Pre-print of published version

Reference:
Wiebe Nijland ; Scott E. Nielsen ; Nicholas C. Coops ; Michael A. Wulder ; Gordon B. Stenhouse. (2014), Fine-Spatial Scale Predictions of Understory Species Using Climate and LiDAR-Derived Terrain and Canopy Metrics. J. Appl. Remote Sens. 8(1), 083572

DOI

http://dx.doi.org/10.1117/1.JRS.8.083572 (published online: 11 August 2014).

Fine-Spatial Scale Predictions of Understory Species Using Climate and LiDAR-Derived Terrain and Canopy Metrics.

Wiebe Nijland ; Scott E. Nielsen ; Nicholas C. Coops ; Michael A. Wulder ; Gordon B. Stenhouse

* corresponding author. Nicholas.coops@ubc.ca

ABSTRACT

Food and habitat resources are critical components wildlife management and conservation efforts. Grizzly bear (*Ursus arctos*) have diverse diets and habitat requirements particularly for understory plant species which are impacted by human developments and forest management activities. In this paper use Light Detection and

Ranging (LiDAR) data to predict the occurrence of 14 understorey plant species relevant to bear forage and compare our predictions to more conventional climate- and land cover-based models. We use boosted regression trees to model each of the 14 understory species across 4435 km^2 using occurrence (presence-absence) data from 1,941 field plots. Three sets of models were fitted: climate-only, climate and basic land and forest cover from Landsat 30m imagery, and third a climate and LiDAR-derived model describing both the terrain and forest canopy. Resulting model accuracies varied widely among species. Overall, 8 of 14 species models were improved by including the LiDAR-derived variables. For climate-only models, mean annual precipitation and frost-free period were most important variables. With inclusion of LiDAR-derived attributes, Depth to water table, terrain-intercepted annual radiation, and elevation were most often selected. This suggests fine-scale terrain conditions affect the distribution of the studied species more than canopy conditions.

Keywords:

LiDAR, Species distribution modelling, Grizzly bear (Ursus arctos), Understory vegetation.

1. Introduction

Developing a comprehensive understanding of food and habitat resource use for large mammals is a critical component for their conservation and management, as well as for assessing cumulative effects of human impacts and estimating habitat-based carrying capacities for species of management concern (Gordon et al., 2004). This is particularly true for grizzly bear (*Ursus arctos*) populations in western Alberta, Canada where resource extraction is expanding (e.g., forestry, exploration and mining, and urban expansion) along with human use of the landscape, resulting in concern for the long term survival of the species in this region (Clark et al., 1996; Nielsen et al., 2006; S. Nielsen et al., 2004; Nielsen et al., 2008).

Grizzly bears are considered habitat generalists (Schwartz et al., 2003) with diverse, seasonal diet and habitat requirements. Within this region, optimal habitat is a mosaic of forested and non-forested areas (S. E. Nielsen et al., 2004a). Large gaps within forest stands, alpine meadows, and areas regenerating after fire offer bears an abundance of understory species, including seasonal fruits, ants, ungulates, green herbaceous vegetation, roots, other subterranean foods, and grasses which can form a major part of the species diet for at least some part of the year (Martin, 1983; Munro et al., 2006; Zager et al., 1983). Forest harvesting can provide similar habitat, as regenerating stands share many understory species. As a result, forest cutblocks provide an alternate habitat resource to otherwise open areas (Benn and Herrero, 2002; S. Nielsen et al., 2004). However, ongoing forestry and resource extraction activity is also a threat to grizzly bears; management agencies within Alberta are actively trying to balance economic development needs with the conservation needs of the species (Wilkinson et al., 2008). Key to this conservation priority is a comprehensive understanding of how food resource availability and abundance may vary in response to forest management, where the ultimate goal is to better understand the observed habitat use of bears in western-central Alberta, Canada (S. E. Nielsen et al., 2004b).

Grizzly bears have three distinct foraging seasons: hypophagia, early hyperphagia, and late hyperphagia (Nielsen et al., 2006). During hypophagia, grizzly bears in our study area feed on the roots of *Hedysarum* spp. (sweetvetch) and other early herbaceous material. During early hyperphagia, their diet extends to green herbaceous material such as *Heracleum lanatum* (cow-parsnip) and *Equisetum* spp. (horsetail), while in the later season berries such as *Shepherdia canadensis* (buffalo berry) and *Vaccinium* spp. (huckleberry, blueberry, and lingonberry) make up the majority of their diet. As fruit availability declines in the fall, grizzly bears once again dig for sweetvetch roots (Munro et al., 2006; Nielsen et al., 2010, 2006, 2005). While animal matter and insects are an important food resource for grizzly bears during spring and early summer, the variety of vegetable matter (including roots, forbs, and fruit) makes up the majority of their diet between late June and early October.

A comprehensive understanding of the horizontal distribution of understory flora is required to accurately predict wildlife-habitat relationships (Linderman et al., 2005; Lindzey and Meslow, 1977; MacArthur and MacArthur, 1961; Roughgarden et al., 1991). Neglecting to consider the influence of understory vegetation in broad-scale habitat studies has limited the relevance of fine-scale monitoring of animals(e.g., radio collar movement data) for accurate conservation and management planning (Linderman et al., 2005). Remote sensing vegetation studies provide the opportunity to bridge this gap.

The distribution of these plant species is influenced by local landscape conditions and canopy structure (S. E. Nielsen et al., 2004a). For example, many berry species have higher yields in open canopies, while plant species (e.g., clover and dandelions) thrive in recently opened areas such as forest clearings associated with anthropogenic disturbance. The most common approaches to species abundance and occurrence modelling rely on empirical correlations with environmental variables to develop “niche” or “bioclimatic envelope” models (Austin, 1985; Iverson and Prasad, 1998; Mckenzie et al., 2003; Thuiller et al., 2008). These models usually empirically relate presence / absence data to environmental variables, most often with climate (but sometimes including soil and physiographic features), using an array of statistical methods including multiple regression techniques, neural networks, and regression tree analysis (Iverson and Prasad, 2001). Climate surfaces are effective predictors of broad scale patterns and a number of studies have linked climate and land cover information derived from optical remote sensing data (Nielsen et al., 2010, 2003), such as forest cover classes derived from Landsat Thematic Mapper imagery

(Mcdermid, 2005). Land cover attributes derived from optical remote sensing have been shown to increase the predictive power of models, but they are still unable to fully represent the fine scale processes related to stand and canopy conditions, particularity in areas where forest management regularly changes the structure of the forest.

Light detection and ranging (LiDAR), is an increasingly well understood and established remote sensing technology which is able to detect both topographic and canopy features within forest ecosystems at previously unavailable levels of accuracy (Wulder et al., 2008). Airborne LiDAR systems function by emitting and receiving laser energy that measure distance to target surfaces. Laser systems, when combined with Global Positioning Systems (GPS) and orientation systems (e.g., Inertial Navigation Systems), allow the location of surfaces intercepted by the beam to be precisely computed (Gaveau and Hill, 2003) with vertical and horizontal accuracies approximately within 40 cm (Davenport et al., 2004). Especially for understory applications, a key advantage of LiDAR over conventional optical remote sensing imagery (such as Landsat, or high spatial resolution imagery, like Quickbird and Worldview) is the ability to describe the canopy structure in three dimensions including information on areas otherwise obscured by the tree canopy.

