INTERNATIONAL JOURNAL OF REMOTE SENSING
2020, VOL. 41, NO. 10, 4022–4056
https://doi.org/10.1080/01431161.2019.1711245

Taylor & Francis
Taylor & Francis Group

Reconstructing past forest composition and abundance by using archived Landsat and national forest inventory data

Bina Thapa[a], Peter T. Wolter[a], Brian R. Sturtevant[b] and Philip A. Townsend[c]

[a]Department of Natural Resource Ecology and Management, Iowa State University, Ames, IA, USA; [b]Institute for Applied Ecosystem Studies, Northern Research Station, USDA Forest Service, Rhinelander, WI, USA; [c]Department of Forest and Wildlife Ecology, University of Wisconsin—Madison, Madison, WI, USA

ABSTRACT

Effective modelling of forest susceptibility to defoliating insect outbreaks requires a better understanding of outbreak dynamics, which includes detailed knowledge of the pre- and post-outbreak forest status as well as subsequent feedback mechanisms. In this paper, we strive to fill the forest status need by combining archived Landsat sensor data (pre- and post-outbreak) with different formats and dates of the U.S. Forest Service's Forest Inventory and Analysis (FIA) data (periodic [1970s, 1990s] and annual [2003–2006]). Specifically, we explore the utility of these FIA ground data for calibrating models of forest species and type abundance for mapping past forest composition in the Border Lakes Ecoregion (BLE) of Upper Midwest of the US. Model calibration results between Landsat reflectance and FIA ground data for both total forest basal area and balsam fir (*Abies balsamea*) relative basal area, a preferred host of the spruce budworm (SBW, *Choristoneura fumiferana*), were poor to moderate (R^2_{adj} 0.39 and 0.48, respectively). Results for aspen (*Populous tremuloides*) and spruce (*Picea glauca* and *P. mariana*) abundance yielded substantially better accuracies (R^2_{adj} 0.64 and 0.78; RMSE 15.56 and 10.65 m^2 ha^{-1}, respectively). Groupings of tree species into broadleaved and conifers substantially improved model calibration result (R^2_{adj} range: 0.72–0.91), except for the SBW host group (*A. balsamea, P. glauca*, and *P. mariana*). Periodic FIA ground data from the early 1990s generated stronger models compared to other FIA-Landsat date combinations tested. A paired t-test of abundance differences between undisturbed forest from the older 1977 and 1990 periodic inventories was significant (p-value < 0.0001), suggesting possible effects of variable FIA sampling protocol or ground plot positional accuracy through time. However, a similar paired t-test of abundance difference between periodic FIA (1990) and annual FIA (2003–2006) was not significant (p-value = 0.249). We posit four potential factors that may have contributed to weak Landsat-FIA calibration results for species abundance: 1) variation in FIA subplot arrangement and sampling protocols through time, 2) variability in species abundance and heterogeneity among FIA sampling across adjacent Landsat orbital paths, 3) understory species (balsam fir) that are largely hidden from remote detection, and 4) cloud cover and orbital phase mismatches preventing capture of key forest phenology aids. While past and present FIA sampling protocols were not specifically

ARTICLE HISTORY
Received 10 July 2019
Accepted 14 November 2019

CONTACT Bina Thapa btha001@iastate.edu

designed for integration with 30-meter satellite sensor data, careful pairing of FIA ground data (past or present) with Landsat sensor data can facilitate reasonable estimates, of forest abundance for generalized forest types, and possibly forest species when heterogeneity is low. Nevertheless, we recommend that FIA subplot sampling protocols be augmented to include measurements of forest conditions that are more amenable to integration with 30-meter Landsat sensor data.

1. Introduction

Outbreaks of forest insects have had profound impacts on the dynamics of northern forest ecosystems (Franklin et al. 2002; Dymond et al. 2010; Sommerfeld et al. 2018; Sturtevant et al. 2015) and associated socio-economic processes (Dale et al. 2001; Turner 2010). Economic losses to insect and pathogen disturbance approach 1.5 billion dollars annually in the U.S. alone (Dale et al. 2001). More alarming is the fact that stand-killing insect outbreaks have the potential to advance the accumulation of CO_2 in Earth's atmosphere by converting forests from net carbon sinks to sources across relatively short time frames (Dymond et al. 2010).

A unique aspect of insect outbreaks as a forest disturbance agent is that damage is generally limited to a restricted taxonomy defined by the range of susceptible hosts for a given herbivore (Sturtevant et al. 2004; Koricheva, Klapwijk, and Björkman 2012), in many cases restricted to a particular genus or family (Jactel and Brockerhoff 2007). Broad-scaled risk assessments quantifying forest susceptibility to specific herbivores therefore requires detailed compositional information (e.g., Crocker et al. 2016). Evidence suggests that outbreak dynamics (i.e., outbreak duration, frequency, intensity, or spatial synchrony) can be related to the landscape concentration or arrangement of host (Volney and McCullough 1994; Roland 2005; Wesołowski and Patryk 2006; Haynes, Liebhold, and Johnson 2009; Robert et al. 2018) and the relative proportion of associated non-host species (e.g., Cappuccino et al. 1998; Raffa, Powell, and Townsend 2013). Hence, for insect pests that cause significant mortality in their forest host, such as the endemic eastern spruce budworm (SBW, *Choristoneura fumiferana*), accurate mapping of host before and after major outbreaks is required.

