Forest Ecology and Management 372 (2016) 19–27

Contents lists available at ScienceDirect

Forest Ecology and Management

journal homepage: www.elsevier.com/locate/foreco

Using forest structure to predict the distribution of treed boreal peatlands in Canada

Dan K. Thompson [a,*], Brian N. Simpson [a], André Beaudoin [b]

[a] Natural Resources Canada, Canadian Forest Service, Northern Forestry Centre, Edmonton, Alberta T5H 1Y7, Canada
[b] Natural Resources Canada, Canadian Forest Service, Laurentian Forestry Centre, Quebec City, Québec G1V 4C7, Canada

ARTICLE INFO

Article history:
Received 29 December 2015
Received in revised form 29 March 2016
Accepted 30 March 2016

Keywords:
Boreal
Peatland
Distribution mapping
Organic soil depth
Boosted regression trees
Logistic regression

ABSTRACT

Mapping peatland extent in Canada would contribute important information concerning carbon balance and hydrology. While such mapping, based on air photo interpretation and remote sensing data, has recently improved, maps have been limited to 1:1 million scale. We hypothesized that forest structure information from forest inventory plots could be used to predict the presence of forested and treed peatlands in boreal Canada at the ground plot-level, and that a resulting model could be used to predict the distribution of forested and treed peatlands across Canada. Inventory ground plots from the Canadian National Forest Inventory (NFI) with organic soil depth measurements were used to create a model of the presence of treed to forested (canopy cover ranging from sparse to closed) peatlands (greater than 40 cm organic soil depth) in boreal Canada. The presence of black spruce (*Picea mariana*) or larch (*Larix laricina*), in combination with low stand height and stand age greater than 75 years, were the strongest predictors of the presence of peatlands. Bioclimatic variables related to high diurnal and annual temperature variation, consistent with a continental climate, also contributed to the increased predicted presence of treed peatlands. Both logistic and boosted regression tree models showed similar results, with ~87% accuracy in the discrimination of treed peatlands when validated against an independent set of ground plots. The boosted regression tree model was propagated across Canada using forest attribute raster data layers at 250 m resolution from the NFI along with bioclimatic layers. Estimates of treed peatland extent agreed with data points from peat cores with 85–95% accuracy in the Boreal Shield ecozone, although prediction was less accurate in the more southern boreal and Great Lakes forest areas. The resulting map can be used as an input to forest carbon modelling, and the improved knowledge of treed peatland extent will be useful in modelling wildfire or peatland drainage.

1. Introduction

Peatlands represent hotspots of carbon (C) storage in the North American boreal forest, covering 110 Mha of land with local accumulations exceeding 200 kg C m^{-2}. Peatland soils are defined as organic deposits greater than 40 cm in depth, which corresponds to the maximum rooting depth of most boreal trees (Canada Soil Survey Committee, 1978). These organic soils cover approximately 22% of the land area in Canada's boreal and subarctic regions; 97% of all peatland areas in Canada fall within these boreal and subarctic regions (Tarnocai et al., 2011). Densely forested peatlands, with a closed tree canopy and tree canopy height greater than 5 m, constitute approximately 17% of the peatland area in Canada (Zoltai and Martikainen, 1996). In such densely forested peatlands, the higher wood volume allows for wood fibre extraction (Zoltai and Martikainen, 1996), though peatland drainage and active forest management of peatlands in Canada has never extended beyond a few limited trials (Haavisto et al., 1995). However, trees also play a major role in the C balance (Wieder et al., 2009), ecology (Miller et al., 2015), and hydrology (Kettridge et al., 2013) of open-canopy treed peatlands where tree biomass is too small to economically harvest. Because of the low economic value of these peatlands, they largely remain unmapped except when found adjacent to timber extraction areas in the southern portion of the Canadian boreal forest. In Canada, many of these open treed peatlands are classified as forested lands under the Marrakesh Accord, in which a threshold of 25% crown closure and potential tree height of 5 m is used to define forested lands (UNFCC, 2002). Here, we use the term "treed" to refer to peatlands that have, at a minimum, the density of a treed peatland, but also up to and including closed-canopy densely forested peatlands.

---

[*] Corresponding author.
E-mail addresses: daniel.thompson@canada.ca (D.K. Thompson), brian.simpson@canada.ca (B.N. Simpson), andre.beaudoin@canada.ca (A. Beaudoin).

http://dx.doi.org/10.1016/j.foreco.2016.03.056
0378-1127/

Mapping peatland extent in Canada originated in ecological inventory efforts, principally through the Canada Land Inventory (Pratt, 1965). Early efforts were largely limited to polygon-mapping products of 1:1 million or greater scale, although there were attempts to characterize the abundance of differing peatland types regionally (Zoltai et al., 1975). A more exact quantification of peatland area and C stock was required by the 1990s as early C budgets of northern peatlands were developed (Gorham, 1991). This effort was quickly followed by research on plot-level carbon cycling in boreal peatlands (e.g., Hogg et al., 1992; Moore and Dalva, 1993; Frolking et al., 1998).

Polygon-mapping products depicting proportional peatland area of differing classes in Canada at the 1:1 million scale have been available for some time (e.g., Tarnocai et al., 1995, 2011) and were based on air photo interpretation and manual delineation of LANDSAT imagery (Lacelle, 1998). More detailed analyses were available for selected locations, such as north-central Alberta (Vitt et al., 1995). Air photo interpretation of forested peatlands relies on two surface expressions of peatlands: forest structure and landform (Zoltai and Vitt, 1995). Forest attributes (e.g., tree height, density, volume), particularly when they indicate an open, short, conifer canopy, were used to delineate forested peatlands from upland conifer stands (Zoltai and Vitt, 1995). In addition to forest structure, landforms such as permafrost palsas, strings, flarks, and other features visible in aerial photography (approximately 100 m or larger in size) assist in the manual identification of peatlands (Zoltai and Vitt, 1995). Remote sensing approaches, including optical (e.g., Palylyk and Crown, 1984), synthetic aperture radar (Touzi et al., 2007), and composite methods involving radar and optical (Li and Chen, 2005) offer high accuracy (>80%) of peatland delineation over small spatial areas. The delineation of large, open, treeless peatlands in Canada has been available for some via the Earth Observation for Sustainable Development (EOSD) landcover product (Wulder et al., 2008). The identification of smaller, more heavily forested peatlands with little or no patterning is more difficult, as such peatlands resemble upland conifer forests. Given that peatland disturbances such as wildfire are more severe in more heavily forested peatlands (Lukenbach et al., 2015) that are associated with shallow peat (Bhatti et al., 2006), the efficient identification of these more hidden peatland areas becomes even more important for C accounting. Contemporary C accounting practices and models increasingly require high-resolution maps of soil C loads in order to accurately model emissions from disturbances such as fire (Anderson et al., 2015).

Peat depth has rarely been measured in a systematic way in Canada; notable datasets of peat depth from peat coring studies by Zoltai et al. (2000) and Riley and Michaud (1989) lack information on forest structure and only provide a cover percentage by tree species. Peat depth measurements are available alongside metrics of forest structure from the Canadian National Forest Inventory (NFI), a national network of aerial and ground forest inventory plots, spanning all provinces, territories, and forested ecozones. While not as numerous as provincial silviculture permanent sample plots that measure large merchantable trees, the NFI plots offer the advantage of representing both merchantable and non-merchantable forest stands with a harmonized methodology across Canada. The ground plots available through the NFI (913 in total to date) include at least one soil pit measuring total organic soil depth (OSD) down to 100 cm. Compared with PSP networks, the NFI dataset has the additional advantage that the deep OSD (in the form of peat) is actually measured, as are the attributes of trees as small as 1.3 m in height, which are common in forested peatlands.