LiDAR data has shown promise in estimating understory structural attributes despite there being a limited number of studies of which most are associated with fire fuel prediction. Seielstad and Queen(2003) investigated the potential of using LiDAR data to quantify fuel for a widely applied fuel model. Riaño et al. (2003)used cluster analysis, based on the minimum Euclidean distance, to distinguish understory from overstory returns in a mixed conifer and deciduous two-tiered forest. Maltamo et al. (2005) applied regression models to estimate the number and mean height of suppressed understory trees in a boreal forest using LiDAR data with r^2 values of 0.87 and 0.76, respectively.

This study aims to evaluate the integration of LiDAR data into large area studies on species distribution. To do so we assess the effectiveness of using LiDAR remote sensing data to predict species occurrence for 14 understory plant species relevant to bear habitat and food. We compare these to more conventional climate- and land cover-based models of species occurrence to evaluate whether LiDAR data improves our understanding of the local distribution of bear foods. We compiled, and derived a number of topographic and canopy metrics from airborne LiDAR data, and combined them with climate and land cover data, to model the distribution of 14 key plant species in the Alberta foothills region. Model performance and spatial patterns of the three sets of models were compared. In addition we assessed variable importance within the models to increase our understanding of the main environmental drivers on plant distribution in the study area and our ability to capture those drivers in different data sources.

2. Materials and Methods

2.1 Study area:

The study area is situated in the Rocky Mountains and Foothills area in western Alberta stretching from the North Saskatchewan River (Highway 11) in the south to Grande Prairie in the north, with elevations ranging between 600m and 3300m. Variations in climate and topography generate a vegetation species gradient from the south west to the north east. Higher elevation and rugged, conifer-dominated forests of the Subalpine and Upper Foothills transitions to a lower elevation, gently rolling terrain that is characteristic of the Lower Foothills and Central Mixedwood subregions. Generally the overstory structure of the stands is relatively simple, with regeneration after large wildfires common. As a result stand structure is relatively constant with respect to age (Cieszewski and Bella, 1989; Kirby, 1975) The dynamics of forest stand height and cover increase quickly when the stand is younger, slowing considerably when the stand is older.

2.2 Plant distribution data

Field data were collected between the years 2001 and 2008 as part of the Foothills Research Institute Grizzly Bear Project (Figure 1). In total, 2,338 plots were sampled as described in detail by Nielsen et al. (2010; 2004), within a study area of 4435 km^2. Vegetation was sampled at 5 places (0,5m^2) on 20m transect. The transect center was registered by GPS with an accuracy within 10m. To avoid issues with the temporal differences between collection of the plot and remotely sensed data, we utilised a Landsat-derived disturbance layer (Hilker et al., 2009) which provided information on recent harvest and fire events within the region. All plots which occurred in recently disturbed areas and areas disturbed between the plot visit and remote sensing data acquisition were excluded from the analysis, as were plots outside of the LiDAR coverage (429 excluded, 1,944 plots used). Fourteen species formed the basis of the analysis including species important for root digging, herbivory, and fruiting (Table 1).

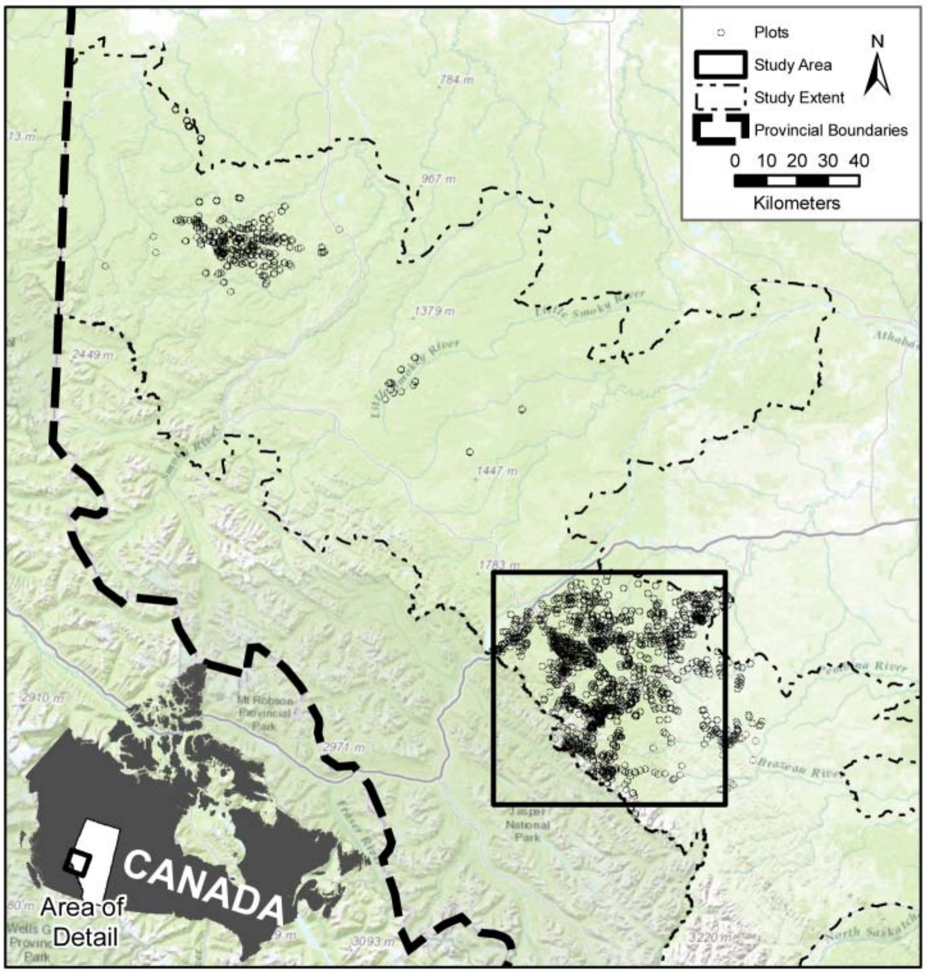

Figure 1: Overview of the study area and plot locations. The black square indicates the area displayed in species specific figures (Figure 2).

Table 1: Focus understory plant species

Scientific name	Common name	Plots Present
Root digging		
Hedysarum alpinum	alpine sweetvetch	7.6%

Lathyrus ochroleucus	creamy peavine	5.9%
Herbivory		
Equisetum arvense	Horsetail	21.3%
Heracleum lanatum	cow-parsnip	6.2%
Taraxacum officinale	dandelion	18.2%
Frugivory		
Arctostaphylos uva-ursi	bearberry	12.9%
Fragaria virginiana	wild strawberry	38.3%
Rosa acicularis	prickly rose	32.5%
Rubus idaeus	wild red raspberry	13.5%
Rubus pedatus	five leaf bramble	9.4%
Shepherdia Canadensis	buffaloberry	8.7%
Vaccinium caespitosum	dwarf blueberry	24.4%
Vaccinium vitis-idaea	lingonberry	38.8%
Viburnum edule	highbush cranberry	8.8%

2.3 Environmental Covariates

Climate: Spatial predictors of the region included a number of seasonal and annual climate variables which were derived from long term (1961-1990) climate records, using the CLIMATE-WNA (Wang et al., 2012) which uses a PRISM down-sampling (Daly et al., 1993) approach to create surfaces at a 500*500m resolution. These included mean maximum and minimum temperature, growing degree days (base 0°C), frost free period, mean annual precipitation during the growing season, and summer moisture index (Table 2).