Data from the national Forest Inventory and Analyses program (FIA) provide detailed information on past forest composition and structure, changes over time, and incidences of mortality from insects and other agents. Moreover, these FIA-based estimates are design-unbiased, with associated estimates of uncertainty, which most map-based estimates cannot provide (M.D. Nelson, pers. comm.). However, the distribution of FIA sample points across the landscape is far too sparse (i.e., one-sample plot per 2,400-ha hexagon, Bechtold and Patterson 2005) to enable finer-scale analyses of spatial process mechanics, such as those required for the analyses of eastern spruce budworm dynamics in Minnesota, USA and neighbouring Ontario, Canada (James et al. 2011). Reliable methods are needed whereby FIA data may be used as ground truth to enable finer-scale, wall-to-wall mapping of past forest composition and structure to enable more robust (i.e., long term) understanding of the relationships of forest response to insect outbreak dynamics. Such scaling-up of past forest inventory is possible via satellite remote sensing (Wolter et al. 2008), as the continuous

Landsat archive dates back to ca. 1972 and 1984 (57-metre Landsat Multispectral Scanner [MSS] and subsequent 30-metre Landsat series of sensors [TM, ETM+, and OLI]). Towards this end, many examples of the utility of Landsat sensor data for ecosystem monitoring and mapping exist (Franklin 1986; Hall et al. 1991; Moore and Bauer 1990; Tucker 1978; Turner et al. 1999; Vogelmann et al. 2001; Wolter et al. 1995, and many more). Typically, estimates of forest structure are derived via calibration of satellite sensor data with ground data. In some instances, ground sampling protocols are specifically designed to take advantage of the sensor's spatial resolution (Wolter et al. 2008; Wolter, Townsend, and Sturtevant 2009; Wolter and Townsend 2011). Few studies have successfully combined FIA and Landsat data to calibrate models for mapping continuous forest structure (McRoberts et al. 2007; Nelson et al. 2009; Wilson, Knight, and McRoberts 2018). Rather, many efforts focused on developing categorical maps of forest attributes: forest versus non-forest (Franco-Lopez, Ek, and Bauer 2001), age-class (Song, Schroeder, and Cohen 2007; Liu et al. 2008), basal area (Franco-Lopez, Ek, and Bauer 2001), or disturbance (Nelson et al. 2009; Thomas et al. 2011). Other studies have successfully combined FIA and other ancillary information with coarse-scale satellite sensor data such as the Moderate Resolution Imaging Spectroradiometer (MODIS, 250–500 m) sensor for regional mapping (Zhu and Evans 1994; Ruefenacht et al. 2008; Blackard et al. 2008; Nelson et al. 2011; Wilson, Lister, and Riemann 2012; Wilson, Woodall, and Griffith 2013; Goerndt, Wilson, and Aguilar 2019). All such studies highlight the wide potential of combining FIA and remote sensing data to meet forest ecosystem research needs.

The spatial quality and temporal distribution of these FIA ground data are germane to the objectives of this study. Past and current FIA data include two general sets of sampling protocols termed periodic and annual, respectively. These two sampling protocols differ by subplot number and configuration (Figure 1), sampling design, and types of measurement (Goeking 2015; McRoberts et al. 2005). Former FIA surveys (before 1998) were termed periodic as return times to a particular state for re-measurement ranged from 6 to 18 years (McRoberts et al. 2005). The USDA Forest Service recognized issues with the periodic sampling protocol, including bias and uncertainty due to regionally variable sampling protocols and definitions (Goeking 2015), which, when compounded by land-use change over long sample periods, made area and volume estimation increasingly difficult over large regions (McRoberts et al. 2005). To alleviate such problems, the 1998 Farm Bill laid out a plan to initiate an annual protocol of FIA sampling (after 1998) that would standardize a national set of core variables. This included a common set of definitions, sampling protocols, and a simplified, unbiased plot design (Figure 1). Under the new, annual FIA sampling protocol, 20% of all FIA plots in a given region are sampled every year so that all plots are measured once every 5 years (Vandendriesche and Haugen 2008).

Thus, two primary objectives exist for this research conducted across a landscape that encompasses northern Minnesota and adjacent Ontario in the Border Lakes Ecoregion (BLE) (Figure 2). First, produce spatially explicit, continuous forest composition maps (including spruce budworm host: balsam fir [*A. balsamea*] and spruce [*Picea spp.*]) for two periods in the past (ca. 1985 and ca. 2005) at 30-meter grain-size using the combination of archived satellite sensor data (Landsat TM) and archived FIA data. These two periods correspond roughly to pre- and post-outbreak stages of a past spruce budworm disturbance event in this region (Robert, Kneeshaw, and Sturtevant 2012; Robert et al. 2018). Second, assess respective FIA-Landsat calibration performance differences between former (periodic) and current (annual) FIA sampling protocols, as forest composition maps linked to each