Recently, forest attribute data derived from a combination of manually interpreted aerial photography and modelling, available from the 2 × 2 km photo plot database, have been interpolated between photo-plot sites using optical remote sensing data from the Moderate Resolution Imaging Spectrometer (MODIS), using a *k*-nearest neighbour (*k*-NN) interpolation approach for imagery from 2001 (Beaudoin et al., 2014). The resulting data products are 250 m resolution raster maps of numerous forest canopy attributes available from the NFI photo plots, such as stand height, merchantable volume, proportion of biomass by species, and crown closure. Measurements exclusive to ground plots, such as OSD, were therefore not included in the analysis. However, some of the forest attributes, such as open-crown, short, pure conifer stands, on the raster maps produced by Beaudoin et al. (2014) have been shown to be indicators of the presence of boreal forested peatlands (Zoltai and Vitt, 1995). This combination of a comprehensive ground plot methodology relevant to forested peatlands, coupled with complete national coverage of forest attribute layers serving to predict peatland presence, can be used to produce a spatial presence–absence model, in the same way as point count datasets for birds (Venier et al., 2004) or plants (Ohse et al., 2009) have been used.

We hypothesized that (i) forest structure information from forest inventory plots could be used to predict the presence of forested and treed peatlands in boreal Canada at the ground plot-level, and that (ii) a subset of the forest structure variables available from both ground plots and national raster datasets could be used to calibrate such a model, which could then be spatialized using the raster datasets to predict the distribution of forested and treed peatlands across Canada. Therefore, the objectives of this study were to (1) create models predicting the presence of forested peatlands using NFI ground inventory plots for which forest, soil, and climatic attributes are available; (2) spatialize these models across targeted ecozones of Canada using available national raster maps; and (3) assess the accuracy of output maps using independent validation sets of ground plots.

2. Methods

2.1. Area and input data

The spatial predictions targeted all boreal and taiga regions of Canada (Ecological Stratification Working Group, 1996) where forested peatlands are a significant landscape feature (including Taiga Shield, Taiga Plains, Hudson Plains, and Boreal Shield ecozones; and Great Lakes–St. Lawrence, Mid-Boreal Shield, Lake of the Woods, and Southern Boreal Shield ecoprovinces, the last three being subdivisions of the Boreal Shield ecozone; Fig. 1). Four hundred and fifty NFI ground plots (Gillis et al., 2005) located within these ecozones or ecoprovinces were used as input data for model calibration. Only variables available in both the NFI ground plots (400 m^2 area) and the photo plots used in the creation of the *k*-NN dataset were used in constructing the model. If data on tree cover were missing or no trees were measured in the ground plot (20 × 20 m), sites were excluded. This step eliminated treeless peatlands, such as those with only moss, sedge, and shrub cover. Sites were classified as peatlands if they had a reported OSD greater than 37.5 cm, following the 40 cm threshold used in the Canadian System of Soil Classification (Canada Soil Survey Committee, 1978). The 37.5 cm definition allows for a marginal thickness of moss cover in addition to peat cover, as the NFI ground plots group all duff, litter, and peat in measuring OSD thickness.

2.2. Model creation

Two models predicting the presence or absence of OSD greater than 40 cm were calibrated and compared to evaluate their relative merits and limitations. First, a logistic regression model of forested

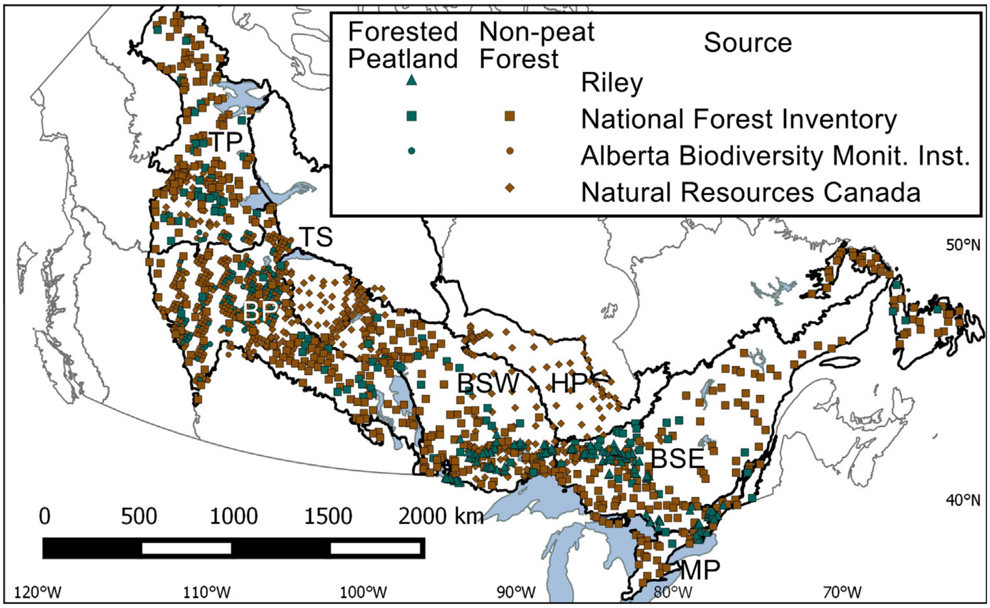

Fig. 1. Locations of ground plots used for model creation and validation (National Forest Inventory and Alberta Biodiversity Monitoring Institute plots), as well as for spatial model validation (Riley and Michaud and Natural Resources Canada control points). Note that the Riley and Michaud data indicate presence only while the Natural Resources Canada data show absence only. For details on datasets used in model construction, refer to Table 1. Ecozones, defined by the Ecological Stratification Working Group (1996), are labelled as follows: TP – Taiga Plains, TS – Taiga Shield, BP – Boreal Plains, BSW – Boreal Shield West, HP – Hudson Plains, BSE – Boreal Shield East, MP – Mixedwood Plains.

peatland presence–absence from forest structure attributes (Table 1) was created using the base *stats* package in the *R* statistical programming language (version 3.1.2). A parsimonious model with only linear terms and without interactions was chosen for maximum reproducibility. Second, a boosted regression tree (BRT) model was created using the *dismo* package version 1.0-5 (Hijmans et al., 2015). Both models used the same input dataset. The *dismo* package defaults (bag fraction of 0.5 with a learning rate of 0.05) were used to parameterize the BRT model calibration. The BRT was limited to the optimal number of boosting trees using *k*-fold cross-validation using the *gbm.step* (Elith et al., 2008) function in the *dismo* package. The model was trimmed to include only variables explaining more than 10% of the variance. The logistic regression model is used here as a baseline model against which the more complex BRT model can be compared against.

Spatial predictions were made using the BRT and rasters of the same independent variables as listed in Table 1. Forest structure information at 250 m resolution was used from the *k*-NN dataset of Beaudoin et al. (2014; Fig. 2). Additionally, bioclimatic variables (Table 1) from ANUSPLIN at 2.5 arc-minutes were resampled to the *k*-NN grid using the WorldClim dataset (Hijmans et al., 2005). Slope was calculated from the 90 m EarthEnv digital elevation model (DEM; Robinson et al., 2014; Table 1) and also resampled to 250 m. A matrix of Spearman correlation coefficients (ρ) was computed to ensure that no paired input variables were highly correlated (absolute ρ values below 0.4 confirmed low correlation; Table 2).