Table 2. Environmental covariates utilised in Boosted Regression Tree Modelling

	Range			unit
Climate				
max mean maximum monthly temp.	10.3	–	22.6	°C
min mean minimum monthly temp.	-18.9	–	-16.1	°C
degree days base 0°C	1134.0	–	2014.0	days * °C
frost free period	46.4	–	101.4	days
growing season precipitation	377.5	–	532.7	mm
mean annual precipitation	516.9	–	965.9	mm
summer moisture index	0.2	–	3.0	–
Forest				
landcover class (14 classes)	0	–	13	categorical
regenerating forest mask	0	–	1	binary
canopy cover	0.0	–	95.0	%

percent conifers	0.0	–	99.0	%
Lidar				
Max height above ground	0.1	–	35.1	meter
Mean height above ground	0.0	–	17.7	meter
05th percentile	-0.6	–	0.9	meter
50th percentile	0.0	–	21.0	meter
95th percentile	0.0	–	27.7	meter
fraction points above 2m	0.0	–	94.1	%
relative height ratio	0.0	–	0.7	–
Skewness	-2.2	–	21.4	–
standard deviation	0.0	–	11.2	meter
Elevation	858.9	–	2266.6	meter
Slope	0.1	–	40.3	degrees
Aspect	0.0	–	360.0	degrees
terrain solar index	0.9	–	2.0	–
canopy solar index	1935.0	–	2618.0	–
canopy and terrain solar index	2218.0	–	3810.0	–
Wet Area Map	0.0	–	257.0	meter

Land and Forest Cover: Landsat-derived land cover information was available for the study region and included information on land cover, canopy cover (%), and percent of pixel dominated by conifer overstory species (McDermid et al., 2009) The products were based on Landsat images acquired between 2005 and 2009 and have a 30*30m resolution, geolocation accuracy is typical within one pixel (Lee et al., 2004)

LiDAR Data: LiDAR data were provided by the Alberta Environment and Sustainable Resource Department (AESRD), who compiled a globally unique compilation of LiDAR datasets acquired from 2003 – 2008. The compiled LiDAR dataset covers the majority of the forested areas of the Province of Alberta extending over 25 million hectares. The LIDAR dataset was compiled by the Government of Alberta from a variety of sources including forestry, mining and exploration companies. The extremely large area covered by the compilation allow broad-scale environmental issues (such as species habitat relations) to be addressed. Typical characteristics of the multiple data acquisitions were re multiple return, small footprint, acquisition from a fixed wing platform with nominal post spacing of approximately 0.75 points per square meter, vertical and horizontal accuracies are typically within 40 cm (Davenport et al., 2004).To minimise the impact of different survey configurations and acquisition dates (e.g., hit density, or leaf-on/off), the data were thinned to produce a consistent 1m spacing dataset, which, despite being lower than many typical

LiDAR datasets (Wulder et al., 2008), ensured consistent density and coverage over the entire 4435 km^2 study area. From the thinned LiDAR point cloud, a bare Earth DEM (Digital Elevation Model) and a canopy height model were provided at 1m raster resolution. From the bare Earth DEM, slope aspect and elevation were extracted for each plot location. A suite of forest canopy metrics were then developed for each 25*25m pixel, including a calculation of percentiles from 5 to the 95^{th} in steps of 5%, where a given height percentile was calculated as the height greater than a given percentage of LiDAR first returns (Means et al., 2000) mean height, maximum height, fraction points above 2m, relative height ratio(mean height/ max height), skewness of the percentile height, and standard deviation of heights were also computed for each plot. In addition to the canopy and topographic metrics, information on the annual radiation regime for the bare Earth DEM, canopy height, and terrain and canopy elevation for each plot was calculated from the LiDAR data using a hemispherical viewshed algorithm (Fu and Rich, 2002; Rich et al., 1994), which incorporates extraterrestrial solar flux, the relative optical path (determined by the solar zenith angle and elevation above sea level), the duration of a defined time interval, and the effect of the surface orientation (Garnier and Ohmura, 1968). Lastly, a Wet-Areas Mapping (WAM) layer was available, providing an estimate of depth to water table using the shape and orientation of the terrain (White et al., 2012), the WAM was based on the same LiDAR elevation models, created a 1m raster resolution and resampled to 25m. Input variables used for the models are listed in Table 2. For the LiDAR derived variables we made a selection capturing different aspects of terrain and vegetation cover while limiting the overall number of variables and multi-collinearity, based on other studies including (Coops et al., 2010; Ferster et al., 2009) Previous LiDAR approaches involving the direct detection of understory structure (Martinuzzi et al., 2009; Wing et al., 2012) was not possible due to insufficient point density and limitation in separating low vegetation and ground returns in the compiled dataset. As a result, overstorey and terrain characteristics were used as surrogate predictors of understory structure.

2.4 Modelling: Boosted Regression Trees

Distribution models were built for the 14 plant species using Boosted Regression Trees from the 'gbm' package in R statistical software (R Development Core Team, 2013), it follows the methods described in Friedman (Friedman, 2001, 2002). Boosted Regression Tree modelling is a relatively new technique which is gaining popularity in the distribution modelling community (Elith et al., 2008). Benefits include flexibility in combining different types of variables (e.g., continuous, categorical, nominal), flexibility in statistical distributions, and demonstrate high predictive power (Elith et al., 2008). Up to 1500 individual trees were fit with a 5 level tree depth and a learning rate of 0.005 to avoid over fitting of collinear variables. The optimum number of trees was selected using a 10 fold cross validation within the training data. To verify the selected model, we made a random 80-20 split of all plot data before the model building and calculated model fit using the separated 20% of the plots. Model performance was assessed using the "Area under the Receiver-Operator Characteristic Curve" (AUC) (Jiménez-Valverde, 2011) with values ranging from 0.5 to 0.7 generally viewed as 'low' model accuracy, values between 0.7 and 0.9 considered 'good', and values greater than 0.9 considered 'high' model accuracy (Manel et al., 2001; Swets, 1988). The kappa coefficient also was calculated, although disputed by some (Pontius and Millones, 2011) it is a widely used ,metric useful particularly in ecological research (see review by Monserud and Leemans(1992)). This statistic calculates the proportion of specific agreement across presence and absent classes.

3. Results

Overviews of the three sets of models developed for the individual species in Table 3 show a wide variety of model accuracy. Model AUC values ranged from 0.70 – 0.85, while K statistic values ranged between 0.09 and 0.48 (i.e., poor to moderate, based on Landis and Koch's (1977) thresholds for the K statistic). Apart from the three model sets shown, we also tested models using LiDAR or Landsat based information only, but these had poor performance with average validation AUC around 0.65 as they fail to capture the larger scale patterns in the study area. The most accurately predicted species was *Hedysarum alpinum* (sweet vetch) while the poorest was *Equisetum arvense* (horsetail). *Hedysarum alpinum* is a critical spring root-digging resource for bears, whereas horsetail produces a high-protein, succulent and herbaceous food resource at green-up (Table 3).

Overall, 8 of the 14 most accurate species models were developed using a combination of climate and LiDAR-derived variables, with an average increase in AUC of 5% and the greatest model improvement of up to 12% for *Arctostaphylos uva-ursi* (bearberry). For three species, the most accurate model derived was from climate and broader scale land and forest cover information, and three were equally supported.