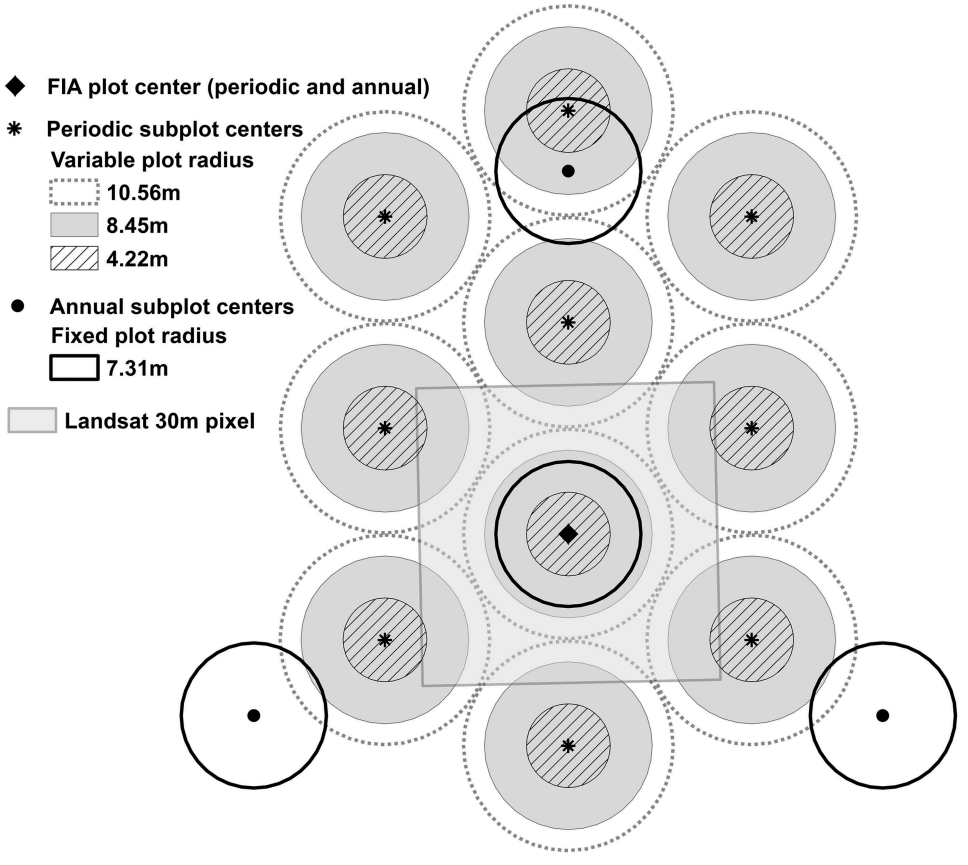

Figure 1. Forest Inventory and Analysis (FIA) sample plot designs for the old, periodic FIA protocol and the new annual FIA protocol adopted in 1998. Periodic FIA data were collected using variable-radius sampling methods. Hence, sub-plots are shown with different radii (i.e., 4.22 m, 8.45 m, and 10.56 m) associated with tree bole diameters of 25.4 cm, 50.8 cm, and 63.5 cm, respectively. Whereas, annual FIA sub-plots in the Minnesota region use a fixed, 7.31 m radius. The superimposed square represents a single 30-meter Landsat pixel. In addition, a black diamond symbol represents periodic and annual FIA full plot center as well as hypothetical Landsat pixel center.

inventory must be comparable across time. Hence, calibration tests using these FIA data include comparisons between periodic and annual FIA sampling design (Figure 1), lag time between FIA sampling and satellite image acquisition dates, and heterogeneity in forest cover proximal to FIA plot centers. To date, we are unware of studies that investigated forest structure modelling where both Landsat sensor data and ground calibration data (i.e., FIA) were collected greater than 13 years before the analysis period, which is the case in this study: 1985 and 2005 (34 and 14 years, respectively).

In this study, forest basal area (BA) and relative basal area (RBA) information derived from the national FIA program were used as metrics of tree abundance. Such forest metrics are standard in forest ecosystem studies for quantifying the abundance of all or individual forest species in a given area that takes into account both stem count and tree size (Reineke 1933; Wolter et al. 2008). In this research, forest structure information at individual species- or genus-level and groupings of species is pertinent to understanding