Table 1
Predictive variables used as input in the prediction of peat horizon presence at the ground plot scale.

Variable name	Description
	Biophysical[a]
Larilar_prop	Proportion of total aboveground biomass as *Larix laricina*[b]
Picemar_prop	Proportion of aboveground biomass as *Picea mariana*[b]
Site_age	Mean age of dominant and co-dominant trees (years)
Site_height	Mean height of dominant and co-dominant trees (m)
Slope	Ground surface slope (%)
	Bioclimatic[c]
BIO2	Mean diurnal range; mean of monthly max temp–min temp (°C)
BIO4	Seasonality; standard deviation of monthly temperature (°C)

[a] From the NFI ground plot database (Gillis et al., 2005).
[b] Expressed as a proportion from 0 to 1, rather than as an integer percentage.
[c] Sampled from v1.4 of the *WorldClim* database (Hijmans et al., 2005).

2.3. Model validation

2.3.1. Ground plot validation

Both ground plot models and spatialized predictions based on the BRT model were validated using various datasets including ecological inventory plots and geodetic survey markers (Table 3). For the ground plot models of peatland presence or absence, an independent dataset of 255 ecological inventory ground plots from the Alberta Biodiversity Monitoring Institute (ABMI, 2015) were used to validate the model's accuracy (Fig. 1). Tree lists from each ABMI ground plot site were supplemented with allometric models from Lambert et al. (2005) in order to calculate the percentage of total aboveground biomass by species, as defined in *k*-NN raster datasets. Site height for each plot was calculated as the Lorey height (Nakai et al., 2010; Lorey, 1878), as this was the best proxy measure for the biomass-weighted height from the *k*-NN rasters. Since the exact locations of the ground plots in the ABMI are not available, they were not used to validate any spatial predictions of the BRT model.

2.3.2. Spatial model validation

In addition to the validation of the ground plot model of OSD based on forest structure, two additional independent datasets were used for the validation of the spatial version of the BRT model. A dataset of peatland core locations from Riley and Michaud (1989) containing 126 peatlands from Ontario was used as an independent dataset of peatland locations. To validate non-peatland locations, passive control stations from the Canadian Spatial Reference System (Natural Resources Canada, 2015a) were

Table 2
Spearman correlation matrix of input variables used for model construction; data are from NFI ground plots ($n = 540$). See Table 1 for variable descriptions.

Variable name	*Slope*	*Site_age*	*Site_height*	*Picemar_prop*	*Larilar_prop*	*BIO2*	*BIO4*
Slope							
Site_age	0.02						
Site_height	**0.18**	−0.05					
Picemar_prop	**−0.14**	**0.30**	**−0.40**				
Larilar_prop	**−0.14**	**0.06**	**−0.23**	0.01			
BIO2	−0.06	**−0.23**	**0.24**	−0.02	0.02		
BIO4	**−0.19**	**−0.13**	**−0.25**	**0.22**	**0.14**	**0.25**	

Bold values are significant correlations at $p < 0.05$.

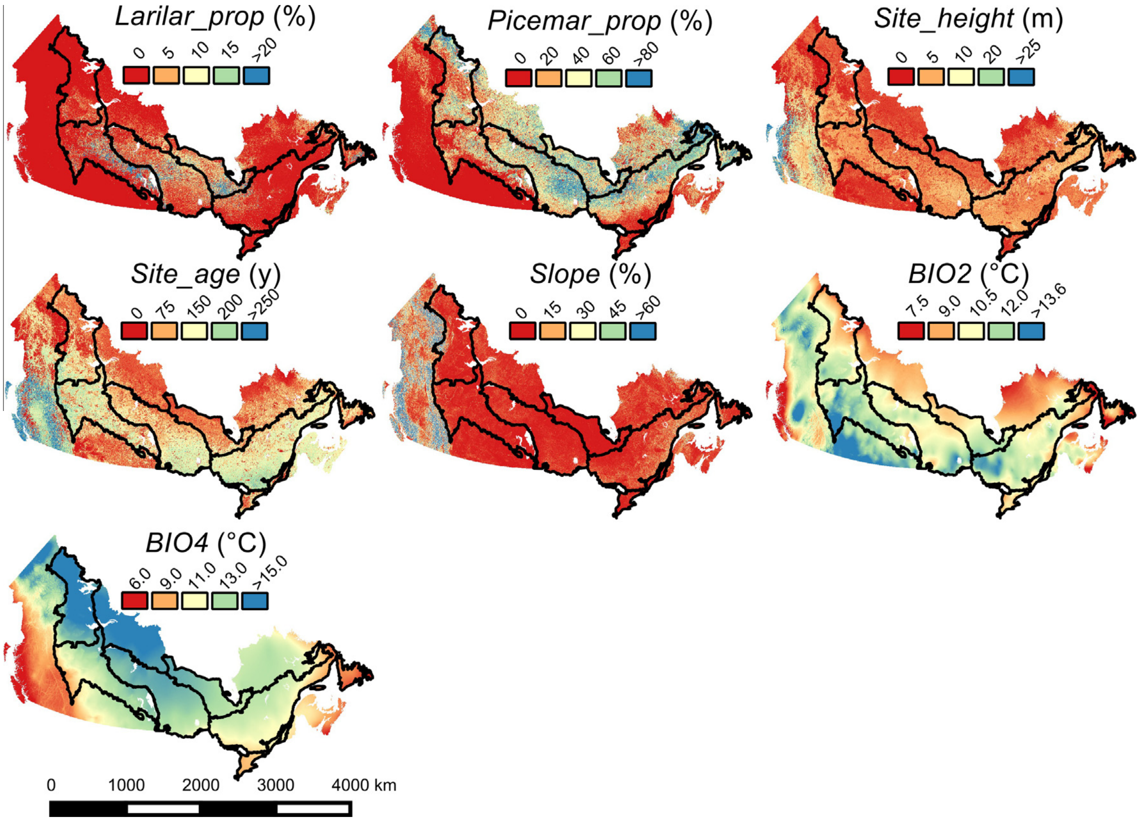

Fig. 2. Raster layers of predictive variables (see Table 1 for description of variables) at 250 m resolution used for the spatial implementation of a boosted regression tree model of forested peatland presence.

Table 3
Description of datasets used in this study.

Dataset	*n*	Forest structure info	Spatial accuracy	Pseudo-located[a]
National Forest Inventory	563	Yes	±5 km	Yes
Alberta Biodiversity Monitoring Institute	429	Yes	±5 km	Yes
Riley and Michaud (1989)	126	No	See note[b]	No
Natural Resources Canada	606	No	±1 m	No

[a] Pseudo-located data are artificially rounded to the nearest 5 km to protect the integrity of the plot, but represent a ground plot of approximately 400 m^2.

[b] Note that the location used for the Riley and Michaud dataset is the centroid of a manually mapped peatland polygon.

used. A total of 606 control points from the Taiga Plains, Boreal Shield and Hudson Plains ecozones were selected, keeping only those points that contained a reference to bedrock, sandy soils, or pine trees in the text description of the control point as contraindicators of peatland presence (Fig. 1). Validation of spatial predictions was performed within ecozones and their ecoprovince subdivisions that contained validation points. Additional qualitative validation was undertaken by comparing peatlands delineated by interpretation of aerial photography from the Alberta Geological Survey (Fenton et al., 2013) with those from the BRT method.