Examining the spatial predictions of the species models, the differences in the spatial resolution of the input parameters was apparent. Figure 2 shows the probability of occurrence for a number of species for subset of the study area (*E.arvense* (horsetail), *H.alpinum* (sweet vetch), *T.officinale* (dandelion), and *V.vitis-idea* (loganberry) based on the 3 different sets of variables. Overall, the coarser nature of the climate data (500m) results in a coarse model output which is unable to reflect changes in forest patterns associated with management or fine scale topographic features across the landscape. In contrast, models developed using either the 30m Landsat-derived land cover or the LiDAR-derived canopy and terrain information were much finer, allowing management and topographic variation to be represented in greater spatial detail.

Examining variable importance (Figure 3) for the climate-only models, mean annual precipitation and frost free period were selected as the most critical variables predicting species occurrence for most species, followed by degree days, growing season precipitation, and summer moisture index. No other climate variables were selected as important in the climate-only model predictions. When forest and land cover variables were added into the models, their overall effect was minor; only percent conifers were additionally selected as an important variable for a single species model.

Table 3: Model results

				Variable Importance		
	Training AUC	Validation AUC	Validation Kappa	1 st	2nd	3rd
Arctostaphylos uva-ursi						

Climate	0.87	0.67	0.15	Growing season Precip.	Mean annual Precip.	Frost free period
Climate + Forest	0.89	0.69	0.26	Growing season Precip.	Frost free period	Degree days base 0
Climate + LiDAR	0.95	0.79	0.34	Terrain Solar	Growing season Precip.	Mean annual Precip.
Equisetum arvense						
Climate	0.80	0.69	0.20	Growing season Precip.	Mean annual Precip.	Frost free period
Climate + Forest	0.78	0.67	0.15	Mean annual Precip.	Growing season Precip.	Frost free period
Climate + LiDAR	0.81	0.69	0.19	Wet Area	Elevation	Mean annual Precip.
Fragaria virginiana						
Climate	0.77	0.67	0.24	Mean annual Precip.	Degree days base 0	Growing season Precip.
Climate + Forest	0.81	0.67	0.21	Mean annual Precip.	Degree days base 0	Frost free period
Climate + LiDAR	0.84	0.74	0.30	Mean annual Precip.	Degree days base 0	Wet Area
Hedysarum alpinum						
Climate	0.94	0.88	0.47	Degree days base 0	Frost free period	Mean annual Precip.
Climate + Forest	0.94	0.89	0.46	Degree days base 0	Frost free period	Mean annual Precip.
Climate + LiDAR	0.96	0.91	0.48	Degree days base 0	Frost free period	Mean annual Precip.
Heracleum lanatum						
Climate	0.87	0.71	0.10	Mean annual Precip.	Frost free period	Growing season Precip.
Climate + Forest	0.88	0.70	0.09	Mean annual Precip.	Growing season Precip.	Frost free period
Climate + LiDAR	0.92	0.80	0.15	Wet Area	Mean annual Precip.	Slope
Lathyrus ochroleucus						
Climate	0.82	0.73	0.10	Mean annual Precip.	Frost free period	Summer moisture index
Climate + Forest	0.84	0.76	0.11	Mean annual Precip.	Frost free period	Degree days base 0
Climate + LiDAR	0.94	0.76	0.21	Mean annual Precip.	Terrain Solar	top of canopy Solar
Rosa acicularis						
Climate	0.77	0.69	0.26	Mean annual Precip.	Degree days base 0	Summer moisture index
Climate + Forest	0.79	0.71	0.23	Mean annual Precip.	Frost free period	Degree days base 0
Climate + LiDAR	0.84	0.75	0.32	Terrain Solar	Mean annual Precip.	Elevation
Rubus idaeus						
Climate	0.85	0.74	0.16	Frost free period	Mean annual Precip.	Degree days base 0
Climate + Forest	0.88	0.79	0.26	Frost free period	Mean annual Precip.	Degree days base 0
Climate + LiDAR	0.91	0.81	0.30	Wet Area	Frost free period	Maximum Height
Rubus pedatus						
Climate	0.84	0.79	0.20	Mean annual Precip.	Growing season Precip.	Degree days base 0
Climate + Forest	0.87	0.84	0.27	Growing season Precip.	Mean annual Precip.	Degree days base 0
Climate + LiDAR	0.91	0.87	0.29	Terrain Solar	Growing season Precip.	Mean annual Precip.
Shepherdia canadensis						
Climate	0.92	0.76	0.19	Degree days base 0	Mean annual Precip.	Summer moisture index
Climate + Forest	0.92	0.78	0.26	Degree days base 0	Frost free period	Growing season Precip.
Climate + LiDAR	0.94	0.76	0.25	Degree days base 0	Terrain Solar	Mean annual Precip.
Taraxacum officinale						
Climate	0.79	0.64	0.10	Mean annual Precip.	Frost free period	Growing season Precip.
Climate + Forest	0.85	0.73	0.24	Frost free period	Mean annual Precip.	Growing season Precip.
Climate + LiDAR	0.88	0.72	0.17	Frost free period	Mean Height	Wet Area
Vaccinium caespitosum						
Climate	0.80	0.67	0.16	Frost free period	Mean annual Precip.	Degree days base 0
Climate + Forest	0.83	0.72	0.24	Frost free period	Mean annual Precip.	Degree days base 0
Climate + LiDAR	0.87	0.73	0.32	Wet Area	5th percentile	Frost free period
Vaccinium vitis-idaea						
Climate	0.83	0.75	0.35	Frost free period	Degree days base 0	Summer moisture index
Climate + Forest	0.86	0.77	0.37	Frost free period	percent conifers	Degree days base 0
Climate + LiDAR	0.89	0.77	0.38	Frost free period	Minimum monthly temp.	Elevation
Viburnum edule						

Climate	0.92	0.80	0.21	Frost free period	Mean annual Precip.	Degree days base 0
Climate + Forest	0.93	0.85	0.31	Frost free period	Mean annual Precip.	Degree days base 0
Climate + LiDAR	0.93	0.82	0.19	standard deviation	canopy Solar	Mean annual Precip.