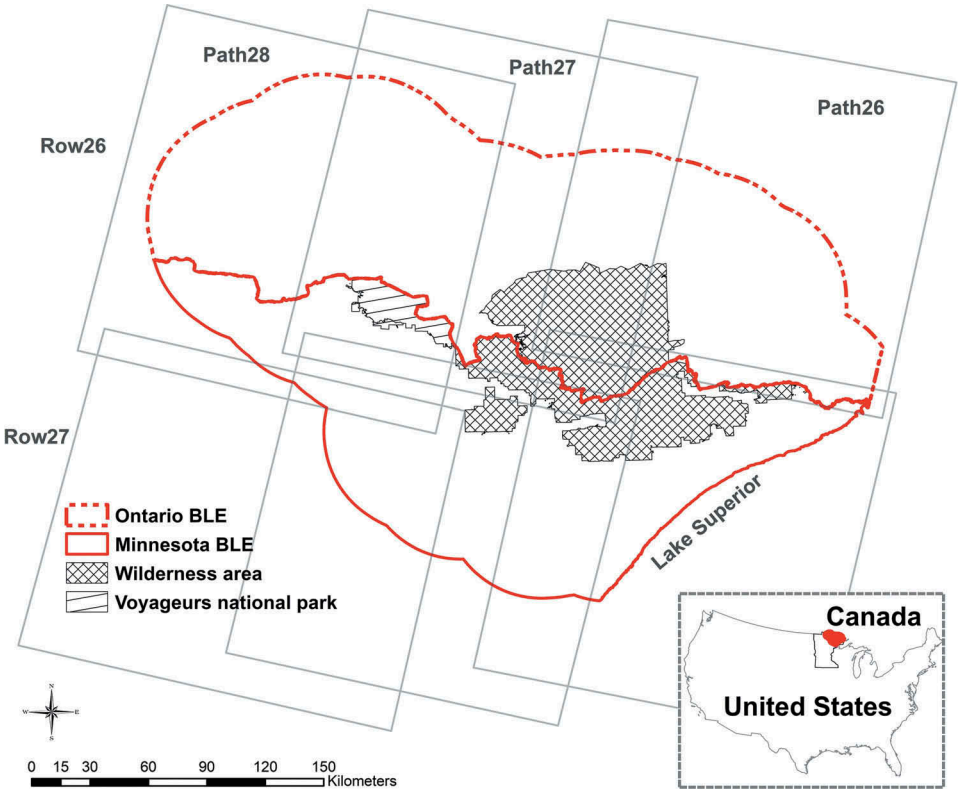

Figure 2. The Boarder Lakes Ecoregion (BLE) study area in northern Minnesota, USA and adjacent Ontario, Canada. The area is a mixture of managed forest (outlined in dotted and solid red lines) and unmanaged natural or wilderness forests (simple hatch and crosshatch in black). Overlapping grey rhombus polygons represent the six Landsat footprints that cover the study area (Landsat WRS-2 paths 26–28 and rows 26–27).

the mechanisms of spruce budworm (*Choristoneura fumiferana*) outbreaks in this region (James et al. 2011; Sturtevant et al. 2015). Given the importance, we explore the relative strengths and weakness of calibrating forest structure models for mapping purposes using the newer annual and older periodic FIA data with Landsat sensor data.

There has been substantial use of nearest-neighbour techniques, such as Most Similar Neighbour (MSN) (Moeur and Stage 1995), Gradient Nearest Neighbour (GNN) (Ohmann and Gregory 2002), and k-Nearest Neighbour (McRoberts, Nelson, and Wendt 2002; McRoberts et al. 2007; Wilson, Knight, and McRoberts 2018), to relate FIA ground data to satellite sensor data for forest parameter estimation and mapping purposes. In particular, k-NN has gained much traction, because it is non-parametric, multivariate, transparent, and a relatively simple method for producing estimates of continuous forest structure variables (McRoberts, Nelson, and Wendt 2002; McRoberts et al. 2007). The past remote sensing research has clearly demonstrated that the accuracy of forest cover type mapping can be substantially improved by careful selection of many additional satellite images that coincide with various species-specific phenology phenomena (Wolter et al. 1995, 2008; Wilson, Lister, and Riemann 2012). While other studies have successfully linked Landsat and FIA data for mapping forest basal area

(McRoberts et al. 2007; Franco-Lopez, Ek, and Bauer 2001; Wilson, Knight, and McRoberts 2018), finer-scale mapping (i.e., 30-meter) of near species-level forest composition and abundance have not yet been reported. Here, we apply the iterative exclusion partial least squares (xPLS) regression approach (Wolter et al. 2008, 2012) designed to handle numerous, collinear input image predictors. This basic approach for modelling and mapping forest structure has been successfully tested and validated in this region with Landsat and other satellite sensor data (Radarsat-1, PALSAR, SPOT-5) (Wolter et al. 2008; Wolter, Townsend, and Sturtevant 2009; Wolter and Townsend 2011), but has not yet been tested where FIA data represent ground truth.

2. Methods

2.1. Study area

The Border Lakes Ecoregion (BLE) encompasses northern Minnesota and neighbouring Ontario and requires six Landsat thematic mapper (TM) images (each 170 km × 183 km with ca. 35% side-lap) to cover this area: three Worldwide Reference System (WRS-2) Paths (P26-P28) and two WRS-2 Rows (R26-R27) (Figure 2). The study area is a transitional sub-boreal forest (Baker 1989; Heinselman 1973). Within this region, balsam fir (*A. balsamea*) and white spruce (*P. glauca)*) are the preferred host tree species of the spruce budworm, with black spruce (*P. mariana*) serving as a less-preferred host. Non-host tree species includes broadleaved species: aspen (*Populus tremuloides, P. grandidentata*), paper birch (*Betula papyrifera*), maples (*Acer rubrum, A. Saccharum*), ash (*Fraxinus spp*.), and other coniferous species: jack pine (*Pinus banksiana*), big-pines (*P. resinosa, P. strobus*), tamarack (*Larix larcinia*), and northern white cedar (*Thuja occidentalis*). The current forest species composition in the BLE reflects the legacy of fire suppression over the last century (Baker 1992; Frelich and Reich 1995; Scheller et al. 2005) and intensive management for pulpwood (Blais 1983; Wolter and White 2002; Pastor, Sharp, and Wolter 2005).