3. Results

3.1. Forested peatland presence modelling using ground plot data

3.1.1. Logistic regression model

The parsimonious logistic regression is shown in Table 4. The logistic regression explained 29.8% of the deviance, with an accuracy of 88.2% and an area under the receiver operating characteristic curve (AUC) of 0.931 (Table 5). The logistic model output threshold for peatland presence was 0.315, meaning that model outputs greater than the threshold value were classified as treed peatlands in a binary map. While the top predictors (*PICEMAR_prop*, *BIO2*, *Larilar_prop*, *slope*) were common to both the logistic regression and the BRT, in the logistic regression the variables *BIO4*, *site_age*, and *site_height* (see Table 2 for definitions of variables) were not statistically significant predictors (at $P < 0.05$) (Table 4).

3.1.2. Boosted regression tree model

The BRT for peatland presence based on NFI ground plot data explained 48.0% of the total variability, with an accuracy of 87.1% and an AUC of 0.934 (Table 5). The BRT model output threshold for peatland presence was 0.372. The strongest predictor in the model was the variable (*Picemar_prop*) for the proportion (0–1) of aboveground biomass consisting of black spruce (*Picea mariana* (Mill.) Britton, Sterns & Poggenburg). This predictor explained 24.4% of the model's predictive power (Fig. 3) and showed a roughly linear response through its marginal effect on the logit (log-odds) of a peatland's presence, with a crossover from a negative marginal effect (less likely to be a peatland) to a positive one (more likely to be a peatland) at a threshold where black spruce constituted a 0.5 proportion of aboveground biomass. The variable (*Larilar_prop*) describing the proportion of a plot's biomass consisting of larch (*Larix laricina* (Du Roi) K. Koch) was another strong predictor, accounting for 13.9% of the model's explanatory power. *Larilar_prop* showed a strong negative marginal effect when the proportion of biomass consisting of larch was less than 0.1 (Fig. 3), with a strong inflection point at 0.2 proportion of biomass, and no difference in the marginal effect from 0.4 upwards. *Site_age* and *site_height* contributed 14.7% and 12.8% to the model's explanatory power, respectively, and both showed a unimodal optimum-type relationship, with the greatest marginal effect at 175 years and 8 m stand height. Slope showed a strong positive marginal effect from 0% to 10% slope, with little effect afterwards, contributing 10.8% to the model's power. Mean diurnal range (*BIO2*) contributed 13.3% of the model's power, with the marginal effect switching from a negative to a positive at 12 °C. Monthly temperature standard deviation (*BIO4*) showed a bimodal effect, with peaks of positive marginal effect at 13.5 °C and at greater than 16 °C, and made the smallest contribution (10.1%) to model power. While *Picemar_prop* and *Larilar_prop* came out as strong predictors in both the logistic and BRT models, *site_height* was ranked as a stronger predictor in the BRT than in the logistic regression, while the opposite was true for *BIO2* and *slope*.

3.2. Spatial prediction based on BRT model

Based on the comparison of the models above, we selected the BRT model to conduct the spatial prediction of forested peatland presence across Canada (Fig. 4) using input raster data at 250 m resolution from the Beaudoin et al. (2014) dataset. Overall, the areas of the highest density of forested peatlands are in the southern Hudson Plains ecozone, with other areas of high density in the southern Boreal Plains and southern Taiga Shield ecozones. At the national scale, these areas of high density of forested peatlands correspond well with the 1:1 million scale mapping of Tarnocai et al. (2011). At the regional scale, the spatial predictions obtained with the BRT model correspond with larger peatlands identified via air photo interpretation in central Alberta, where the majority of peatlands are forested (Fig. 4 inset).

Table 4
Logistic model parameters, ranked by *Z* score. See Table 1 for description of predictors.

Predictor	*Z* score	*P*	Odds ratio
Intercept	−4.72	<0.001	
Picemar_prop	5.43	<0.001	11.9
BIO2	4.14	<0.001	1.76
Larilar_prop	3.50	<0.001	29.0
Slope	−3.16	0.0015	0.86
Site_age	1.93	0.053	1.055
Site_height	−1.71	0.087	0.928
BIO4	−0.73	0.462	0.994

Table 5
Accuracy measures for the logistic regression and boosted regression tree (BRT) models of peatland presence based on ground plot data. Validation statistics are derived from comparison of both models comes against the independent ABMI ground plot database (Alberta Biodiversity Monitoring Institute, 2015).

	Regression model	
Metric	Logistic	BRT
Sensitivity (%)	76.7	84.6
Specificity (%)	92.9	87.7
Accuracy (%)	88.2	87.1
AUC[a]	0.931	0.934

[a] Area under the receiver operation characteristic curve.

The accuracy of the spatial model varied widely by ecoprovince, with the correct prediction of peatland location as high as 95.3% in the 64 points in the Mid-Boreal Shield of Ontario, and 84.2% in the Lake of the Woods ecoprovince. Although considered part of the Boreal Shield ecozone by Natural Resources Canada (2015b), the Lake of the Woods ecoprovince shares deep glacial sediments and expansive peatland complexes with the Boreal Plains ecozone (Table 6; Heinselman, 1963). Further south, spatial predictions of peatland location were only 48.0% accurate, while in southern Ontario, the model was only 5.5% accurate. The 5.5% accuracy in southern Ontario is no different from chance, since less than 5% of the region is covered by peatlands (Tarnocai et al., 2011). Spatial prediction of uplands (non-peatland forests with mineral soils), the corollary of prediction of peatlands, was overall consistently more accurate than prediction of peatlands. The model was 77–90% accurate in the prediction of uplands in the boreal and taiga forest, except in the Boreal Shield ecozone, where prediction of uplands was only 63.8% accurate (Table 6). The total area of treed and forested peatland in the Boreal and Taiga ecozones of Canada estimated by this method using a model output threshold of 0.40 (108.8 Mha) is close to the sum area of treeless and treed peatlands estimated by Tarnocai et al. (2011) of 109.5 Mha.

4. Discussion

4.1. Ground plot data

In the boreal forest of western North America, black spruce abundance is strongly related to organic matter accumulation and poor drainage conditions, as the faster-growing aspen (*Populus tremuloides* Michx.) and pines (*Pinus banksiana* Lamb./*Pinus contorta* Dougl. ex Loud.) dominate in upland stands (Johnstone and Chapin, 2006). The presence of black spruce has been used as an indicator of potential paludification (transformation process to peatlands), *Sphagnum* moss dominance, and peat formation in the boreal forests of Ontario and Quebec (Lafleur et al., 2015). However, black spruce does occur as co-dominant with hardwoods or other conifers; indeed, 18% of NFI ground plots with black spruce are mixed with trees from *Populus* spp. or *Pinus* spp. Larch is found more exclusively in wetlands, owing to its shade intolerance (Duncan, 1954) and flooding tolerance (Girardin et al., 2001). In the maritime boreal forest of Newfoundland in Canada, larch is reported to be more widespread (Hall, 1989). Although larch is also found near the treeline in Canada (Payette and Gagnon, 1979), the use of bioclimatic variables to distinguish southern boreal from treeline areas was useful in separating larch stands of similar structure. The marginal effect of larch biomass in Fig. 3 reflects this, with a sharp increase in the marginal effect from 0 to 0.20 proportion of larch biomass, but no further effect once a site exceeds a 0.20 proportion of larch.

