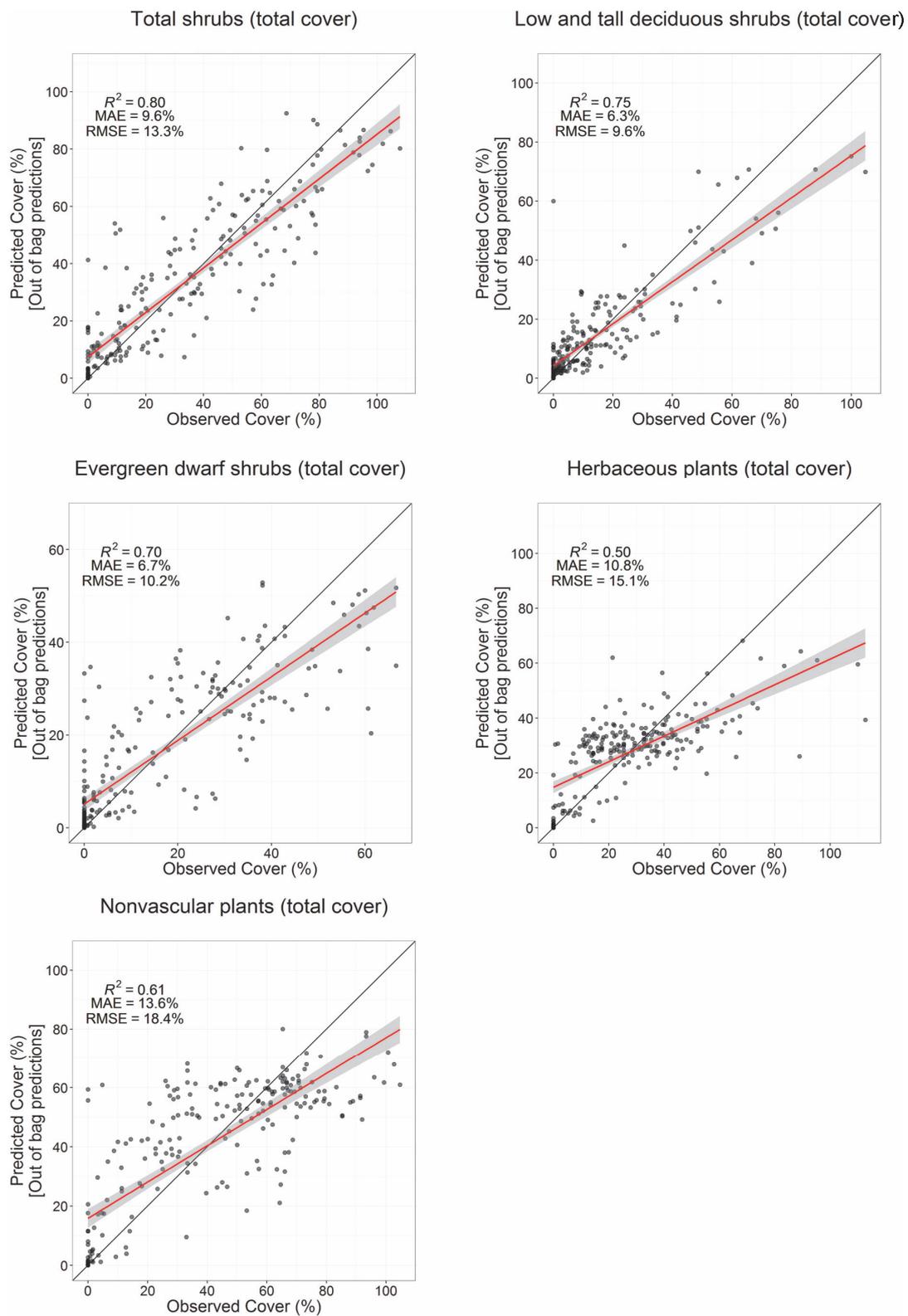


Figure 11. Scatterplots comparing observed and predicted cover values from training and validation plots for selected PFTs using the full set of 78 predictors. Predicted values for training data are “out-of-bag” predictions. The red lines are linear fits with 95% confidence intervals (grey).

Table 5. Ensemble performance metrics for cover models using full set of 78 predictors, based on cross-validation with training data and independent validation data, North Slope, Alaska.

PFT or Cover Type	Cover Metric	R^2		RMSE (% Cover)		MAE (% Cover)	
		Internal Cross-Validation ^a	Independent Validation ^b	Internal Cross-Validation ^a	Independent Validation ^b	Internal Cross-Validation ^a	Independent Validation ^b
Tall deciduous shrub	Total	0.49 ± 0.07	0.56 ± 0.20	5.8 ± 0.9	5.2 ± 2.5	2.2 ± 0.2	2.2 ± 0.7
Low deciduous shrub	Total	0.62 ± 0.04	0.65 ± 0.13	10.5 ± 0.5	10.4 ± 1.9	6.7 ± 0.3	6.6 ± 1.1
Low and tall deciduous shrub	Total	0.74 ± 0.04	0.74 ± 0.12	9.8 ± 0.6	9.6 ± 1.8	6.4 ± 0.2	6.3 ± 0.9
Dwarf deciduous shrub	Total	0.23 ± 0.05	0.24 ± 0.11	3.6 ± 0.2	3.6 ± 0.6	2.3 ± 0.1	2.3 ± 0.3
Total deciduous shrub	Total	0.74 ± 0.03	0.73 ± 0.12	10.3 ± 0.5	10.4 ± 1.8	7.0 ± 0.3	7.1 ± 1.0
Dwarf evergreen shrub	Total	0.69 ± 0.02	0.70 ± 0.07	10.4 ± 0.4	10.1 ± 1.4	6.9 ± 0.3	6.8 ± 1.0
Total shrub	Total	0.79 ± 0.02	0.79 ± 0.06	13.5 ± 0.5	13.2 ± 1.8	9.8 ± 0.4	9.7 ± 1.3
Sedges	Total	0.38 ± 0.03	0.38 ± 0.10	14.6 ± 0.6	14.6 ± 2.2	10.4 ± 0.4	10.3 ± 1.3
Grasses	Total	0.14 ± 0.05	0.21 ± 0.18	8.1 ± 1.1	7.5 ± 3.2	3.5 ± 0.5	3.5 ± 0.9
Forbs	Total	0.39 ± 0.05	0.43 ± 0.16	7.3 ± 0.4	7.2 ± 1.5	4.6 ± 0.3	4.7 ± 0.8
Total herbaceous	Total	0.48 ± 0.03	0.49 ± 0.09	15.5 ± 0.6	15.0 ± 2.4	11.2 ± 0.4	10.9 ± 1.5
Total vascular plants	Total	0.81 ± 0.01	0.80 ± 0.05	17.4 ± 0.6	17.2 ± 1.9	12.9 ± 0.5	13 ± 1.6
Bryophytes	Total	0.49 ± 0.03	0.48 ± 0.09	18.4 ± 0.5	18.1 ± 1.6	14.0 ± 0.5	13.9 ± 1.5
Lichens	Total	0.54 ± 0.03	0.54 ± 0.11	6.9 ± 0.4	6.9 ± 1.3	4.4 ± 0.2	4.5 ± 0.7
Total nonvascular plants	Total	0.60 ± 0.02	0.61 ± 0.08	18.6 ± 0.5	18.1 ± 1.7	13.9 ± 0.5	13.6 ± 1.5
Total vascular plants	Top	0.75 ± 0.02	0.76 ± 0.05	12.0 ± 0.3	11.8 ± 1.2	9.2 ± 0.3	9.1 ± 1.0
Total nonvascular plants	Top	0.35 ± 0.04	0.35 ± 0.10	7.7 ± 0.4	7.7 ± 1.1	5.5 ± 0.3	5.5 ± 0.7
Litter	Top	0.52 ± 0.02	0.50 ± 0.09	12.0 ± 0.3	12.2 ± 1.2	9.1 ± 0.3	9.4 ± 1.1
Bare ground	Top	0.88 ± 0.01	0.86 ± 0.08	8.6 ± 0.5	8.8 ± 2.2	4.6 ± 0.3	4.8 ± 1.1
Water	Top	0.91 ± 0.01	0.88 ± 0.09	9.3 ± 0.7	9.3 ± 2.3	4.6 ± 0.3	4.7 ± 1.1