Substantial patch size contrast exists in clear-cut harvest legacies between northern Minnesota and Ontario, which has created a unique landscape dichotomy (Wolter et al. 2012b). Historically, managed forests in the Ontario portion of the BLE were more homogenous compared to managed forests in Minnesota (Wolter and White 2002) due to a combination of past forest management practices and substantial differences in both the magnitude of public land ownership and average cut-block size between the two countries (Sturtevant et al. 2014; Wolter et al. 2012b). While management differences were evident in these landscapes in the past, cut-block sizes in Ontario have decreased in recent years (Sturtevant et al. 2014; Wolter et al. 2012b). Wilderness forests in the BLE, including the Boundary Waters Canoe Area Wilderness (BWCAW), Voyageurs National Park in Minnesota, and the Quetico Provincial Park in Ontario, bisect the two dominant forest management zones in respective countries (Figure 2).

Recent analyses suggest that divergent land management legacies across the study area have influenced outbreak dynamics of SBW (Robert et al. 2018). Moreover, host in mixture with non-host species, especially broadleaved, can decrease the severity of SBW outbreaks on host species (Su, Needham, and MacLean 1996; Cappuccino et al. 1998), which is similar within western conifer forest attacked by native bark beetles (Raffa, Powell, and Townsend 2013). The last major SBW outbreak in the region attained

maximum defoliation in the mid-1990s over a large portion of the study area. Other important native forest defoliators in this region include forest tent caterpillar (*Malacosoma disstria*) affecting hardwood tree species, jack pine budworm (*Choristoneura pinus pinus*), and potentially eastern larch beetle (*Dendroctinus simplex*) (Crocker et al. 2016). Additional invasive insect pests, such as European gypsy moth (*Lymantria dispar dispar*) and emerald ash borer (*Agrilus planipennis*) currently threaten the region.

2.2. Field data

Forest Inventory and Analysis (FIA) field plot data from 1977 and 1990 (periodic) and from 2003, 2004, 2005 and 2006 (annual) for northern Minnesota were obtained from the USDA Forest Service under a memorandum of understanding between the US Forest Service and Iowa State University. We did not obtain forest inventory data from Ontario's Ministry of Natural Resources and Forestry to represent the Canadian side of the study area since a specific goal was to evaluate the use of FIA. Moreover, Canada's national forest inventory differs substantially from FIA in sampling design, subplot configuration, and measurement protocol (McRoberts et al. 2009), which would confound validation efforts. Hence, validation of Landsat-FIA calibrations for the entire Border Lakes Ecoregion (BLE) is based solely on Minnesota FIA data. The FIA field plot locations are permanent positions established across the United States using a stratified random, equal probability sampling technique (Bechtold and Patterson 2005). Periodic FIA plots were designed as a 10-point cluster of variable-radius subplots, but were plagued by spatial, temporal, and methodological inconsistencies and poorly documented sampling strategies that precluded reliable long-term forest change analyses (Goeking 2015). The new, annual FIA protocol has a national standard sampling strategy across a simplified plot design: a central subplot surrounded by three satellite subplots separated by 120 degrees with centers 36.6 m (120 ft) from the center subplot (Figure 1), which represents the full FIA plot center (Bechtold and Patterson 2005). The reader should note that while sample intensity can still vary among states, within states, and between years, the annual FIA sample intensity in Minnesota has been consistent within and between years (M.D. Nelson, pers. comm.).

In this study, we used only forested FIA plots (both periodic and annual) where total forest basal area was greater than 14.5 m^2 ha^{-1}, according to Wolter et al. (2008), and all subplots (10 periodic and 4 annual, Figure 1) had the same forest condition code (i.e., accessible forest land). This procedure served to avoid sources of spatial heterogeneity in ground data related to mixtures of forest and non-forest sub-plot conditions co-occurring within a full FIA plot. Under broad, homogeneous forest conditions, a relatively small FIA sub-plot (i.e., 7.32 m radius) can certainly represent the conditions of a much larger 30-meter pixel. Under heterogeneous forest conditions, individual sub-plot species diversity may be quite different from the overarching heterogeneity associated with the larger pixel area sampled by the satellite sensor (Wilson, Lister, and Riemann 2012). Moreover, positional uncertainty in the ortho-rectified and geo-corrected Landsat products provided by the USGS can be in error by as much as 12 m on the ground (Irons, Dwyer, and Barsi 2012). And in cases where a pixel center is not perfectly aligned with the FIA plot center, a maximum additional positional variance of up to 21.2 m (0.5 × diagonal pixel dimension) is possible. Hence, FIA plot-wise averages of the sub-plot BA data (Figure 1) were calculated and used as response variables to

further reduce the impact of fine-grain spatial variability when compared to the area represented by a 30-meter Landsat pixel, including inherent spatial error due to imperfect image geo-registration. Because of the various spatial mismatches discussed above (pixel-to-ground, FIA-to-ground, and FIA-to-pixel), we did not attempt to parse FIA sub-plots by pixel or groupings of pixels for any Landsat-FIA calibrations. Past attempts to do so have reported mixed effects on accuracy (See McRoberts 2009; Ohmann, Gregory, and Roberts 2014 for more detail). Thus, we believe averaging basal area values across all sub-plots provides a more robust representation of ground conditions within the range of potential space occupied by a single Landsat pixel. These averages included total forest basal area (TOTBA, m^2 ha^{-1}) and associated relative basal areas (RBA, entity BA/TOTBA) of each major forest species, or genus, and species grouping such as spruce budworm (SBW) host species (*A. balsamea* + *P. glauca* + *P. mariana*), all conifers combined (CON), and all broadleaved species combined (BLF) (Table1). Species and species grouping were defined based on insect's host and non-host species as well as to enhance the model accuracy.