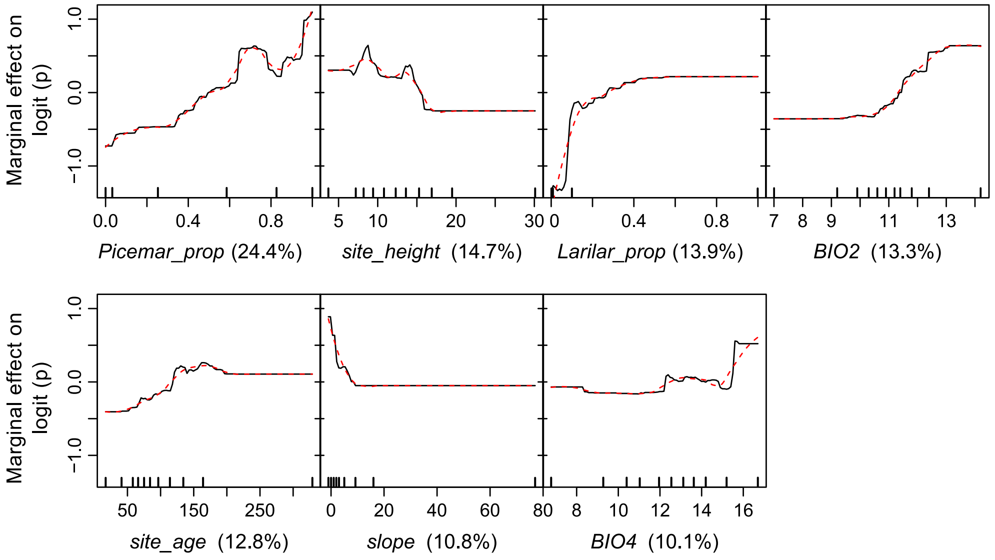

Fig. 3. Response splines of the marginal effect of the top seven predictors in the boosted regression tree model. Percentages correspond to the proportion of the model's power explained by the variable. Dashed lines represent the smooth response function. Vertical tick marks along the *x*-axis of each plot represent deciles of the distribution of each variable.

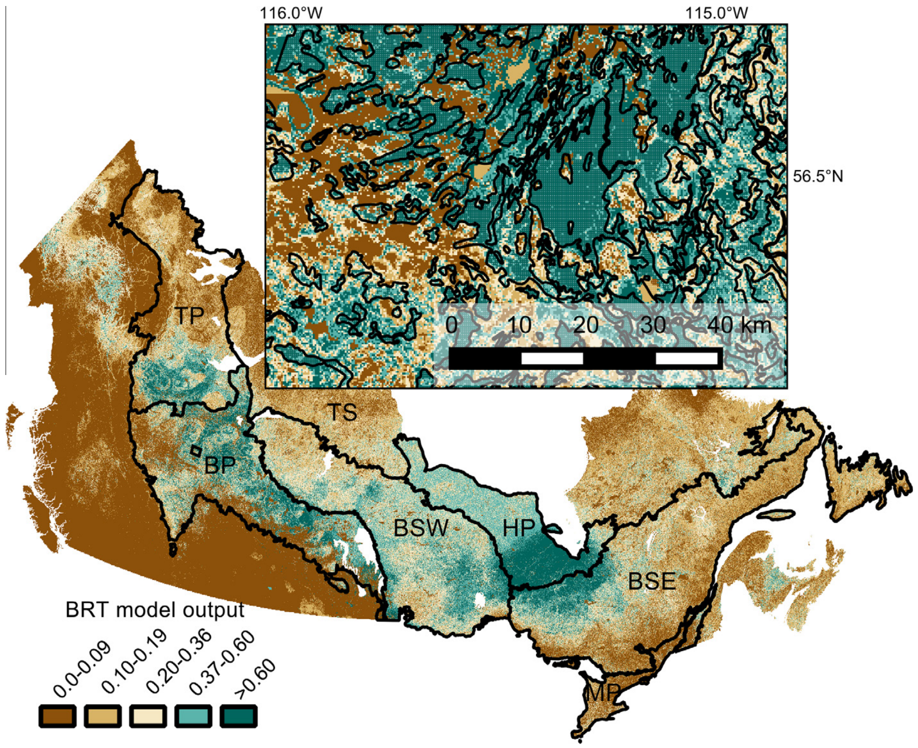

Fig. 4. Spatial prediction of peatland presence using the boosted regression tree approach. The two categories in green represent peatlands (lower threshold = 0.37). Ecozones with available ground plots for model creation and validation (Fig. 1) are outlined in black. Inset map shows the spatial prediction over an area in the Boreal Plains (BP) in north-central Alberta, with an overlay showing the delineation of peatlands from air photos from the Alberta Geological Survey (Fenton et al., 2013). Ecozone acronyms as follows: BSE = Boreal Shield East; BSW = Boreal Shield West; BP = Boreal Plains; HP = Hudson Plains; MP = Mixedwood Plains; TS = Taiga Shield; TP = Taiga Plains.

The *k*-NN model performance in predicting *Larilar_prop*, as reported by Beaudoin et al. (2014), was the lowest among species-composition variables used as predictors, with an R^2 of only 0.07, despite this term being a strong predictor of treed peatland presence at the plot scale. The root-mean-square deviation of *Larilar_prop* was 5%, suggesting this metric is more useful in predicting the presence–absence of *Larix*, and thus the difference between stands with minor *Larix* composition and those dominated by *Larix*. In this study, the BRT splines suggest that it is not the actual proportion of larch biomass that is most informative to the model. However, the sharp increase in the log-odds from 0 to 0.20 proportion larch suggests that any presence, even as a small proportion of the total biomass, provides strong predictive power to the model. Unfortunately, the model statistics from Beaudoin et al. (2014) do not provide a metric of the errors of omission or commission of larch presence, as species composition was

Table 6

Accuracy of the spatial prediction of peatlands and non-peatlands. Accuracy of peatland locations was assessed using an independent validation dataset from Riley and Michaud (1989), while non-peatland accuracy was assessed using horizontal control points from Natural Resources Canada (2015b).

Region	Accuracy (%)	*n*
Peatlands		
Mid-Boreal Shield	95.3	64
Lake of the Woods	84.2	19
Southern Boreal Shield	48.0	25
Great Lakes–St. Lawrence	5.5	18
Total	71.4	126
Non-peatlands		
Taiga Shield (TS)	89.5	19
Taiga Plains (TP)	77.3	97
Hudson Plains (HP)	83.6	67
Boreal Shield (BS)	63.8	105
Boreal Plains (BP)	82.9	318
Total	72.2	606

estimated as a continuous variable, although presence–absence might have been a more appropriate metric in this specific application of the dataset.

Slope was a significant predictor of peatland presence, with 68 out of 103 peatland NFI ground plots reporting zero slope, and another 23 reporting 1–2% slope. For non-permafrost peatlands, topographic indices such as low standard deviations of surface elevation and low DEM ruggedness have been used to distinguish peatlands from uplands (Millard and Richardson, 2013). At the landscape scale, the proportion of areas with slope <2% has been used as a proxy for peatland areas when modelling peatland contributions to riverine dissolved organic C (Olefeldt et al., 2014).

Oceanic boreal regions in Canada feature predominantly nontreed, moss-dominated peatlands (Wells, 1981), owing to the higher annual rainfall and low annual potential evaporation (Price, 1991), two metrics that are not directly captured in the WorldClim database. Greater monthly average temperature range (*BIO4*) is a muted version of the diurnal range, with the two variables being weakly but significantly correlated ($\rho = 0.25$, $P < 0.05$; Table 2), and therefore *BIO4* is a weaker predictor in the model. Like *BIO2*, it adds discrimination power in isolating oceanic versus continental climates, although the contrast at northern latitudes is not as strong.