^a Data on model performance were based on 100 runs using internal cross-validation (out-of-bag predictions). ^b Data on model performance were based on a validation subset from 100 runs using independent 80%/20% training/validation splits.

Table 6. Model performance metrics for selected cover models by plant functional type (PFT) at different levels of model parsimony, based on cross-validation, North Slope, Alaska. Bold numbers indicate the highest value for each performance metric within each PFT.

Plant Functional Type	Model Parsimony	Number of Predictors	R^2	RMSE (% Cover)	MAE (% Cover)
Dwarf evergreen shrub	Full	78	0.70	6.7	10.2
Dwarf evergreen shrub	Auto trimmed	60	0.71	6.6	10.0
Dwarf evergreen shrub	Manual trimmed	10	0.68	7.0	10.5
Dwarf evergreen shrub	Midsummer	14	0.63	7.9	11.4
Low and tall deciduous shrub	Full	78	0.75	6.3	9.6
Low and tall deciduous shrub	Auto trimmed	73	0.76	6.3	9.5
Low and tall deciduous shrub	Manual trimmed	9	0.76	6.4	9.5
Low and tall deciduous shrub	Midsummer	14	0.66	7.0	11.2
Total shrub	Full	78	0.80	9.6	13.3
Total shrub	Auto trimmed	67	0.80	9.5	13.1
Total shrub	Manual trimmed	11	0.78	10.0	13.9
Total shrub	Midsummer	14	0.72	11.2	15.5
Total herbaceous	Full	78	0.50	10.8	15.1
Total herbaceous	Auto trimmed	63	0.51	10.8	15.0
Total herbaceous	Manual trimmed	15	0.52	10.7	14.8
Total herbaceous	Midsummer	14	0.39	11.8	16.7
Total nonvascular	Full	78	0.61	13.6	18.4
Total nonvascular	Auto trimmed	60	0.63	13.4	18.0
Total nonvascular	Manual trimmed	14	0.64	13.1	17.8
Total nonvascular	Midsummer	14	0.52	15.1	20.5

Table 7. Mean model performance by model parsimony for all PFTs and other cover types, North Slope, Alaska.

Model Parsimony	Mean R^2	Mean MAE (%)	Mean RMSE (%)
Full	0.589	7.31	10.83
Auto trimmed	0.597	7.26	10.75
Manual trimmed	0.598	7.29	10.80
Midsummer	0.505	8.08	12.01

Scatterplots comparing the observed and out-of-bag predictions for selected PFTs (Figure 11) show that there is a tendency towards overprediction at the lowest observed cover values and underprediction at higher observed cover values. This is most pronounced for herbaceous and nonvascular models, which have lower overall performance based on the other metrics. The effect is more modest for the shrub PFTs, which represent the dominant canopy-forming vegetation and are better modeled by spectral predictors. Future modeling efforts should focus on identifying predictors or modeling approaches to improve the representation of low and high cover values for each PFT. A linear adjustment to the estimates is one approach that has been applied to reduce the bias of cover estimates in the taiga-tundra ecotone [57].

We focused on full models because the improvements in model performance from trimmed models were minor and because the use of full models is simpler to implement for updating the PFT cover maps in the future. Model performance varied widely across PFTs and was generally highest for shrub PFTs (Figure 11, Table 4). Based on cross-validation with out-of-bag samples, the overall fit of full models for the total cover of individual PFTs was best for dwarf evergreen shrub ($R^2 = 0.70$, RMSE = 10.2%) and low deciduous shrub ($R^2 = 0.66$, RMSE = 10.1%). Overall model fit was intermediate for lichen ($R^2 = 0.55$, RMSE = 6.8%), bryophyte ($R^2 = 0.51$, RMSE = 18.1%) and tall deciduous shrub ($R^2 = 0.49$, RMSE = 5.7%) and generally poor for the three herbaceous PFTs (R^2 range 0.13–0.42). Model fit increased substantially, however, when similar PFTs were aggregated, such as low and tall deciduous shrub ($R^2 = 0.75$), total shrub ($R^2 = 0.80$), total vascular ($R^2 = 0.81$) and total nonvascular ($R^2 = 0.61$). For top cover, model performance was very good for water ($R^2 = 0.91$, RMSE = 8.9%) and bare ground ($R^2 = 0.88$, RMSE = 8.5%) and intermediate for litter ($R^2 = 0.55$, RMSE = 11.7%).

Among shrub PFTs, cross-validated MAE was lowest for tall deciduous shrubs (2.0%) and dwarf deciduous shrubs (2.2%) (Table 4); however, the R^2 values were relatively low at 0.49 and 0.25, respectively. Both these PFTs had low abundance overall, with the mean total cover at field plots of 2.2% and 2.5%, respectively, so the MAE was of similar magnitude to the mean total cover, and RMSE was greater than the mean cover. The other shrub PFTs and aggregated shrub PFTs had much higher cover ($\geq 11.6\%$) based on field data, and though the MAE and RMSE errors are higher in absolute magnitude, they are lower as a proportion of the mean cover observed in the field. For example, the Low and Tall Deciduous Shrub PFT had a mean cover of 13.8%, and the modeled cover had MAE of 6.4% and RMSE of 9.6%.