We performed additional screening of FIA plots to guard against instances where the forest state may have changed due to disturbance between ground sampling and associated image capture dates. Forest masks for each period (1985 and 2005) were developed (discussed below) and applied to ensure all recent disturbances in that period were excluded from analyses (Table 2). After all screening was complete, 1082 of the periodic FIA plots (inventory year 1990) and 1053 of the annual FIA plots (inventory years 2003–2006) remained for analyses. We withheld 20% of these FIA plots from model development for validation. However, due to the smaller relative size of US land area in WRS-2 P26 (Figure 2), the number of annual FIA plots needed for both calibration and validation was limited (182 plots vs. 428 and 443 plots in WRS-2 P27 and P28, respectively). Therefore, we obtained additional ground data (122 plots) from another Minnesota study (Wolter et al. 2008) solely for P26 model validation purposes.

2.3. Landsat data

Landsat Thematic Mapper (TM) imagery for WRS-2 P26-P28 and rows R26-R27 was downloaded from www.earthexplorer.org as fully processed surface reflectance products that were both ortho-rectified and geo-corrected as part of the Landsat Ecosystem

Table 1. Dependent FIA plot data represent forest species and groups of species for model calibration with Landsat sensor data.

FIA species variable	Type level	Species
Ash	Genus	Black and green ash (*Fraxinus nigra* and *F. pennsylvanica*)
Aspen	Genus	Quaking and big-toothed aspen (*Populus tremuloides* and *P. grandidentata*)
Maple	Genus	Red and sugar maple (*Acer rubrum* and *A. saccharum*)
Birch	Species	Paper birch (*Betual papyrifera*)
Pine	Genus	Red, white and jack pine (*Pinus resinosa, P. strobus* and *P. banksiana*)
Big-pines	Group	Red and white pine (*P. resinosa* and *P. strobus*)
Balsam fir	Species	Balsam fir (*Abies balsamea*)
Spruce	Genus	White and black spruce (*Picea glauca* and *P. mariana*)
Black spruce	Species	Black spruce (*P. mariana*)
Cedar	Species	Northern white cedar (*Thuja occidentalis*)
Tamarack	Species	Tamarack (*Larix laricina*)
Broadleaves (BLF)	Group	Birch, maple, aspen, and ash
Conifers (CON)	Group	Pine, balsam fir, spruce, cedar, and tamarack
Spruce budworm (SBW) host	Group	Balsam fir, white spruce and black spruce

Table 2. Landsat-5 TM image dates, associated bands (B1, blue; B2, green; B3, red; B4, NIR; B5, SWIR; and B7, SWIR), and vegetation indices used for each WRS-2 path and period. Bold image dates indicate the images used for the calculation of neighbourhood spectral variance. The asterisk next to an image date for P27 indicates a Landsat MSS image.

Path	Circa	Image Date	Abbr.	B1	B2	B3	B4	B5	B7	NDVI	MSI	SVR	NBR	AI	NDMI	TC1	TC2	TC3	DI	VVI
P26	1985	24 June 1985	J	×	×	×	×	×	×	×	×	×	×	×	×					
		14 October 1985	O	×	×	×	×	×	×	×	×	×	×	×	×					
		26 May 1986	M				×	×	×	×	×	×	×	×	×					×
		26 September 1990	**S**	×	×	×	×	×	×	×	×	×	×	×	×					
		18 December 1991	D	×	×	×	×			×	×	×	×		×	×	×	×	×	
	2005	14 May 1999	M	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
		04 June 2001	J	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
		24 September 2001	S	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
		26 August 2008	A	×	×	×	×	×	×	×	×	×	×		×	×	×	×	×	×
		03 March 2008	R	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
		03 October 2010	**O**	×	×	×	×	×	×	×	×	×	×		×	×	×	×	×	×
P27	1985	16 September 1984*	S1		×	×	×	×		×	×				×					
		28 April 1985	P	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	
		25 September 1987	S2	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	
		12 May 1990	M	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	
		31 July 1990	**Y**	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	
		11 February 1992	F	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	
	2005	20 May 2004	M	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
		11 October 2004	O				×	×	×		×		×		×					
		18 March 2005	R	×	×	×	×	×	×	×	×	×	×		×	×	×	×	×	
		26 September 2005	S1				×	×	×		×		×		×					
		13 September 2006	**S2**	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
P28	1985	01 October 1984	O		×	×	×	×		×	×				×					
		21 May 1985	M	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	
		05 September 1989	S	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	
		07 August 1990	**A**	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	
		15 February 1991	F	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	
	2005	20 May 2002	M	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
		08 February 2006	F	×	×	×	×			×										
		04 September 2006	S1	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
		22 August 2007	**A**	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
		23 September 2007	S2	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×

Note: Abbreviations indicate specific month, however, month dates vary by WRS-2 path. When two images from same month are used in a particular WRS-2 path and year, we follow the abbreviation with a number (e.g., S1 and S2 represent images from September for 1984 and 1987 for WRS-2 P27, ca. 2005). Landsat predictor variables are then referenced by their month abbreviation followed by and underscore with the specific band or index abbreviation. For example, S_1 indicates Landsat B1 from September; Y_NDVI indicates NDVI index from February.