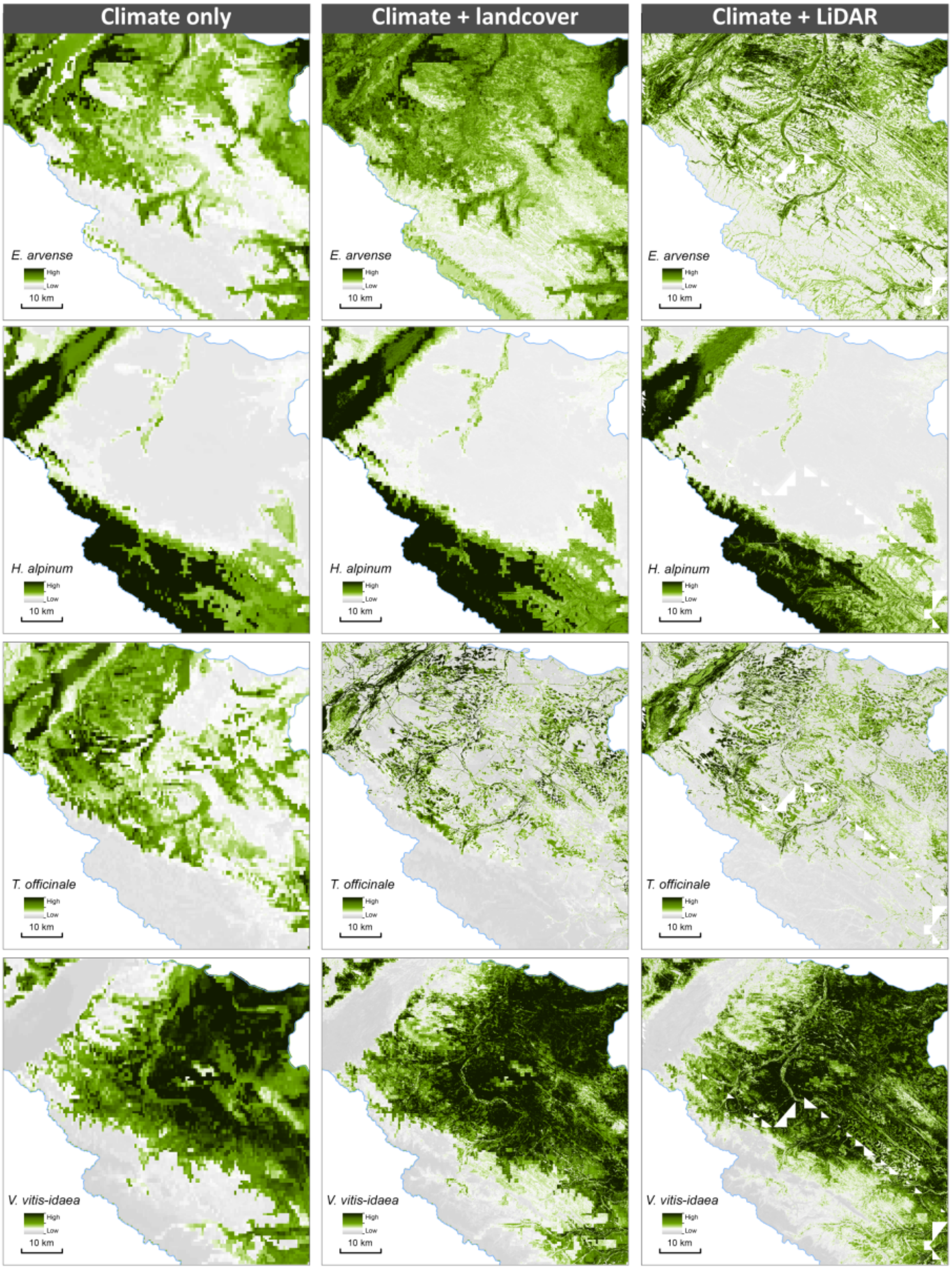

Figure 2: Probability of occurrence maps based on Climate (left), Climate + forest cover (middle), and Climate + Lidar (right) data *for E.arvense, H.alpinum, T.officinale, and V.vitis-idea*

This result implies that the addition of Landsat-scale variables on land cover and forest cover do not add significant power to the understory models in this area. When LiDAR-derived canopy and terrain attributes were added to the models, variable selection changes markedly. A large number of LiDAR-derived variables were selected as important in model prediction. Wet area information, derived from the LiDAR DEM, was the most common variable added into the models, selected in 6 of the 14 species models. This was followed by the annual terrain-intercepted radiation (which is indicative of radiation regimes imposed by terrain) and then elevation. The addition of these three variables indicates the importance of higher spatial resolution in terrain patterns as it affected species distribution models. In response to these additions related to terrain attributes, there was a reduction in the importance of climate variables including frost free period and degree days while annual precipitation remained critical to model predictions. Of the LiDAR-derived canopy attributes, height and the solar regime of the canopy (i.e., shading of sites based on canopy cover and canopy gaps) were selected most often.

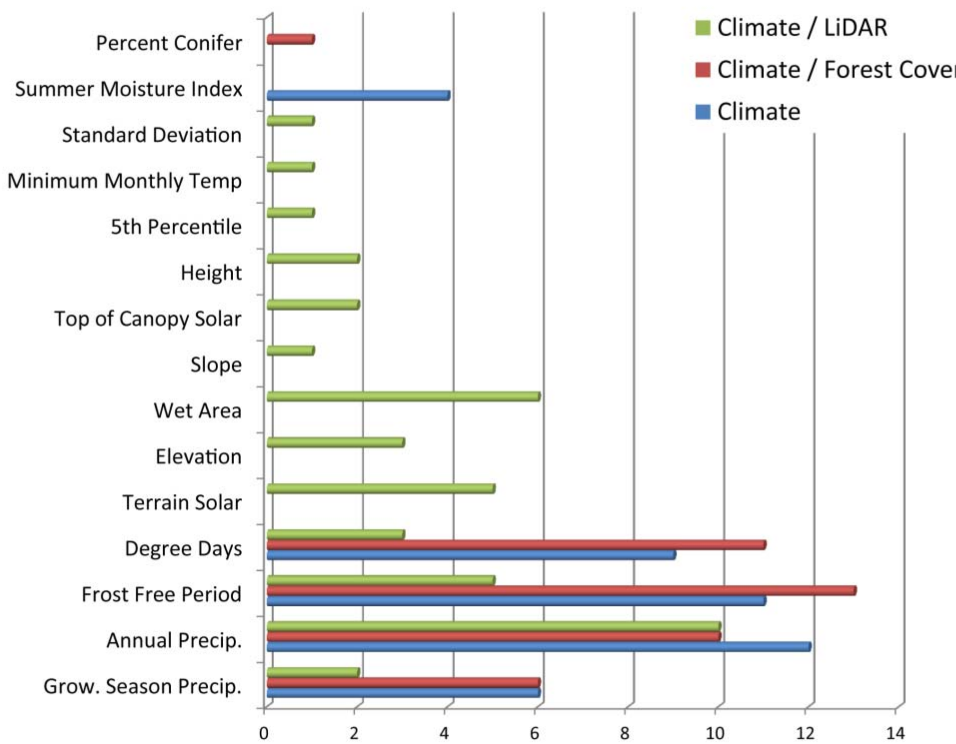

Figure 3: Frequency of variables selected in the top-three predictors for each species for the three model suites.

The individual response graphs of the most important variables of the combined models for 4 species are shown in Figure 4 and indicate the relationship between species occurrence and environmental drivers. For *E. arvense,* it is apparent that species occurrence is driven by presence of wet areas within the landscape at lower elevations. The *H. alpinum* model did not incorporate any LiDAR-derived terrain or canopy information and had a bi-modal response for degree days, reflecting its occurrence in cold, high elevation meadows, and warmer low elevation sites in stream valleys. *T. officinale* is predicted to occur in sites with longer frost free periods and lower mean canopy height, predominantly in sites having vegetation cover less than 5m in height. Lastly, *V. vitis-idea* occurs in sites with intermediate frost free period lengths and in cooler, lower elevation sites.