3.4. Comparison with Other Vegetation Maps

Beck et al. [29] produced a ca. 2000 baseline map of shrubs for the North Slope of Alaska in which shrub cover was defined as the deciduous species “dwarf birch (*Betula nana*), alder (*Alnus viridis* ssp. *fruticosa*) and willow (*Salix* spp.).” This map was based on a random forest regression model trained by binary classifications of shrub presence/absence for 1–5-m pixels. For the portion of the map that overlapped our study area, shrub cover in the ca. 2000 baseline map averaged 62.5%, compared to 46.8% total shrub total cover in our map. Binned scatterplots comparing the Beck et al. [29] shrub cover to our total shrub cover and low and tall deciduous shrub cover (Figure 12) show that the Beck et al. map more closely represents our total shrub cover map across the full range of cover values. Though Beck et al. defined their shrub cover as primarily consisting of the three deciduous species described above, it is likely that they included extensive evergreen shrub cover in their photo-interpretation of shrub cover. The approach of defining shrub presence/absence for 1–5-m pixels also could fail to accurately resolve many open shrub canopies, which often have gradients of shrub cover. The laser pointer used in our field plots can accurately resolve shrub cover in open and mixed shrub canopies and very rarely resulted in observed total cover $\geq 100\%$ (Figure 11). Our models for shrub PFTs were more sensitive to a range of cover percentages and seldom reached 100% predicted cover, when measured cover was high, but less than 100%, while large portions of the Beck et al. map depicted 100% shrub cover. Based on our intensive field dataset and tundra vegetation sampling experience, $\geq 100\%$ shrub cover is rare on the landscape in the study area.

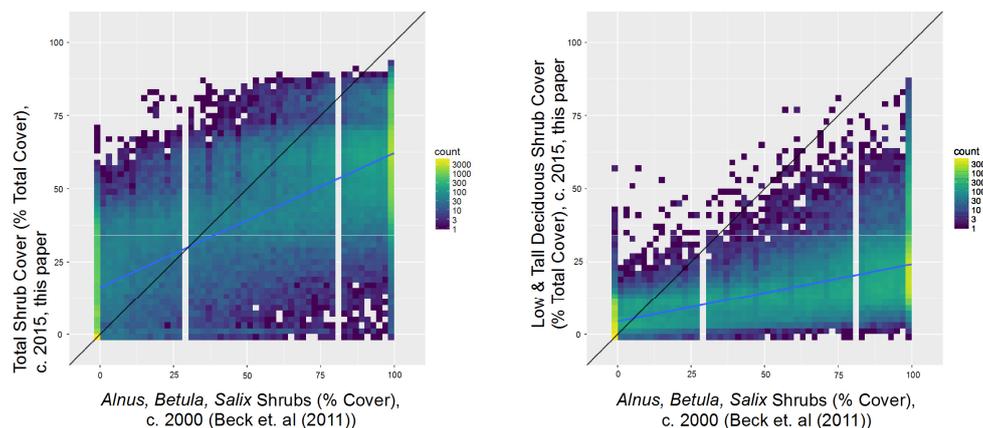


Figure 12. Comparison of shrub cover from Beck et al. [29] circa 2000 North Slope shrub map and the current work, North Slope, Alaska. The black line is 1:1 line, and the blue line is the linear best fit.

We summarized our results for low and tall deciduous shrubs and dwarf evergreen shrub cover for 17 widespread land cover classes represented in the NSSI Land Cover map [5] using box-and-whisker plots (Figure 13). These summaries provide support for the overall consistency of the two 30-m products and also highlight variability within NSSI classes. The PFT cover mapping provides detail that the categorical mapping does not.