Disturbance Adaptive Processing System (LEDAPS) (USGS 2015, detail in Vermote et al. 1997; Masek et al. 2006). We carefully evaluated our selection of Landsat images based on five conditions:

- temporal proximity to FIA field data;
- potential capture of species-specific forest phenology phenomenon (Wolter et al. 1995, 2008);
- avoid haze or cloud contamination;
- avoid images concurrent with the ca. 1995 SBW outbreak in this region (Robert, Kneeshaw, and Sturtevant 2012); and
- imagery for a given WRS-2 row in a given WRS-2 path must have an acceptable date complement for either the preceding or successive row in that path (Table 2).

Hence, ca. 1985 (hereafter simply 1985) pre-outbreak forest composition maps were based on the calibration of 1977 and 1990 FIA field data with Landsat imagery from 1984 (Landsat-5 launched March 1984) to 1992 window. Similarly, post-outbreak maps were based on calibration of 2003–2006 FIA field data with Landsat imagery from a 2000–2008 window. However, we allowed images from 14 May 1999 (aspen, leaf-flush) and 3 October 2010 (maple and aspen senescent foliar colouration) as exceptions for 2005 WRS-2 P26 to capture unique and distinct forest phenology (Wolter et al. 1995). As indicated above, 1977 FIA data were collected 8 years before Landsat-5 TM was launched. Though earlier Landsat Multi-spectral Scanner (MSS) data from ca. 1977 exist, expanded use of these more poorly resolved Landsat-MSS sensor data (Moore and Bauer 1990) from 1972–1983 would have certainly confounded the goals of this study. With that said, we allowed the use of a specific Landsat-4 (MSS) image from 16 September 1984 in P27 to capture the narrow timing of maple foliar senescence (Wolter et al. 1995). Ultimately, for both mapping timeframes (1985 and 2005) wider date ranges were needed to procure suitable TM imagery, as instances of excessive cloud cover and a few instances of orbital cycle-to-phenology mismatches prevented tighter temporal constraints on our image search, which is typical in such studies (Vogelmann et al. 2001; Wolter, Johnston, and Niemi 2006). Given the five constraints listed above, a total of 36 Landsat images (Table 2) were used in this study to map SBW host species abundance and distribution from FIA observations.

As briefly discussed in the FIA weeding protocols above, forest masks were developed for each WRS-2 path using the most recent Landsat image in each mapping timeframe (1985 and 2005). In doing so, changes in forest state (e.g., due to harvest or natural disturbance) that occurred between FIA ground sampling date and the most recent, respective, Landsat image dates could be excluded from model development. In general, winter conditions (i.e., with snow ground cover) were considered ideal to produce forest/non-forest masks. Under such conditions, white snow patches indicate non-forest and stand-replacing forest disturbances, which are clearly visible and easily distinguished from both intact conifer (Wolter et al. 2008) and deciduous forests (Wolter et al. 2012). For this study, a simple threshold applied to the green band (band 2, 520–600 nm) of the respective winter images discriminated forest from non-forested areas. We visually assessed this non-forest masking process using the 'Historical Imagery' functionality available within the Google Earth Pro program (Earth version 7.3.2).

2.4. FIA-Landsat model calibrations

2.4.1. Predictor variable selection

Landsat image selections for each of the three respective WRS-2 paths (P26-P28) resulted in a set of images that were date-matched by row within each WRS-2 path, but not matched between paths (Table 2). Hence, Landsat-FIA calibration models for the two separate periods (1985 and 2005) were path-specific, resulting in six sets of models. In addition, we developed five sets of models to test various FIA data stratifications. Initial Landsat predictor variables for a given WRS-2 path consisted of the individual sensor bands (three visible, one near-infrared, and two shortwave infrared bands) and commonly derived indices and ratios (Tables 2 and 3) used in past forest structure mapping studies (Wolter et al. 2008; Cohen and Spies 1992; Hansen et al. 2001, and many more).

Calibrations between respective image predictor variables and dependent FIA ground data were performed using iterative exclusion partial least squares (xPLS) regression (Wolter et al. 2012), which is an iterative variant of standard partial least squares (PLS) regression (Geladi and Kowalski 1986). In general, the use of PLS regression is appropriate in cases where independent or dependent regression variables are highly collinear (Geladi and Kowalski 1986), which is the case with Landsat sensor data stacks (Wolter et al. 2008). The xPLS routine is unique in that it iteratively weeds out predictor variables that show little or no response to dependent variables based on intermediate model performance evaluation (minimum root mean squared error) using a leave-one-out cross-validation (LOOCV) approach (Gong 1986). The xPLS routine results in a parsimonious set of predictor variables that minimizes dependent variable estimation error (Wolter et al. 2012). Final predictor variable sets are then used in a similar PLS routine (nonlinear iterative PLS [NIPALS], Geladi and Kowalski 1986) with 1000 random permutations to generate coefficients for mapping forest species for each path and period. Resulting models include estimates of total BA (TOTBA) and relative BA (RBA) for all dominant forest species (or genera) and important groupings of species (e.g., SBW host, all conifers [CON], and all broadleaved species [BLF]).

Table 3. List of vegetation indices (VI) and associated formulations used in this study.