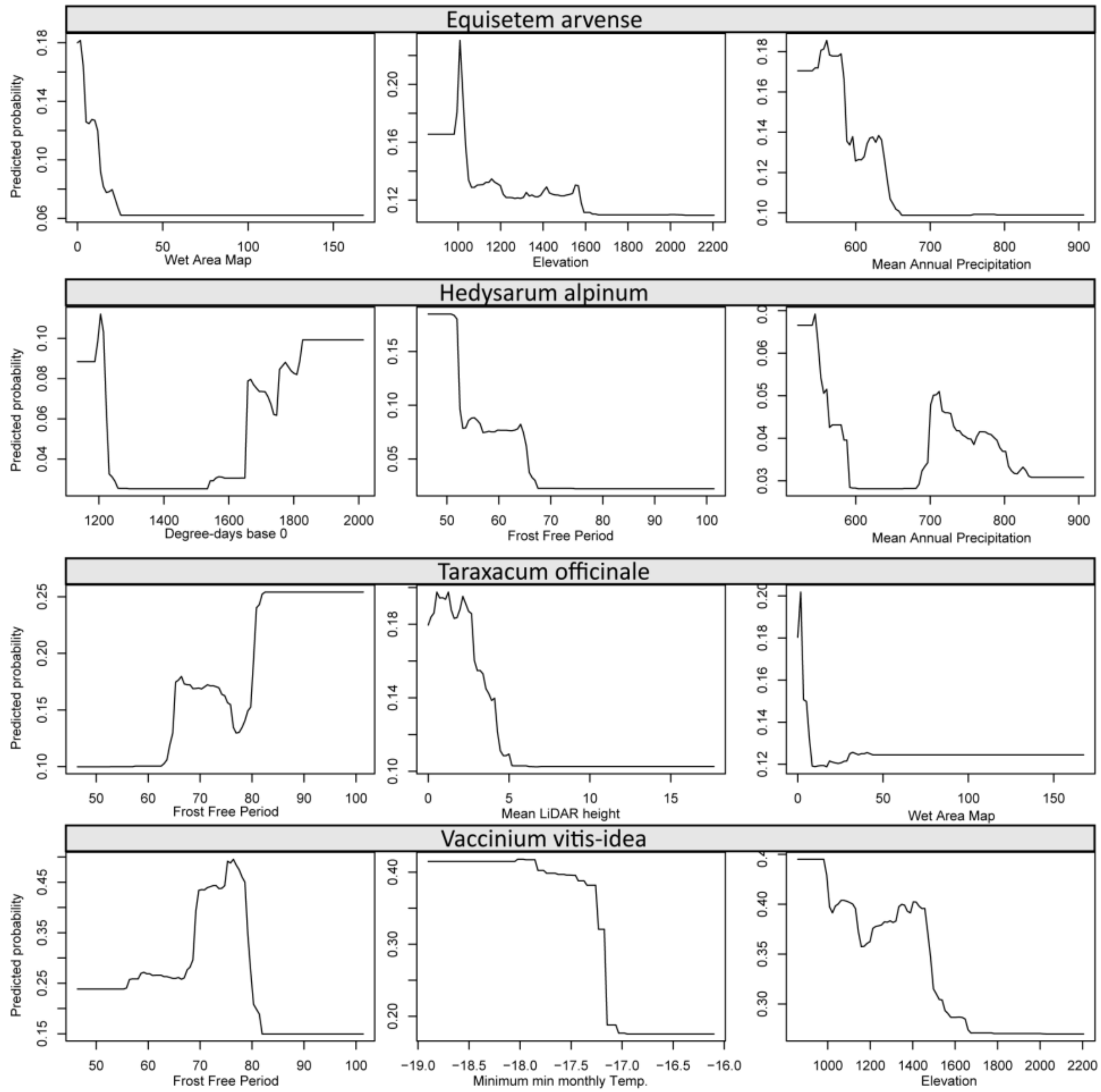

Figure 4: Response graphs of the first three most important variables for the Climate + Lidar model for *E.arvense, H.alpinum, T.officinale* and *V.vitis-idea.*

4. Discussion

In this study we investigated the added benefit of incorporating LiDAR-derived terrain and forest canopy information into understory species models relevant for grizzly bear species habitat modelling. Our use of boosted regression trees for model development enabled the combination of multiple data types as well as the inclusion of complex relationships which are often not possible to represent within standard linear models. Boosted regression trees allow representation of the main variables used in the developed models, as well as response graphs between individual plant occurrence and the most important variables. The past five years has seen these models increasingly selected in ecological research because of a number of features, including an ability to deal with collinear datasets, to exclude insignificant variables, and to allow for asymmetrical distribution of samples (De'Ath, 2002; Melendez et al., 2006; Schwalm et al., 2006). We recognize that a limitation of boosted regression approaches is that many observations are required for reliable model building, making model development of rare and more localised understory species more problematic, and should be undertaken with caution (Coops et al., 2011).

In addition, we are cognisant of potential issues surrounding the quality of both remote sensing and plot observation data given that the datasets were acquired across multiple years. Most of the LIDAR acquisitions occurred within a 5 year difference window, however in a small number of cases temporal gaps of up to 7 years may exist. As indicated in the study area description the relationship between stand dominant height and age for common pine species in the area (Cieszewski and Bella, 1989; Kirby, 1975) indicates that the relative height difference between stands older than 50 years is markedly lower than a height difference between younger stands. Similar effects can be observed for stand volume per hectare (Tait et al., 1988). For example, theoretical dominant height difference between two pine dominated stands in the area, between ages of 80 and 100, and site index of 25, is 6.7%. Similar difference calculated for stands with ages 20 and 40 is 36.8% (Cieszewski, 1991). As a result over the 5 year time frame between plot data measurement and LIDAR data acquisition we anticipate in areas of no disturbance overall structural conditions to have remained relatively constant.

In areas of recent harvest or fire, there may be marked discrepancies between the plot conditions prior to the disturbance and the LIDAR data acquired post disturbance. Using disturbance data on fire, harvesting, and resource extraction from the temporal Landsat time series we were able to detect stands which have had marked disturbance over the 8 year study period, and removed these sites from the analysis (Hilker et al., 2009). However, some issues may remain with LiDAR and field observations not directly coinciding.

The focus of this study is on application of LiDAR data in a large area ecological modelling study. To accommodate the size of the study area, and the level of pre-processing of the LIDAR data prior to analysis, detailled analysis of forest structure using dense LiDAR point clouds below 2m was not possible. We believe what is lost in our ability to derive understorey structure directory, is gained by the large extent of the dataset and to demonstrate how it can be applied to large area projects.

5. Conclusions

In this paper we demonstrate that plant distribution models developed with a combination of both broad-scale climate data, as well as with LiDAR-derived terrain and canopy information, provided the best overall performance, capturing more fine scale spatial variation than models using climate data alone. The inclusion of the LiDAR attributes suggest that these variables provide a more detailed explanation of the fine scale site conditions, such as access to water, solar radiation regime at the site caused by terrain shading, in addition to overall site elevation and slope (White et al., 2012). Information on canopy height, gaps, shading, and height variations also appear to affect distributions for some species but to a lesser degree than the finer site condition measured by LiDAR (Figure 5). The inclusion of site level measures from LiDAR resulted in a reduction of importance of growing degree days and frost free periods. This shift implies that the inclusion of LiDAR data allows a more comprehensive description of the thermal and radiation regimes of individual sites, replacing the need for broader scale descriptions of the thermal load of each site.