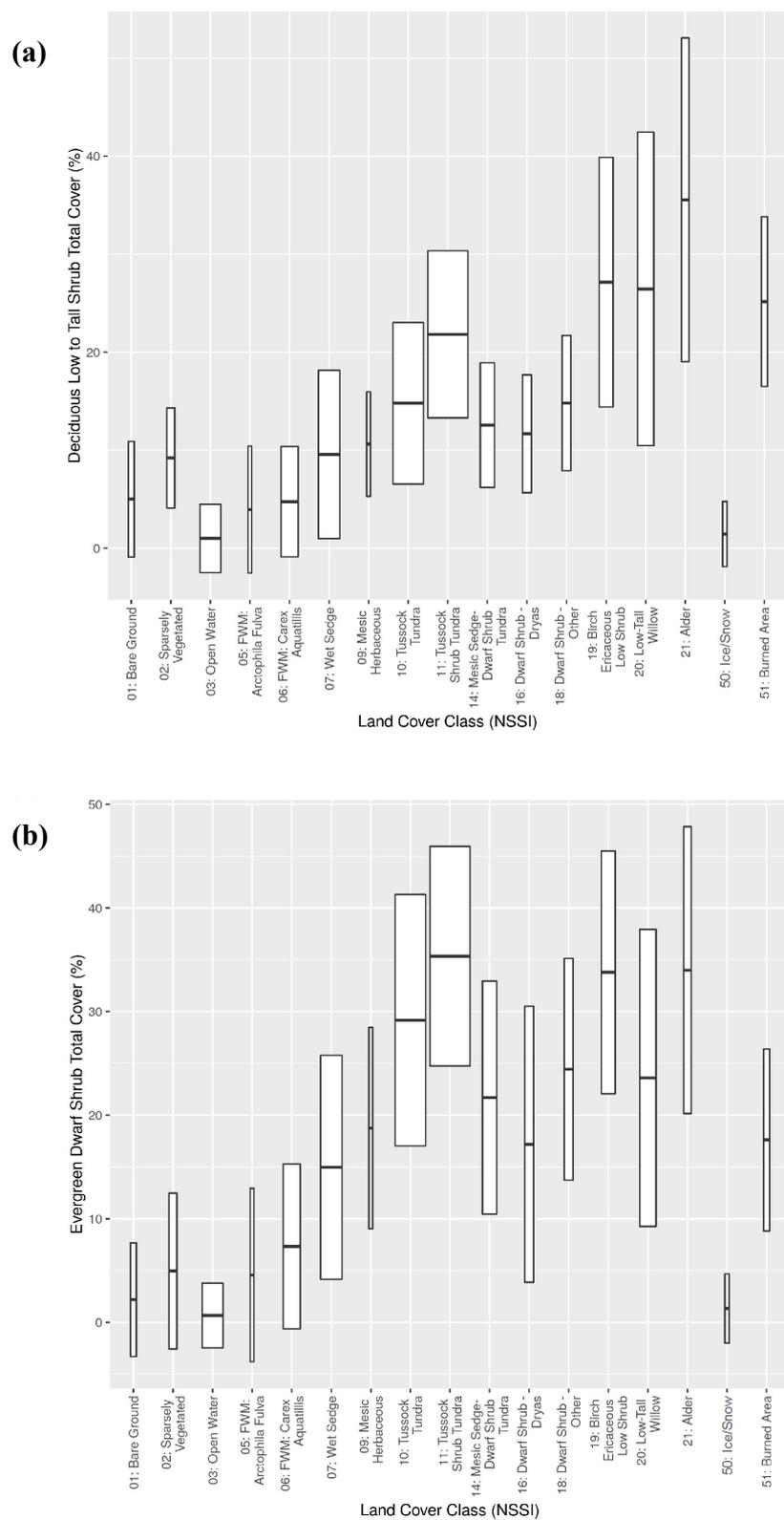


Figure 13. Box and whisker plots of (a) low and tall deciduous shrub cover and (b) dwarf evergreen shrub cover, summarized by the North Slope Science Initiative (NSSI) Land Cover [5] map class. The width of boxes is proportional to the square root of the area of each map class.

4. Discussion

4.1. Correspondence of PFTs and Environmental Gradients

The continuous-cover maps for 15 PFTs provide the most detailed landscape-scale assessments of vegetation cover available for Arctic Alaska and demonstrate rich spatial patterns at regional- and landscape scales that reflect environmental gradients in the mapping area. Many of these patterns are broadly consistent with previous field studies and thematic vegetation maps. At the regional scale, the total cover of most vascular PFTs reflects the coast-inland summer climate gradient, with the highest cover values predicted in warmer, southerly parts of the mapping area (Supplemental Figure S2).

The modeled total cover of low and tall deciduous shrubs (shrub height ≥ 0.2 m) is closely linked to the summer climate gradient (Figure 6), but the 30-m resolution of the map provides representation of landscape-scale variability. This aggregate PFT is most abundant in the foothills on major river floodplains, particularly the Colville River. Abrupt increases in the total cover of low and tall deciduous shrubs evident in the central and western foothills correspond to the footprints of two large, pre-20th century tundra fire scars, a pattern corroborated by field studies [58]. However, representation of deciduous shrub PFTs was somewhat limited by the poor performance of models for individual height classes. Model fits were lowest for dwarf deciduous shrubs; this suggests those erect PFTs, which are unlikely to be overtopped by other PFTs, are more readily modeled from spectral predictors. However, model fit for tall deciduous shrubs was much lower than for low deciduous shrubs. We believe this is because many of the dominant deciduous shrub species, such as diamond leaf willow (*Salix pulchra*) and Siberian alder (*Alnus viridis* ssp. *fruticosa*), have a high degree of phenotypic plasticity and assume different growth forms depending on local environmental conditions. Thus, floristically similar shrub communities have similar spectral properties regardless of canopy height. However, improved representation of tall deciduous shrubs is highly desired because of their recent expansion and associated influence on ecosystem properties in northern Alaska. Future efforts to distinguish and monitor this PFT could combine optical remote sensing predictors with canopy height models derived from emerging tools such as LiDAR and drone photogrammetry; recent work in the forest-tundra ecotone highlights the potential for the use of LiDAR in tandem with optical satellite data to map vegetation canopy properties [57].

The high model fit of dwarf evergreen shrubs was interesting, given the much lower performance for other low-statured PFTs, and it supports the value of developing seasonal reflectance composites. Many of the spatial patterns evident in the map of dwarf evergreen shrub total cover are corroborated by our own training data and independent field studies. For example, dwarf evergreen shrubs are the dominant shrub growth form on the well-drained, nutrient-poor soils that predominate in this area. Component species, such as *Ledum*, *Dryas* and *Vaccinium vitis-idaea* are widespread across Arctic Alaska, but their distribution and changes in their abundance have received far less attention than larger shrubs.

Spatial models for the aggregate PFT total herbaceous achieved good model fits, but our efforts to quantify individual herbaceous PFTs were generally unsuccessful. In most environments in northern Alaska, sedges are the predominant herbaceous PFT and are virtually ubiquitous in a variety of soil environments, ranging from hydrophilic species on the coastal plain to extensive *Eriophorum vaginatum* tussocks on mesic sites. Grasses and forbs are also common, but they seldom contribute much of the total cover except in specific landscape positions, such as active river floodplains. Nonetheless, the spatial models for total herbaceous cover provide improved representation of PFTs that have traditionally been neglected in thematic vegetation maps.

In contrast to cover models for shrub PFTs, total cover of nonvascular plants (Figure 9) is not closely linked to the summer climate gradient, with intermediate cover values predicted over most of the mapping area. Landscape-scale variability in predicted total nonvascular cover indicates that cover is not closely linked to aboveground biomass; modeled total nonvascular cover is similar in both high-biomass shrub-tussock tundra (red areas in color-infrared imagery) and in the drained lake

basin at upper left (brownier signature). We are unaware of previous efforts to map the total cover of bryophytes. This PFT (primarily mosses) is virtually ubiquitous in Arctic environments and forms nearly continuous live cover in many of them. For many applications; however, quantitative estimates of the thickness of mosses rather than total cover are highly desirable due to the strong insulative properties of moss and their influence on permafrost thermal regime [30]. Such estimates remain elusive; however, our spatial models of bryophyte cover provide a starting point for efforts to predict moss thickness, particularly in areas with high total cover. Lichen cover estimates are highly desirable for efforts to monitor winter forage for Arctic caribou herds [34].

4.2. Sources of Uncertainty

One potential source of uncertainty is a timing mismatch between the field data and the predictor datasets. We constrained this by minimizing the temporal disparity between the fieldwork (2012–2014) and the spectral predictor datasets, most of which were derived from Landsat data from 2010–2015. Though the year range of Landsat imagery was wider than the field data, the field data were acquired near the middle of the range. We determined that ~6 years of Landsat data were required to consistently characterize normal spectral/environmental conditions for most seasonal time windows, with more years required for the late summer window when cloud cover was more common.