This study has similarities to species distribution modelling (e.g., Venier et al., 2004; Ohse et al., 2009), in that the goal is to predict the presence or absence of the phenomenon of interest, in this case organic soils greater than a threshold depth, rather than to predict the absolute depth of peat. Efforts to predict the actual peat depth at a location have primarily been conducted in United Kingdom and Irish blanket peat systems (Holden and Connolly, 2011; Parry et al., 2012), where topography plays a greater role. In the case of the NFI ground plot data, many of the measurements of peat depth were truncated to 40 cm (the threshold for organic accumulation in a peatland soil in the Canadian System of Soil Classification), and recorded as such even if the actual depth may have been greater. This artifact of the data limits any prediction of absolute peat depth; more thorough peat depth measurements such as those from Zoltai et al. (2000) unfortunately do not feature the same detail of the forest cover as the NFI ground plots.

4.2. Spatial modelling

The method used here, like digital soil mapping (DSM; e.g., McBratney et al., 2003), aims to create a continuous surface of soil properties, but in this instance the predictive surface layers are not solely a DEM and optical remote sensing, but rather modelled rasters of forest properties. Moreover, most DSM benefits from knowledge of the exact locations of the ground plots used to derive the model. In the case of the NFI ground plots, however, the exact plot locations are not made available in order to protect the integrity of the plots. As a result, for validation data for the spatial model we had to rely instead on known peatland locations that lacked the same forest attribute measurements.

The current implementation of this modelling framework treats each 250 m pixel as an independent point for model prediction and is not influenced by the local abundance of peatlands in surrounding pixels. Many DSM efforts employ Kriging or other forms of spatial autocorrelation for spatial prediction of continuous variables, especially when accurately located point data are available (McBratney et al., 2003). Future efforts to extend this model may incorporate the local abundance of peat in other pixels as a proxy for depth or may factor in the presence of large peatland complexes, which occur more frequently in the Boreal Plains and Taiga Plains ecozones than in the granitic environment of the Canadian Shield in the Boreal and Taiga Shield ecozones (Halsey et al., 1997).

4.3. Spatial modelling by region

High agreement was observed between the estimate of total treed peatland area in Boreal and Taiga regions of Canada calculated here at the regional estimates from Tarnocai et al. (2011). There is likely some compensating error where the method shown here maps densely treed peatlands not captured by other methods, while at the same time not mapping completely treeless peatlands. As the area estimates of Tarnocai et al. (2011) area regional by design (estimating portion of peatland area per region), it is difficult to directly compare the two products.

The accuracy of spatial prediction of peatlands was very high in both the Boreal Shield and the Lake of the Woods regions of Ontario (85–95%), in part reflecting peatland complexes (and their centroids, used as coordinates) on the order of 1000 ha in size or larger in the validation dataset. The high accuracy in the Boreal Plains (uplands predicted with 83% accuracy) contrasts with a higher false-positive rate in the Boreal Shield, where uplands were predicted with only 64% accuracy. This high false-positive rate could be due in part to the smaller size of many of the peatlands in the granitic terrain of the Canadian Shield (many of which are smaller than the 250 m pixel size used in the *k*-NN data from Beaudoin et al. (2014)), where peatlands sit among a complex pattern of bedrock uplands (Branfireun and Roulet, 1998). The low accuracy in the Southern Boreal Shield ecoprovince in Ontario could be due in part to the higher prevalence of non-boreal trees in forested peatlands in the region, as the genera *Betula*, *Acer*, and *Thuja* (Devito and Dillon, 1993) begin to appear in abundance. These genera are not captured by the species preference for black spruce and larch built into the model's input data. Indeed, peatlands in southern Ontario are at the very southern limit of both black spruce and forested peatlands (Bonan and Sirois, 1992).

Some of the forest attribute data from the NFI *k*-NN rasters used here are also available as polygons in provincial forest inventories. Leading species, merchantable volume, and canopy closure are the primary attributes commonly mapped from aerial images at a finer scale than the 250 m rasters used here (OMNR, 2009; AESRD, 2005). However, such forest inventory data are often acquired by forest lease holders in Canada and are considered proprietary. Moreover, only areas under active forest management are typically mapped in detail, which excludes large portions of the northern Boreal Shield and the majority of the Taiga Plains and Taiga Shield ecozones.

The model underperformed in the boreal foothills of western Alberta at the westernmost extent of the Boreal Plains. Peatlands in this region are more often found on a higher slope than are the majority of boreal peatlands, as groundwater-fed fens are more dominant than ombrotrophic bog systems in these regions (Vitt et al., 1995). In addition to the lower density of ground points, lower model performance in the Taiga Plains ecozone could be due in part to the presence of permafrost. Such peatland plateaus are typically dominated by black spruce, with little to no larch present (Zoltai and Tarnocai, 1971; Zoltai, 1972). Overall, the input data of the NFI ground plots used for model construction is significantly skewed towards more southerly, productive forests, so the accurate prediction of peatland in permafrost areas is unlikely with the current dataset. Furthermore, NFI *k*-NN rasters are likely less accurate outside the spatial distribution of NFI photo plots used to build the dataset; thus, predictions of peatland presence in northern portions of Taiga Shield and Hudson Plains, for example, should be viewed with caution.

5. Conclusions

Forest structure data successfully predicted the extent of treed peatlands across Canada using forest inventory ground plots, with the proportion of aboveground biomass consisting of black spruce and larch being the best vegetation predictors and slope as a key geomorphic predictor. Bioclimatic variables describing annual and daily temperature range were useful in discriminating treed peatlands in the boreal ecozones from areas near the northern treeline in Canada that share similar forest structure. For ecozones where peatlands are common, NFI *k*-NN raster maps of forest structure data derived from NFI photo plot inventory data were used successfully to propagate the plot-level model at the national level. Resulting unique new spatial predictions of peatland location were very accurate in the Boreal Plains, although less accurate in the Boreal Shield, probably because the pattern of small peatland and upland areas in much of the Canadian Shield is difficult to capture at 250 m resolution. The resulting map of the presence of organic soils to a depth of 40 cm or more can be used as an input to spatially explicit forest C modelling, and would be particularly useful in modelling disturbances such as wildfire or peatland drainage through a better knowledge of forested peatland extent.

Acknowledgements

Arlene Hilger provided archival peatland locations. Cindy Shaw provided helpful discussions on forest inventory and soils data. Caitlin Mader provided the inspiration to use species distribution modelling methods. Graham Stinson assisted in the interpretation of the NFI data.

References

Alberta Biodiversity Monitoring Institute (ABMI), 2015 <http://www.abmi.ca/home/data/welcome-data-portal.html> (accessed Nov. 20, 2015).

Alberta Environment and Sustainable Resource Development (AESRD), 2005. Alberta Vegetation Inventory Interpretation Standards 2005 <http://esrd.alberta.ca/lands-forests/documents/AVI-ABVegetation3-InventoryStan-Mar05.pdf> (accessed Nov. 20, 2015).

Anderson, K., Simpson, B., Hall, R.J., Englefield, P., Gartrell, M., Metsaranta, J.M., 2015. Integrating forest fuels and land cover data for improved estimation of fuel consumption and carbon emissions from boreal fires. Int. J Wildland Fire 24, 665–679. http://dx.doi.org/10.1071/WF14142.