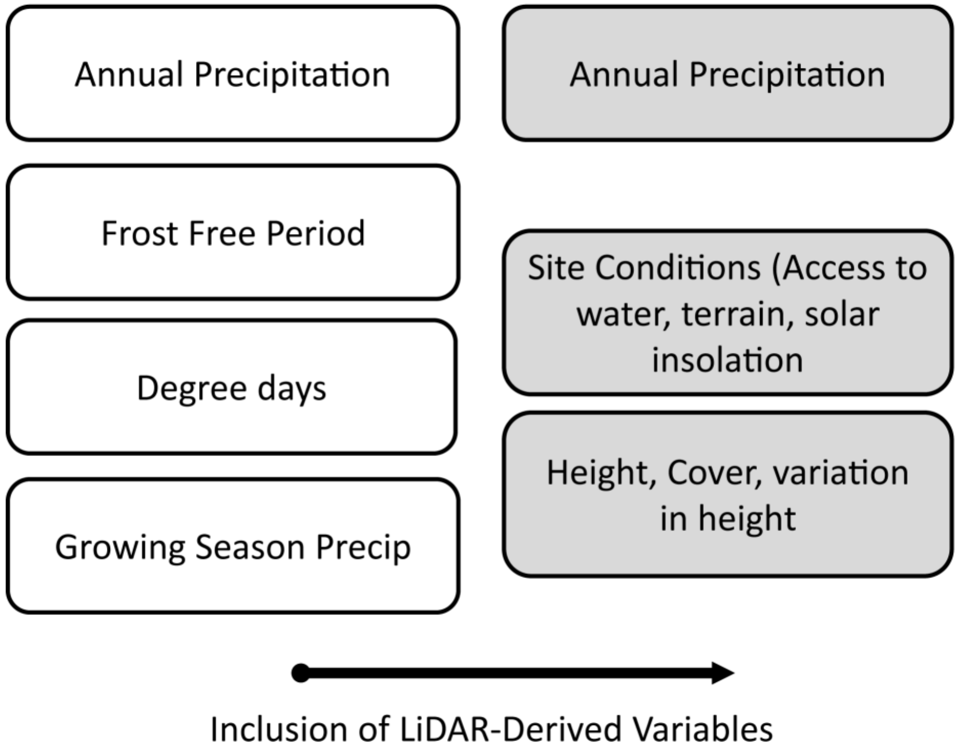

Figure 5: Change in Variables selected by models when incorporating fine scale site and canopy LiDAR derived information

Acknowledgements

Funding for this research was generously provided by the Grizzly Bear Program of the Foothills Research Institute located in Hinton, Alberta, Canada, with additional information available at: www.foothillsri.ca. Additional funds were provided by an NSERC Discovery grants to Coops and Nielson. LiDAR data was generously provided by the Government of Alberta.

References

Austin, M., 1985. CONTINUUM CONCEPT , ORDINATION METHODS , AND NICHE THEORY. Annu. Rev. Ecol. Syst. 16, 39–61.

Benn, B., Herrero, S., 2002. Grizzly bear mortality and human access in Banff and Yoho National Parks, 1971-98. Ursus 13, 213–221.

Cieszewski, C.J., 1991. Polymorphic height and site index curves for the major tree species in Alberta. For. Canada, Northwest Reg. North. For. Centre, Edmonton, Alberta. For. Manag. Note 51.

Cieszewski, C.J., Bella, I.E., 1989. Polymorphic height and site index curves for lodgepole pine in Alberta. Can. J. For. Res. 19, 1151–1160.

Clark, T.W., Paquet, P.C., Curlee, A.P., 1996. Special section: large carnivore conservation in the Rocky Mountains of the United States and Canada. Conserv. Biol. 10, 936–939.

Coops, N.C., Duffe, J., Koot, C., 2010. Assessing the utility of lidar remote sensing technology to identify mule deer winter habitat. Can. J. Remote Sens. 36, 81–88.

Coops, N.C., Waring, R.H., Beier, C., Roy-Jauvin, R., Wang, T., 2011. Modeling the occurrence of 15 coniferous tree species throughout the Pacific Northwest of North America using a hybrid approach of a generic process-based growth model and decision tree analysis. Appl. Veg. Sci. 14, 402–414.

Daly, C., Neilson, R.P., Phillips, D.L., 1993. A statistical-topographic model for mapping climatological precipitation over. J. Appl. Meteorol. 33, 140–158.

Davenport, I.J., Holden, N., Gurney, R.J., 2004. Characterizing errors in airborne laser altimetry data to extract soil roughness. IEEE Trans. Geosci. Remote Sens. 42, 2130–2141.

De'Ath, G., 2002. Multivariate regression trees: a new technique for modeling species-environment relationships. Ecology 83, 1105–1117.

Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees. J. Anim. Ecol. 77, 802–13.

Ferster, C., Coops, N.C., Trofymow, J., 2009. Aboveground large tree mass estimation in a coastal forest in British Columbia using plot-level metrics and individual tree detection from lidar. Can. J. Remote Sens. 35, 270–275.

Friedman, J., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 29, 1189–1232.

Friedman, J.H., 2002. Stochastic gradient boosting. Comput. Stat. Data Anal. 38, 367–378.

Fu, P., Rich, P.M., 2002. A geometric solar radiation model with applications in agriculture and forestry. Comput. Electron. Agric. 37, 25–35.

Garnier, B.J., Ohmura, A., 1968. A method of calculating direct shortwave radiation income of slopes. J. Appl. Meteorol. 7, 796–800.

Gaveau, D.L.A., Hill, R.A., 2003. Quantifying canopy height underestimation by laser pulse penetration in small-footprint airborne laser scanning data. Can. J. Remote Sens. 29, 650–657.

Gordon, I.J., Hester, A.J., Festa-Bianchet, M., 2004. REVIEW: The management of wild large herbivores to meet economic, conservation and environmental objectives. J. Appl. Ecol. 41, 1021–1031.

Hilker, T., Wulder, M.A., Coops, N.C., Linke, J., McDermid, G., Masek, J.G., Gao, F., White, J.C., 2009. A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on Landsat and MODIS. Remote Sens. Environ. 113, 1613–1627.

Iverson, L.R., Prasad, A.M., 1998. Predicting abundance of 80 tree species following climate change in the eastern United States. Ecol. Monogr. 68, 465–485.

Iverson, L.R., Prasad, A.M., 2001. Potential changes in tree species richness and forest community types following climate change. Ecosystems 4, 186–199.

Jiménez-Valverde, A., 2011. Relationship between local population density and environ - mental suitability estimated from occurrence data. Front. Biogeogr. 3, 59–61.

Kirby, C.L., 1975. Site index equations for lodgepole pine and white spruce in Alberta. Inf. Rep. NOR-X-142.

Landis, J.R., Koch, G.G., 1977. An application of hierarchical Kappa-type statistics in the assessment of majority agreement among multiple observers. Biometrics 33, 363–374.

Lee, D.S., Storey, J.C., Choate, M.J., Hayes, R.W., 2004. Four years of Landsat-7 on-orbit geometric calibration and performance. IEEE Trans. Geosci. Remote Sens. 42, 2786–2795.

Linderman, M., Bearer, S., An, L., Tan, Y., Ouyang, Z., Liu, J., 2005. The effects of understory bamboo on broad-scale estimates of giant panda habitat. Biol. Conserv. 121, 383–390.

Lindzey, F., Meslow, E.C., 1977. Home range and habitat use by black bears in southwestern Washington. J. Wildl. Manage. 41, 413–425.

MacArthur, R., MacArthur, J., 1961. On bird species diversity. Ecology 42, 594–598.

Maltamo, M., Packalen, P., Yu, X., Eerikainen, K., Hyyppa, J., Pitkanen, J., 2005. Identifying and quantifying structural characteristics of heterogeneous boreal forests using laser scanner data. For. Ecol. Manage. 216, 41–50.

Manel, S., Williams, H.C., Ormerod, S.J., 2001. Evaluating presence-absence models in ecology: the need to account for prevalence. J. Appl. Ecol. 38, 921–931.

Martin, P., 1983. Factors influencing globe huckleberry fruit production in northwestern Montana, in: Bears: Their Biology and Management, Vol. 5, A Selection of Papers from the Fifth International Conference on Bear Research and Management. pp. 159–165.