Another source of uncertainty is the spatial variability of tundra vegetation compared to sample plots and Landsat pixels. It has been suggested that 10-m spatial resolution is more appropriate for characterizing tundra vegetation than 30-m Landsat pixels (e.g., [59]). However, the Landsat archive has the depth to provide consistent seasonal reflectance composites across large study areas, a capability not yet available for higher resolution sensors. Despite including all available Landsat data over a six year window, some minor artifacts are visible. This demonstrates the challenge of compositing Landsat data in general (e.g., scan-line artifacts associated with most Landsat 7 data), and compositing in tundra regions in particular, where seasonal changes are rapid, the growing season is short and cloud cover is frequent. Our use of six-year compositing periods minimized the occurrence of such artifacts, which would be more extensive with a shorter window. Eventually, data from the Sentinel family of satellites or other sensors may provide composites at higher spatial resolution.

The field training dataset was pooled from independent field efforts conducted by BLM and us. Both field efforts used very similar methods, but there were differences in the way plot locations were selected and the length of sampling lines. Approximately half of field plots (sampled by us) were subjectively located in areas of homogeneous vegetation (distributed across vegetation types), and sample lines were 50 m long; BLM field plots were allocated using a randomized design stratified by NSSI land cover classes, and sample lines were only 25 m long, with final locations also adjusted to locate plots within homogeneous vegetation patches. While our approach to plot locations was more subjective, we did strive to collect data in the variety of vegetation types that occurred in the study area, within logistical constraints. Spatial auto-correlation of sample plots is not expected to be an issue for model development and evaluation, because plots from both field efforts were spaced at least hundreds of meters apart and in different vegetation patches.

We related plant species to shrub height classes, but there can be substantial variability in shrub growth forms within a species [60]. For future work using height measurements from the field, LiDAR mapping and/or photogrammetry could provide more robust shrub height estimates to distinguish dwarf, low and tall shrub PFTs.

We did not apply a terrain illumination correction, though this might have provided local improvements to model results. Most of the study area is flat: >91% of the study had a terrain slope of $\leq 5^\circ$, which is a common threshold below which no correction is applied even when a terrain illumination correction is used elsewhere in a single study area. It would be worthwhile to evaluate a terrain illumination correction for similar modeling efforts in the future, particularly in regions that are less dominated by flat terrain.

4.3. Predictor Importance: Implications for Monitoring

Conditional random forest helped to determine predictor importance, though the auto- and manually-trimmed models provided very small improvements in model fit over full models and minimal computational benefits. The use of conditional random forest was useful in that it identified the predictors that are most important for modeling PFTs in our study area and provide guidance for remote sensing applications in other tundra regions. Variable importance can also suggest underlying biophysical mechanisms that likely explain the importance of specific predictors. For example, spectral predictors were of much greater importance than environmental predictors for most PFTs. For deciduous shrub PFTs, virtually all of the most important predictors were spectral predictors derived from the early July composite. At this time of year, deciduous shrubs are likely to dominate the spectral signal because they are fully leaf-out, whereas the photosynthetically-active biomass (phytomass) of sedges and other herbaceous PFTs is largely obscured by standing litter until later in the summer. In contrast, for dwarf evergreen shrubs, the most important spectral predictors came from the spring snow-free and early June composites; in the spring, evergreen PFTs are among the few plants that would contribute much photosynthetically-active biomass to the signal observed by satellites. The spring snow-free composite represents virtually all terrestrial portions of the mapping area in a snow-free condition, but before deciduous shrub and herbaceous PFTs have developed substantial green leaf biomass.

The use of seasonal composites greatly improved cover estimates for lower-statured PFTs such as evergreen dwarf shrubs and nonvascular plants compared to using a single, midsummer reflectance composite. The spectral signatures of dwarf shrubs and nonvascular plants are often masked by overtopping deciduous shrubs and herbaceous plants. Using reflectance values from the spring snow-free composite, which is representative of “brown” tundra that lacks snow and newly emerged green vegetation, effectively removes most of the overtopping deciduous vegetation from the reflectance signature, thereby emphasizing evergreen and nonvascular plants.

5. Conclusions

We applied a random forest decision tree data-mining technique to predict the cover of tundra plant functional types, litter, water and bare ground at 30-m resolution for an ~125,000-km² study area spanning Alaska’s western and central North Slope. For training data, we used field measurements of tundra vegetation and other cover types collected at 236 plots using a point-intercept sampling approach. Model predictors included surface reflectance values and spectral indices derived from Landsat seasonal composites spanning phenological events from snowmelt to senescence and a suite of quantitative environmental predictors that captured important environmental gradients in the study area.

Our simultaneous mapping of different PFTs will aid in detecting and monitoring both long-term (i.e., climate-driven) and short-term (i.e., disturbance-driven) changes in North Slope tundra. Both the field sampling protocol and the methods used to construct seasonal composites from the long-lived Landsat program are well-suited for periodic replication in future decades. Additionally, the cover models are highly versatile because of their 30-m resolution and support diverse applications from regional to landscape scales. Periodic revision of the PFT maps presented here will effectively leverage ongoing efforts by BLM and others to monitor North Slope ecosystems using the point-intercept sampling approach. These ongoing efforts will also contribute field training data for portions of the mapping area that are poorly represented in the current network of field plots. In the future, UAVs may provide a more efficient way to collect detailed vegetation cover and structure data at a scale (e.g., 0.5–5.0 cm) that matches the vegetation elements of low-stature tundra communities.

The quantitative cover maps significantly improve upon earlier efforts to map shrub functional types. Previous models built on binary vegetation cover data (presence/absence for 1–5-m pixels) overestimated cover when compared to our models, which were based on detailed field vegetation measurements. The use of seasonal composites greatly improved the detection of short-statured

PFTs (e.g., evergreen shrubs, nonvascular plants) compared to traditional methods using midsummer surface reflectance values. Spatial patterns in the model outputs conform to known relationships between tundra vegetation and environmental gradients of the North Slope, at regional and landscape scales. The model outputs are desirable for diverse applications including long-term ecosystem monitoring, Earth-system modeling and wildlife habitat assessments.

Supplementary Materials: Data products are available to download from the North Slope Science Catalog (<http://catalog.northslope.org/catalog/entries/8595>). The following are available online at www.mdpi.com/2072-4292/9/10/1024/s1: Figure S1: Maps of the total cover of (a) sedge, (b) grass and (c) forb based on random forest models with 78 predictors, North Slope, Alaska. Figure S2: Maps of total cover of (a) vascular plants, (b) lichen and (c) bryophytes based on random forest models with 78 predictors, North Slope, Alaska. Figure S3: Maps of top cover of (a) open water, (b) bare ground and (c) litter based on random forest models with 78 predictors, North Slope, Alaska. Table S1. Five most important predictors for cover models by plant functional type (PFT) or other cover type, North Slope, Alaska.