Beaudoin, A., Bernier, P.Y., Guindon, L., Villemaire, P., Guo, X.J., Stinson, G., Bergeron, T., Magnussen, S., Hall, R.J., 2014. Mapping attributes of Canada's forests at moderate resolution through k NN and MODIS imagery. Can. J. For. Res. 44, 521–532. http://dx.doi.org/10.1139/cjfr-2013-0401.

Bhatti, J.S., Errington, R.C., Bauer, I.E., Hurdle, P.A., 2006. Carbon stock trends along forested peatland margins in central Saskatchewan. Can. J. Soil Sci. 86 (Special Issue), 321–333. http://dx.doi.org/10.4141/S05-085.

Bonan, G.B., Sirois, L., 1992. Air temperature, tree growth, and the northern and southern range limits to *Picea mariana*. J. Veg. Sci. 3 (4), 495–506.

Branfireun, B.A., Roulet, N.T., 1998. The baseflow and storm flow hydrology of a precambrian shield headwater peatland. Hydrol. Process. 12 (1), 57–72.

Canada Soil Survey Committee, Subcommittee on Soil Classification, 1978. The Canadian System of Soil Classification. Can Dept. Agric. Publ. 1646. Supply and Services Canada, Ottawa.

Devito, K.J., Dillon, P.J., 1993. The influence of hydrologic conditions and peat oxia on the phosphorus and nitrogen dynamics of a conifer swamp. Water Resour. Res. 29, 2675–2685.

Duncan, D.P., 1954. A study of some of the factors affecting the natural regeneration of tamarack (Larix laricina) in Minnesota. Ecology 35, 498–521.

Ecological Stratification Working Group, 1996. A National Ecological Framework for Canada. Agriculture and Agri-food Canada, Ottawa, Ontario.

Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees. J. Anim. Ecol. 77, 802–813.

Fenton, M.M., Waters, E.J., Pawley, S.M., Atkinson, N., Utting, D.J., Mckay, K., 2013. Surficial Geology of Alberta. Map 601. Alberta Geological Survey, Edmonton, Alberta <http://www.ags.gov.ab.ca/publications/abstracts/Map_601.html>.

Frolking, S.E., Bubier, J.L., Moore, T.R., Ball, T., Bellisario, L.M., Bhardwaj, A., Carroll, P., Crill, P.M., Lafleur, P.M., McCaughey, J.H., Roulet, N.T., Suyker, A.E., Verma, S.B., Waddington, J.M., Whiting, G.J., 1998. Relationship between ecosystem productivity and photosynthetically active radiation for northern peatlands. Glob. Biogeochem. Cycles 12, 115–126.

Gillis, M.D., Omule, A.Y., Brierley, T., 2005. Monitoring Canada's forests: The National Forest Inventory. For. Chron. 81, 214–221. http://dx.doi.org/10.5558/tfc81214-2.

Girardin, M.P., Tardif, J., Bergeron, Y., 2001. Gradient analysis of *Larix laricina* dominated wetlands in Canada's southeastern boreal forest. Can. J. Bot. 79, 444–456. http://dx.doi.org/10.1139/b01-023.

Gorham, E., 1991. Northern peatlands: role in the carbon cycle and probable responses to climatic warming. Ecol. Appl. 1, 182–195.

Haavisto, V.F., Jeglum, J.K., Wearn, V.H., 1995. Operational Drainage: Planning, Costs and Impact Mitigation. Natural Resources Canada, Canadian Forest Service, Great Lakes Forestry Centre, Sault Ste. Marie, Ontario, Frontline Technical Note 63. URL: <http://www.cfs.nrcan.gc.ca/publications?id=9208>.

Hall, J.P., 1989. Growth and development of larch in Newfoundland. In: Murray, M. (Ed.), Proceedings of the Sixth International Workshop on Forest Regeneration: The Yield Advantages of Artificial Regeneration at High Latitudes, pp. 21–26.

Halsey, L., Vitt, D., Zoltai, S., 1997. Climatic and physiographic controls on wetland type and distribution in Manitoba, Canada. Wetlands 17, 243–262.

Heinselman, M.L., 1963. Forest sites, bog processes, and peatland types in the Glacial Lake Agassiz Region, Minnesota. Ecol. Monogr. 33, 327–374.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25, 1965–1978. http://dx.doi.org/10.1002/joc.1276.

Hijmans, R.J., Philips, S., Leathwick, J., Elith, J., 2015. Package "dismo" <http://cran.r-project.org/web/packages/dismo/dismo.pdf> (accessed Nov. 20, 2015).

Hogg, E.H., Lieffers, V.J., Wein, R.W., 1992. Potential carbon losses from peat profiles: effects of temperature, drought cycles, and fire. Ecol. Appl. 2, 298–306.

Holden, N.M., Connolly, J., 2011. Estimating the carbon stock of a blanket peat region using a peat depth inference model. Catena 86, 75–85. http://dx.doi.org/10.1016/j.catena.2011.02.002.

Johnstone, J.F., Chapin III, F.S., 2006. Fire interval effects on successional trajectory in boreal forests of northwest Canada. Ecosystems 9, 268–277. http://dx.doi.org/10.1007/s10021-005-0061-2.

Kettridge, N., Thompson, D.K., Bombonato, L., Turetsky, M.R., Benscoter, B.W., Waddington, J.M., 2013. The ecohydrology of forested peatlands: Simulating the effects of tree shading on moss evaporation and species composition. J. Geophys. Res.: Biogeosci. 118 (2), 422–435.

Lacelle, B., 1998. Canada's soil organic carbon database. In: Lal, R., Kimble, J.M., Follett, R.F., Stewart, B.A. (Eds.), Soil Processes and the Carbon Cycle. CRC Press, Boca Raton, Florida, pp. 93–101.

Lafleur, B., Fenton, N.J., Bergeron, Y., 2015. Forecasting the development of boreal paludified forests in response to climate change: a case study using Ontario ecosite classification. For. Ecosyst. 2, 1–11. http://dx.doi.org/10.1186/s40663-015-0027-6.

Lambert, M.C., Ung, C.H., Raulier, F., 2005. Canadian national tree aboveground biomass equations. Can. J. For. Res. 35, 1996–2018. http://dx.doi.org/10.1139/x05-112.

Li, J., Chen, W., 2005. A rule-based method for mapping Canada's wetlands using optical, radar and DEM data. Int. J. Remote Sens. 26, 5051–5069. http://dx.doi.org/10.1080/01431160500166516.

Lorey, T., 1878. Die mittlere Bestandeshohe. Allg. Forst Jagdztg. 54, 149–155.

Lukenbach, M.C., Hokanson, K.J., Moore, P.A., Devito, K.J., Kettridge, N., Thompson, D.K., Wotton, B.M., Petrone, R.M., Waddington, J.M., 2015. Hydrological controls on deep burning in a northern forested peatland. Hydrol. Process. http://dx.doi.org/10.1002/hyp.10440.

McBratney, A.B., Santos, M.M., Minasny, B., 2003. On digital soil mapping. Geoderma 117, 3–52. http://dx.doi.org/10.1016/S0016-7061(03)00223-4.

Millard, K., Richardson, M., 2013. Wetland mapping with LiDAR derivatives, SAR polarimetric decompositions, and LiDAR–SAR fusion using a random forest classifier. Can. J. Remote Sens. 39, 290–307. http://dx.doi.org/10.5589/m13-038.