Martinuzzi, S., Vierling, L.A., Gould, W. a., Falkowski, M.J., Evans, J.S., Hudak, A.T., Vierling, K.T., 2009. Mapping snags and understory shrubs for a LiDAR-based assessment of wildlife habitat suitability. Remote Sens. Environ. 113, 2533–2546.

Mcdermid, G.J., 2005. Remote sensing for large-area, multi-jurisdictional habitat mapping.

McDermid, G.J., Hall, R.J., Sanchez-Azofeifa, G. a., Franklin, S.E., Stenhouse, G.B., Kobliuk, T., LeDrew, E.F., 2009. Remote sensing and forest inventory for wildlife habitat assessment. For. Ecol. Manage. 257, 2262–2269.

Mckenzie, D., Peterson, D.W., Peterson, D.L., Thornton, P.E., Northwest, P., Station, R., Forestry, W., Division, G.D., 2003. Climatic and biophysical controls on conifer species distributions in mountain forests of Washington State, USA. J. Biogeogr. 30, 1093–1108.

Means, J.E., Acker, S.A., Fitt, B.J., Renslow, M., Emerson, L., Hendrix, C.J., 2000. Predicting forest stand characteristics with airborne scanning Lidar. Photogramm. Eng. Remote Sens. 66, 1367–1371.

Melendez, K. V, Jones, D.L., Feng, A.S., 2006. Classification of communication signals of the little brown bat. J. Acoust. Soc. Am. 120, 1095.

Monserud, R.A., Leemans, R., 1992. Comparing global vegetation maps with the Kappa statistic. Ecol. Modell. 62, 275–293.

Munro, R.H.M., Nielsen, S., Price, M.H., Stenhouse, G.B., Boyce, M., 2006. Seasonal and Diel Patterns of Grizzly Bear Diet and Activity in West-Central Alberta. J. Mammal. 87, 1112–1121.

Nielsen, S., Boyce, M., Stenhouse, G.B., 2004. Grizzly bears and forestry I. Selection of clearcuts by grizzly bears in west-central Alberta. For. Ecol. Manage. 199, 51–65.

Nielsen, S., Boyce, M., Stenhouse, G.B., Munro, R.H.M., 2003. Development and testing of phenologically driven grizzly bear habitat models. Ecoscience 10, 1–10.

Nielsen, S., Graham, K., Larsen, T., Mckay, T., Munro, R., 2010. Grizzly bear habitat productivity models for the Yellowhead, Swan Hills, Grande Cache and Chinchaga population units of Alberta, Methods.

Nielsen, S., Johnson, C.J., Heard, D.C., Boyce, M.S., 2005. Can models of presence-absence be used to scale abundance? Two case studies considering extremes in life history. Ecography (Cop.). 28, 197–208.

Nielsen, S., Stenhouse, G.B., Boyce, M., 2006. A habitat-based framework for grizzly bear conservation in Alberta. Biol. Conserv. 130, 217–229.

Nielsen, S.E., Munro, R.H.M., Bainbridge, E.L., Stenhouse, G.B., Boyce, M.S., 2004a. Grizzly bears and forestry: II. Distribution of grizzlybear foods in clearcuts of west-central Alberta, Canada. For. Ecol. Manage. 199, 67–82.

Nielsen, S.E., Munro, R.H.M., Bainbridge, E.L., Stenhouse, G.B., Boyce, M.S., 2004b. Grizzlybears and forestry: II. Distribution of grizzlybear foods in clearcuts of west-central Alberta, Canada. For. Ecol. Manage. 199, 67–82.

Nielsen, S.E., Stenhouse, G.B., Beyer, H.L., Huettmann, F., Boyce, M.S., 2008. Can natural disturbance-based forestry rescue a declining population of grizzly bears? Biol. Conserv. 141, 2193–2207.

Pontius, R.G., Millones, M., 2011. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. Int. J. Remote Sens. 32, 4407–4429.

R Development Core Team, 2013. R: A language and environment for statistical computing.

Riaño, D., Meier, E., Allgöwer, B., Chuvieco, E., Ustin, S.L., 2003. Modeling airborne laser scanning data for the spatial generation of critical forest parameters in fire behavior modeling. Remote Sens. Environ. 86, 177–186.

Rich, P., Dubayah, R., Hetrick, W., Saving, S., 1994. Using viewshed models to calculate intercepted solar radiation: applications in ecology.

Roughgarden, J., Running, S., Matson, P., 1991. What does remote sensing do for ecology? Ecology 72, 1918–1922.

Schwalm, C.R., Black, T.A., Amiro, B., Arain, M.A., Barr, A.G., Bourque, C.P., Dunn, A.L., Flanagan, L.B., Giasson, M.A., Lafleur, P.M., Margolis, H.A., McCaughey, J.H., Orchansky, A.L., Wofsy, S.C., 2006. Photosynthetic light use efficiency of three biomes across an east-west continental-scale transect in Canada. Agric. For. Meteorol. 140, 269–286.

Schwartz, C.C., Miller, S.D., Haroldson, M.A., 2003. Grizzly Bear 556–586.

Seielstad, C.A., Queen, L.P., 2003. Using airborne laser altimetry to determine fuel models for estimating fire behaviour. J. For. 101, 10−17–10−17.

Swets, J. a, 1988. Measuring the accuracy of diagnostic systems. Science 240, 1285–93.

Tait, D.E., Cieszewski, C.J., Bella, I.E., 1988. The stand dynamics of lodgepole pine. Can. J. Remote Sens. 18, 1255–1260.

Thuiller, W., Albert, C., Araújo, M.B., Berry, P.M., Cabeza, M., Guisan, A., Hickler, T., Midgley, G.F., Paterson, J., Schurr, F.M., Sykes, M.T., Zimmermann, N.E., 2008. Predicting global change impacts on plant species' distributions: future challenges. Perspect. Plant Ecol. Evol. Syst. 9, 137–152.

Wang, T., Hamann, A., Spittlehouse, D.L., Murdock, T.Q., 2012. ClimateWNA—High-Resolution Spatial Climate Data for Western North America. J. Appl. Meteorol. Climatol. 51, 16–29.

White, B., Ogilvie, J., Campbell, D.M.H.M.H., Hiltz, D., Gauthier, B., Chisholm, H.K.H., Wen, H.K., Murphy, P.N.C.N.C., Arp, P. a. a., 2012. Using the Cartographic Depth-to-Water Index to Locate Small Streams and Associated Wet Areas across Landscapes. Can. Water Resour. J. 37, 333–347.

Wilkinson, L., Stenhouse, G.B., Al, E., 2008. Alberta Grizzly Bear Alberta Grizzly Bear Recovery Plan 2008-2013.

Wing, B.M., Ritchie, M.W., Boston, K., Cohen, W.B., Gitelman, A., Olsen, M.J., 2012. Prediction of understory vegetation cover with airborne lidar in an interior ponderosa pine forest. Remote Sens. Environ. 124, 730–741.

Wulder, M.A., Bater, C.C.W., Coops, N.C., Hilker, T., White, J.C., 2008. The role of LiDAR in sustainable forest management. For. Chron. 84, 807–826.

Zager, P., Jonkel, C., Habeck, J., 1983. Logging and Wildfire Influence on Grizzly Bear Habitat in Northwestern Montana, in: Bears : Their Biology and Management , Vol . 5 , A Selection of Papers from the Fifth International Conference on Bear Research and Managament. pp. 124–132.