Acknowledgments: Shell Exploration & Production Company funded the field work, preliminary modeling, and manuscript preparation as part of their multi-year Onshore Environmental Studies Program in northern Alaska. We especially thank Erling Westlien and Michael Macrander at Shell. Additional support for modeling, automation, and manuscript publication came from NASA Arctic Boreal Vulnerability Experiment (ABOVE) Grant NNX15AU05A. We also thank the many field biologists at ABR, Inc.–Environmental Research & Services and BLM that collected the field data; and the BLM Assessment, Inventory, and Monitoring program for sharing their field data.

Author Contributions: M.J.M. conceived of and designed the field data summaries and development of spectral and environmental predictors. M.J.M., G.V.F. and P.R.N. developed and evaluated the models. C.S.S., M.J.M. and P.R.N. developed scripts to facilitate analysis and visualization of the results. M.J.M., G.V.F. and P.R.N. wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest. The funding sponsors had no role in the design of the study; in the collection, analyses or interpretation of the data; in the writing of the manuscript; nor in the decision to publish the results.

Appendix A

Surface water is a widespread and dynamic feature in northern Alaska, especially on the flat coastal plain where drainage is impeded by shallow permafrost. Although perennial lakes and ponds are abundant, the extent of surface water changes dramatically in terrestrial areas through the summer; we therefore produced a continuous-field map of seasonal water frequency by classifying the seasonal composites using the a priori spectral classifier of the Satellite Image Automatic Mapper (SIAM) [61]. We applied the fine-scale SIAM classification consisting of 96 spectral categories and then aggregated the classification for each composite into three classes: water, land and unknown. Next, we calculated the sum of water and land observations for each pixel across all seasonal windows and calculated surface water frequency as the count of water observations divided by the total count of water and land observations; we excluded unknown observations from the water-frequency calculations because they could represent cloud, snow-covered lake ice or tundra. Although we carefully screened the Landsat data, a few cloud shadows remained in most composites and tended to be mapped as water. Therefore, we set a minimal threshold of two water observations to map the frequency of water pixels; if there were only one water observation for a pixel, we set the frequency to zero. Dark fire scars within the boundary of the large Anaktuvuk River fire (burned in 2007) were also occasionally mapped as water, so we set the minimal threshold for that area and other known 1999–2015 fire scars [62] to seven water observations.

References

1. Markon, C.J.; Derksen, D.V. Identification of tundra land cover near Teshekpuk Lake, Alaska using SPOT satellite data. *Arctic* **1994**, *47*, 222–231. [[CrossRef](#)]
2. Muller, S.V.; Racoviteanu, A.E.; Walker, D.A. Landsat MSS-derived land-cover map of northern Alaska: Extrapolation methods and a comparison with photo-interpreted and AVHRR-derived maps. *Int. J. Remote Sens.* **1999**, *20*, 2921–2946. [[CrossRef](#)]

3. Jorgenson, M.T.; Heiner, M. *Ecosystems of Northern Alaska*; The Nature Conservancy: Anchorage, AK, USA, 2003.
4. Jorgenson, M.T.; Roth, J.E.; Miller, P.F.; Macander, M.J.; Duffy, M.S.; Wells, A.F.; Frost, G.V.; Pullman, E.R. *An Ecological Land Survey and Landcover Map of the Arctic Network*; Natural Resource Technical Report NPS/ARC/NRTR—2009/270; U.S. Department of the Interior: Fort Collins, CO, USA, 2009; p. 307.
5. North Slope Science Initiative (NSSI). *North Slope Science Initiative Landcover Mapping Summary Report*; Ducks Unlimited, Inc.: Rancho Cordova, CA, USA, 2013; p. 51.
6. Walker, D.A.; Everett, K.R.; Webber, P.J.; Brown, J. *Geobotanical Atlas of the Prudhoe Bay Region, Alaska*; U.S. Army Cold Regions Research and Engineering Laboratory: Hanover, NH, USA, 1980; p. 68.
7. Walker, D.A.; Maier, H.A. *Vegetation in the Vicinity of the Toolik Field Station, Alaska*; Biological Papers of the University of Alaska No. 28; Institute of Arctic Biology: Fairbanks, AK, USA, 2008; p. 2.
8. Ustin, S.L.; Gamon, J.A. Remote sensing of plant functional types. *New Phytol.* **2010**, *186*, 795–816. [[CrossRef](#)] [[PubMed](#)]
9. Huemmrich, K.F.; Gamon, J.A.; Tweedie, C.E.; Campbell, P.K.E.; Landis, D.R. Middleton Arctic tundra vegetation functional types based on photosynthetic physiology and optical properties. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 265–275. [[CrossRef](#)]
10. Hansen, M.C.; Potapov, P.V.; Moore, R.; Hancher, M.; Turubanova, S.A.; Tyukavina, A.; Thau, D.; Stehman, S.V.; Goetz, S.J.; Loveland, T.R.; et al. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* **2013**, *342*, 850–853. [[CrossRef](#)] [[PubMed](#)]
11. Walker, M.D.; Wahren, C.H.; Hollister, R.D.; Henry, G.H.R.; Ahlquist, L.E.; Alatalo, J.M.; Bret-Harte, M.S.; Calef, M.P.; Callaghan, T.V.; Carroll, A.B.; et al. Plant community responses to experimental warming across the tundra biome. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 1342–1346. [[CrossRef](#)] [[PubMed](#)]
12. Serreze, M.C.; Barry, R.G. Processes and impacts of Arctic amplification: A research synthesis. *Glob. Planet. Chang.* **2011**, *77*, 85–96. [[CrossRef](#)]
13. Myers-Smith, I.H.; Elmendorf, S.C.; Beck, P.S.A.; Wilkening, M.; Hallinger, M.; Blok, D.; Tape, K.D.; Rayback, S.A.; Macias-Fauria, M.; Forbes, B.C.; et al. Climate sensitivity of shrub growth across the tundra biome. *Nat. Clim. Chang.* **2015**, *5*, 887–891. [[CrossRef](#)]
14. Chapin, F.S.; Sturm, M.; Serreze, M.C.; McFadden, J.P.; Key, J.R.; Lloyd, A.H.; McGuire, A.D.; Rupp, T.S.; Lynch, A.H.; Schimel, J.P.; et al. Role of land-surface changes in arctic summer warming. *Science* **2005**, *310*, 657–660. [[CrossRef](#)] [[PubMed](#)]
15. Sturm, M.; Douglas, T.; Racine, C.; Liston, G.E. Changing snow and shrub conditions affect albedo with global implications. *J. Geophys. Res.* **2005**, *110*, G01004. [[CrossRef](#)]
16. Loranty, M.M.; Goetz, S.J.; Beck, P.S.A. Tundra vegetation effects on pan-Arctic albedo. *Environ. Res. Lett.* **2011**, *6*, 029601. [[CrossRef](#)]
17. Blok, D.; Heijmans, M.M.P.D.; Schaepman-Strub, G.; Kononov, A.V.; Maximov, T.C.; Berendse, F. Shrub expansion may reduce summer permafrost thaw in Siberian tundra. *Glob. Chang. Biol.* **2010**, *16*, 1296–1305. [[CrossRef](#)]
18. Myers-Smith, I.H.; Hik, D.S. Shrub canopies influence soil temperatures but not nutrient dynamics: An experimental test of tundra snow-shrub interactions. *Ecol. Evol.* **2013**, *3*, 3683–3700. [[CrossRef](#)] [[PubMed](#)]
19. Defries, R.S.; Field, C.B.; Fung, I.; Justice, C.O.; Los, S.; Matson, P.A.; Matthews, E.; Mooney, H.A.; Potter, C.S.; Prentice, K.; et al. Mapping the land surface for global atmosphere-biosphere models: Toward continuous distributions of vegetation's functional properties. *J. Geophys. Res. Atmos.* **1995**, *100*, 20867–20882. [[CrossRef](#)]
20. Epstein, H.E.; Calef, M.P.; Walker, M.D.; Stuart Chapin, F.; Starfield, A.M. Detecting changes in arctic tundra plant communities in response to warming over decadal time scales. *Glob. Chang. Biol.* **2004**, *10*, 1325–1334. [[CrossRef](#)]
21. Wullschleger, S.D.; Epstein, H.E.; Box, E.O.; Euskirchen, E.S.; Goswami, S.; Iversen, C.M.; Kattge, J.; Norby, R.J.; van Bodegom, P.M.; Xu, X. Plant functional types in Earth system models: Past experiences and future directions for application of dynamic vegetation models in high-latitude ecosystems. *Ann. Bot.* **2014**, *114*, 1–16. [[CrossRef](#)] [[PubMed](#)]
22. Walker, D.A.; Epstein, H.E.; Jia, G.J.; Balsler, A.; Copass, C.; Edwards, E.J.; Gould, W.A.; Hollingsworth, J.; Knudson, J.A.; Maier, H.A.; et al. Phytomass, LAI, and NDVI in northern Alaska: Relationships to summer warmth, soil pH, plant functional types, and extrapolation to the circumpolar Arctic. *J. Geophys. Res.* **2003**, *108*, 8169. [[CrossRef](#)]