Miller, C.A., Benscoter, B.W., Turetsky, M.R., 2015. The effect of long-term drying associated with experimental drainage and road construction on vegetation composition and productivity in boreal fens. Wetl. Ecol. Manag. 5, 845–854.

Moore, T.R., Dalva, M., 1993. The influence of temperature and water table position on carbon dioxide and methane emissions from laboratory columns of peatland soils. J. Soil Sci. 44, 651–664.

Nakai, T., Sumida, A., Kodama, Y., Hara, T., Ohta, T., 2010. A comparison between various definitions of forest stand height and aerodynamic canopy height. Agric. For. Meteorol. 150, 1225–1233. http://dx.doi.org/10.1016/j.agrformet.2010.05.005.

Natural Resources Canada, 2015a. Canadian Spatial Reference System <http://webapp.geod.nrcan.gc.ca/geod/csrs/csrs.php> (accessed May 25, 2015).

Natural Resources Canada, 2015b. Ecological Framework <http://geogratis.gc.ca/api/en/nrcan-rncan/ess-sst/d9df394f-8893-11e0-bc90-6cf049291510.xml> (accessed May 25, 2015).

Ohse, B., Huettmann, F., Ickert-Bond, S.M., Juday, G.P., 2009. Modeling the distribution of white spruce (*Picea glauca*) for Alaska with high accuracy: an open access role-model for predicting tree species in last remaining wilderness areas. Polar Biol. 32, 1717–1729. http://dx.doi.org/10.1007/s00300-009-0671-9.

Olefeldt, D., Persson, A., Turetsky, M.R., 2014. Influence of the permafrost boundary on dissolved organic matter characteristics in rivers within the Boreal and Taiga plains of western Canada. Environ. Res. Lett. 9, 035005. http://dx.doi.org/10.1088/1748-9326/9/3/035005.

Ontario Ministry of Natural Resources (OMNR), 2009. Ontario Forest Information Manual 2009. Queen's Printer for Ontario, Toronto, Ontario.

Palylyk, C.L., Crown, P.H., 1984. Application of clustering to Landsat MSS digital data for Peatland inventory. Can. J. Remote Sens. 10, 201–208. http://dx.doi.org/10.1080/07038992.1984.10855070.

Parry, L.E., Charman, D.J., Noades, J.P.W., 2012. A method for modelling peat depth in blanket peatlands. Soil Use Manag. 28, 614–624. http://dx.doi.org/10.1111/j.1475-2743.2012.00447.x.

Payette, S., Gagnon, R., 1979. Tree-line dynamics in Ungava peninsula, northern Quebec. Ecography 2, 239–248.

Pratt, L.E., 1965. The Canada Land Inventory: Objectives, Scope, and Organization Publication number 1088. Department of Forestry, Ottawa.

Price, J.S., 1991. Evaporation from a blanket bog in a foggy coastal environment. Bound. Layer Meteorol. 57, 391–406.

Riley, J.L., Michaud, L., 1989. Peat and Peatland Resources of Northwestern Ontario, vol. 144. Ontario Ministry of Northern Development and Mines, Mines and Minerals Division, Toronto, Ontario.

Robinson, N., Regetz, J., Guralnick, R.P., 2014. EarthEnv-DEM90: a nearly-global, void-free, multi-scale smoothed, 90m digital elevation model from fused ASTER and SRTM data. J. Photogramm. Remote Sens. 87, 57–67. http://dx.doi.org/10.1016/j.isprsjprs.2013.11.002.

Tarnocai, C., Kettles, I., Ballard, M., 1995. Peatlands of Canada. Geological Survey of Canada Open File 3152. Geological Survey of Canada, Ottawa, Ontario. http://dx.doi.org/10.4095/205270.

Tarnocai, C., Kettles, I., Lacelle, B., 2011. Peatlands of Canada. Geological Survey of Canada Open File 6561. Geological Survey of Canada, Ottawa, Ontario. http://dx.doi.org/10.4095/288786.

Touzi, R., Deschamps, A., Rother, G., 2007. Wetland characterization using polarimetric RADARSAT-2 capability. Can. J. Remote Sens. 33, S56–S67. http://dx.doi.org/10.5589/m07-047.

United Nations Framework Convention on Climate Change, 2002. Report of the Conference of the Parties on its Seventh Session. FCCC/CP/2001/13/Add.1. Marrakesh, Morocco. <http://unfccc.int/resource/docs/cop7/13a01.pdf> (accessed Nov. 20, 2015).

Venier, L.A., Pearce, J., McKee, J.E., McKenney, D.W., Niemi, G.J., 2004. Climate and satellite-derived land cover for predicting breeding bird distribution in the Great Lakes Basin. J. Biogeogr. 31, 315–331. http://dx.doi.org/10.1046/j.0305-0270.2003.01014.x.

Vitt, D.H., Halsey, L.A., Thormann, M.N., Martin, T., 1995. Peatland Inventory of Alberta. Peatland Task Force, Alberta Environmental Protection Agency, Edmonton, Alberta.

Wells, E.D., 1981. Peatlands of eastern Newfoundland: distribution, morphology, vegetation, and nutrient status. Can. J. Bot. 59, 1978–1997.

Wieder, R.K., Scott, K.D., Kamminga, K., Vile, M.A., Vitt, D.H., Bone, T., Xu, B., Benscoter, B.W., Bhatti, J.S., 2009. Postfire carbon balance in boreal bogs of Alberta, Canada. Glob. Change Biol. 15, 63–81.

Wulder, M.A., White, J.C., Cranny, M., Hall, R.J., Luther, J.E., Beaudoin, A., Goodenough, D.G., Dechka, J.A., 2008. Monitoring Canada's forests. Part 1: completion of the EOSD land cover project. Can. J. Remote Sens. 34, 549–562.

Zoltai, S.C., 1972. Palsas and peat plateaus in central Manitoba and Saskatchewan. Can. J. For. Res. 2, 291–302.

Zoltai, S.C., Martikainen, P.J., 1996. Estimated extent of forested peatlands and their role in the global carbon cycle. In: Apps, M.J., Price, D.T. (Eds.), Forest Ecosystems, Forest Management, and the Global Carbon Cycle, NATO ASI Series, vol. I 40. Springer-Verlag, Berlin, pp. 47–57.

Zoltai, S.C., Pollet, F.C., Jeglum, J.K., Adams, G.D., 1975. Developing a wetland classification system for Canada. In: Bernier, B., Winget, C.H. (Eds.), Forest Soils and Forest Land Management: Proceedings of the Fourth North American Forest Soils Conference. Les Presses de l'Université Laval, Quebec City, Quebec, pp. 497–511 http://www.cof.orst.edu/cof/fr/people/nafsc4.htm (accessed Nov. 20, 2015).

Zoltai, S.C., Tarnocai, C., 1971. Properties of a wooded palsa in northern Manitoba. Arct. Alp. Res. 3, 115–129.

Zoltai, S.C., Vitt, D.H., 1995. Canadian wetlands: environmental gradients and classification. In: Finlayson, C.M., van der Valk, A. (Eds.), Classification and Inventory of the World's Wetlands. Springer Netherlands, Dordrecht, Netherlands.

Zoltai, S.C., Siltanen, R.M., Johnson, J.D., 2000. A wetland data base for the western boreal, subarctic, and arctic regions of Canada. Can., Can. For. Serv., North. For. Cent., Edmonton, Alberta. Inf. Rep. NOR-X-368.