23. Beyer, H.L.; Haydon, D.T.; Morales, J.M.; Frair, J.L.; Hebblewhite, M.; Mitchell, M.; Matthiopoulos, J. The interpretation of habitat preference metrics under use-availability designs. *Philos. Trans. R. Soc. B Biol. Sci.* **2010**, *365*, 2245–2254. [[CrossRef](#)] [[PubMed](#)]
24. Tape, K.; Sturm, M.; Racine, C. The evidence for shrub expansion in northern Alaska and the Pan-Arctic. *Glob. Chang. Biol.* **2006**, *12*, 686–702. [[CrossRef](#)]
25. Myers-Smith, I.H.; Forbes, B.C.; Wilkening, M.; Hallinger, M.; Lantz, T.; Blok, D.; Tape, K.D.; Macias-Fauria, M.; Sass-Klaassen, U.; Lévesque, E.; et al. Shrub expansion in tundra ecosystems: Dynamics, impacts and research priorities. *Environ. Res. Lett.* **2011**, *6*, 045509. [[CrossRef](#)]
26. Sturm, M.; Schimel, J.; Michaelson, G.; Welker, J.M.; Oberbauer, S.F.; Liston, G.E.; Fahnestock, J.; Romanovsky, V.E. Winter biological processes could help convert arctic tundra to shrubland. *BioScience* **2005**, *55*, 17–26. [[CrossRef](#)]
27. Fraser, R.H.; Lantz, T.C.; Olthof, I.; Kokelj, S.V.; Sims, R.A. Warming-induced shrub expansion and lichen decline in the western Canadian Arctic. *Ecosystems* **2014**, *17*, 1151–1168. [[CrossRef](#)]
28. Selkowitz, D.J. A comparison of multi-spectral, multi-angular, and multi-temporal remote sensing datasets for fractional shrub canopy mapping in Arctic Alaska. *Remote Sens. Environ.* **2010**, *114*, 1338–1352. [[CrossRef](#)]
29. Beck, P.S.A.; Horning, N.; Goetz, S.J.; Loranty, M.M.; Tape, K.D. Shrub cover on the North Slope of Alaska: A circa 2000 baseline map. *Arct. Antarct. Alp. Res.* **2011**, *43*, 355–363. [[CrossRef](#)]
30. Blok, D.; Heijmans, M.M.P.D.; Schaepman-Strub, G.; van Ruijven, J.; Parmentier, F.J.W.; Maximov, T.C.; Berendse, F. The cooling capacity of mosses: Controls on water and energy fluxes in a Siberian tundra site. *Ecosystems* **2011**, *14*, 1055–1065. [[CrossRef](#)]
31. Walker, D.A.; Auerbach, N.A.; Bockheim, J.G.; Chapin, F.S.; Eugster, W.; King, J.Y.; McFadden, J.P.; Michaelson, G.J.; Nelson, F.E.; Oechel, W.C.; et al. Energy and trace-gas fluxes across a soil pH boundary in the Arctic. *Nature* **1998**, *394*, 469–472. [[CrossRef](#)]
32. Russell, D.E.; Martell, A.M.; Nixon, W.A.C. Range ecology of the Porcupine caribou herd in Canada. *Rangifer* **1993**, *13*, 1–168. [[CrossRef](#)]
33. Heggberget, T.M.; Gaare, E.; Ball, J.P. Reindeer (*Rangifer tarandus*) and climate change: Importance of winter forage. *Rangifer* **2002**, *22*, 13–31. [[CrossRef](#)]
34. Joly, K.; Cole, M.J.; Jandt, R.R. Diets of overwintering caribou, *Rangifer tarandus*, track decadal changes in Arctic tundra vegetation. *Can. Field Nat.* **2007**, *121*, 379–383. [[CrossRef](#)]
35. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
36. Gallant, A.L.; Binnian, E.F.; Omernik, J.M.; Shasby, M.B. *Ecoregions of Alaska*; U.S. Geological Survey Professional Paper 1567; United States Government Printing Office: Washington, DC, USA, 1995; p. 73.
37. CAVM Team. *Circumpolar Arctic Vegetation Map (1:7,500,000 Scale)*; Conservation of Arctic Flora and Fauna (CAFF) Map No. 1; U.S. Fish and Wildlife Service: Anchorage, AK, USA, 2003.
38. Brown, J.; Ferrians, O.J.J.; Heginbottom, J.A.; Melnikov, E.S. *Circum-Arctic Map of Permafrost and Ground Ice Conditions*; National Snow and Ice Data Center: Boulder, CO, USA, 2001.
39. Toevs, G.R.; Karl, J.W.; Taylor, J.J.; Spurrier, C.S.; Bobo, M.R.; Herrick, J.E. Consistent indicators and methods and a scalable sample design to meet assessment, inventory, and monitoring information needs across scales. *Rangelands* **2011**, *33*, 14–20. [[CrossRef](#)]
40. Masek, J.G.; Vermote, E.F.; Saleous, N.E.; Wolfe, R.; Hall, F.G.; Huemmrich, K.F.; Gao, F.; Kutler, J.; Lim, T.-K. A Landsat surface reflectance dataset for North America, 1990–2000. *IEEE Geosci. Remote Sens. Lett.* **2006**, *3*, 68–72. [[CrossRef](#)]
41. Irish, R.R.; Barker, J.L.; Goward, S.N.; Arvidson, T. Characterization of the Landsat 7 ETM+ Automated Cloud-Cover Assessment (ACCA) Algorithm. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 1179–1188. [[CrossRef](#)]
42. Zhu, Z.; Woodcock, C.E. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sens. Environ.* **2012**, *118*, 83–94. [[CrossRef](#)]
43. Reynolds, M.; Comiso, J.; Walker, D.; Verbyla, D. Relationship between satellite-derived land surface temperatures, arctic vegetation types, and NDVI. *Remote Sens. Environ.* **2008**, *112*, 1884–1894. [[CrossRef](#)]
44. Macander, M.J.; Swingley, C.S.; Joly, K.; Reynolds, M.K. Landsat-based snow persistence map for northwest Alaska. *Remote Sens. Environ.* **2015**, *163*, 23–31. [[CrossRef](#)]
45. Rouse, J.W.; Haas, R.H.; Scheel, J.A.; Deering, D.W. Monitoring vegetation systems in the Great Plains with ERTS. In Proceedings of the 3rd Earth Resource Technology Satellite (ERTS) Symposium, Greenbelt, MD, USA, 10–14 December 1974; Volume 1, pp. 48–62.

46. Tucker, C.J.; Elgin, J.H.; McMurtrey, J.E.; Fan, C.J. Monitoring corn and soybean crop development with hand-held radiometer spectral data. *Remote Sens. Environ.* **1979**, *8*, 237–248. [[CrossRef](#)]
47. Jiang, Z.; Huete, A.R.; Didan, K.; Miura, T. Development of a two-band Enhanced Vegetation Index without a blue band. *Remote Sens. Environ.* **2008**, *112*, 3833–3845. [[CrossRef](#)]
48. McFeeters, S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [[CrossRef](#)]
49. Goodwin, N.R.; Coops, N.C.; Wulder, M.A.; Gillanders, S.; Schroeder, T.A.; Nelson, T. Estimation of insect infestation dynamics using a temporal sequence of Landsat data. *Remote Sens. Environ.* **2008**, *112*, 3680–3689. [[CrossRef](#)]
50. Hall, D.K.; Riggs, G.A.; Salomonson, V.V. Development of methods for mapping global snow cover using moderate resolution imaging spectroradiometer data. *Remote Sens. Environ.* **1995**, *54*, 127–140. [[CrossRef](#)]
51. Key, C.H.; Benson, N.C. *The Normalized Burn Ratio (NBR): A Landsat TM Radiometric Measure of Burn Severity*; U.S. Geological Survey, Northern Rocky Mountain Science Center: Bozeman, MT, USA, 1999.
52. Liaw, A.; Wiener, M. Classification and regression by randomForest. *R. News* **2002**, *2*, 18–22.
53. Mahoney, C.; Hopkinson, C.; Held, A.; Simard, M. Continental-scale canopy height modeling by integrating national, spaceborne, and airborne LiDAR data. *Can. J. Remote Sens.* **2016**, *42*, 574–590. [[CrossRef](#)]
54. Strobl, C.; Boulesteix, A.-L.; Zeileis, A.; Hothorn, T. Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinform.* **2007**, *8*, 25. [[CrossRef](#)] [[PubMed](#)]
55. Strobl, C.; Boulesteix, A.-L.; Kneib, T.; Augustin, T.; Zeileis, A. Conditional variable importance for random forests. *BMC Bioinform.* **2008**, *9*, 307. [[CrossRef](#)] [[PubMed](#)]
56. Hothorn, T.; Bühlmann, P.; Dudoit, S.; Molinaro, A.; Van Der Laan, M.J. Survival ensembles. *Biostatistics* **2006**, *7*, 355–373. [[CrossRef](#)] [[PubMed](#)]
57. Montesano, P.; Neigh, C.; Sexton, J.; Feng, M.; Channan, S.; Ranson, K.; Townshend, J. Calibration and validation of Landsat tree cover in the taiga–tundra ecotone. *Remote Sens.* **2016**, *8*, 551. [[CrossRef](#)]
58. Jones, B.M.; Breen, A.L.; Gaglioti, B.V.; Mann, D.H.; Rocha, A.V.; Grosse, G.; Arp, C.D.; Kunz, M.L.; Walker, D.A. Identification of unrecognized tundra fire events on the north slope of Alaska. *J. Geophys. Res. Biogeosci.* **2013**, *118*, 1–11. [[CrossRef](#)]
59. Stow, D.A.; Burns, B.H.; Hope, A.S. Spectral, spatial and temporal characteristics of Arctic tundra reflectance. *Int. J. Remote Sens.* **1993**, *14*, 2445–2462. [[CrossRef](#)]
60. Bret-Harte, M.S.; Shaver, G.R.; Zoerner, J.P.; Johnstone, J.F.; Wagner, J.L.; Chavez, A.S.; Gunkelman, R.F.I.; Lippert, S.C.; Laundre, J.A. Developmental plasticity allows *Betula nana* to dominate tundra subjected to an altered environment. *Ecology* **2001**, *82*, 18–32. [[CrossRef](#)]
61. Baraldi, A.; Puzzolo, V.; Blonda, P.; Bruzzone, L.; Tarantino, C. Automatic spectral rule-based preliminary mapping of calibrated Landsat TM and ETM+ images. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 2563–2586. [[CrossRef](#)]
62. AICC (Alaska Interagency Coordination Center). Historical Fire Information. Available online: https://afsmaps.blm.gov/imf_firehistory/imf.jsp?site=firehistory (accessed on 4 October 2017).