Index and Abbreviation	Equation	Reference
Normalized Difference Vegetation Index (NDVI)	(TM4 – TM3)/(TM4 + TM3)	Rouse et al. 1974
Moisture Stress Index (MSI)	TM5/TM4	Rock et al. 1986
Shortwave to Visible Ratio (SVR)	(TM5 + TM7)/(TM1 + TM2 + TM3)	Wolter et al. 2008
Normalized Burn Ratio (NBR)	(TM4 -TM7)/(TM4 + TM7)	Key and Benson 1999
Autumn Index (AI)	TM3/TM1	Wolter and Townsend 2011
Normalized Difference Moisture Index (NDMI)	(TM4 -TM5)/(TM4 + TM5)	Hardisky et al. 1983
Tasselled Cap Brightness (TC1)	(0.2043 × TM1) + (0.4158 × TM2) + (0.5524 × TM3) + (0.5741 × TM4) + (0.3124 × TM5) + (0.2303 × TM7)	Healey et al. 2005
Tasselled Cap Greeness (TC2)	(−0.1603 × TM1) + (−0.2819 × TM2) + (−0.4934 × TM3) + (0.7940 × TM4) + (−0.0002 × TM5) + (−0.1446 × TM7)	Healey et al. 2005
Tasselled Cap Wetness (TC3)	(0.0315 × TM1) + (0.2021 × TM2) + (0.3102 × TM3) + (0.1594 × TM4) + (−0.6806 × TM5) + (−0.6109 × TM7)	Healey et al. 2005
Disturbance Index (DI)	TC1 – (TC2 + TC3)	Healey et al. 2005
Visible Vegetation Index (VVI)	TM2/TM3	

Salient predictor variables from the xPLS routine are reported following Wolter et al. (2008, 2009), as this presentation enables assessment of different spectral features for their sensitivity to dependent variables of interest and, thus, the realism of the prediction models with respect to known spectral characteristics of vegetation types. Image predictor variables have a range in magnitude from 0 to 1 with either positive or negative sign (i.e. −1 to +1). Here, we used '+' or '-' followed by the image predictor variable to denote the sign of magnitude (Table 2). To conserve space, we report only important variables from WRS-2 P27 with a magnitude greater or equal to 0.8 (Wold 1994).

2.4.2. Field data stratification tests

Rather than using all FIA data, we investigated Landsat-FIA model calibration sensitivity to (1) the spectral variability in shortwave infrared reflectance (band 5 SWIR, 1550–1750 nm) within a 3 × 3 pixel neighbourhood (kernel) around each FIA plot centre on the ground, and (2) differences in temporal lag between image acquisition and FIA inventory dates. For the first, we assessed the coefficient of variation (CV) in SWIR reflectance within a 3 × 3 pixel neighbourhood around each FIA plot center. Landsat TM's SWIR band 5 is especially sensitive to forest composition differences, such as conifer/broadleaved and forest/non-forest (Hopkins, Maclean, and Lillesand 1988). If the CV of SWIR reflectance around an FIA plot was less than or equal to the neighbourhood mean plus one standard deviation (SD), then the plot was classified as spectrally homogeneous (HOM), otherwise plot was classified as spectrally heterogeneous (HET).

Similarly, temporal lag effect was assessed by using the separate dates of periodic FIA data (1977 and 1990) with the same suite of 1985 Landsat predictor variables to calibrate two, respective, sets of forest structure models. Because periodic FIA dates have approximate lag-time separations with respect to the 1985 sensor data of 8 and 5 years, our hypothesis was that the 1990 periodic FIA data would produce superior calibration results. For this purpose, we selected Landsat image data from WRS-2 P27, as it overlaps P26 and P28 by roughly 70% at this latitude and encompasses the majority of the study area (Figure 2). We then developed five additional Landsat-FIA regression calibration models to estimate TOTBA and RBA of species and species groups for 1985 using (1) all FIA data, (2) all HET FIA, (3) all HOM FIA, (4) 1977 HOM FIA, and (5) 1990 HOM FIA. Further stratification of HET FIA plots by inventory year was not performed due to small sample sizes. As before, model performances were evaluated based on leave-one-out cross-validation (LOOCV) for each respective set of periodic FIA calibration data used. Preferred calibrations selected for mapping were characterized as having superior adjusted coefficient of determinations (R^2_{adj}) of each forest species or group of species, predicted residual error sum of squares (PRESS) statistics, and a relatively concise set of image predictor variables. The PRESS statistic is computed on centre-scaled data and provides an overall measure of model performance (Allen 1971; Wolter et al. 2008).

In addition, we tested the relative difference in TOTBA by using a non-parametric paired t-test (Wilcoxon signed-rank) for 1977 vs 1990 and 1990 vs 2003–2006 FIA plots that share the same physical ground location (n = 122 and 58, respectively). Further, paired two-tailed student t-tests were conducted to assess the significant difference in forest structure models performance between the inventory period (1977 vs 1990), WRS-2 paths and FIA plot designs (annual vs periodic).

2.5. Assessment of mapping sensitivity

Receiver Operating Characteristics (ROC) curves (Hanley and McNeil 1982; Wolter and Townsend 2011) were applied to the final, continuous estimates to identify the true lower boundary of mapping sensitivity relative to false positives for different BA or RBA cut-off points. Here, we used the Youden index (Youden 1950) as a criterion to assess the maximum of vertical distance of ROC curve from the point on diagonal chance line for different BA cut-off iterations, which maximizes the difference between sensitivity and 1-specificity. Hence, by maximizing sensitivity plus specificity across various thresholds, the optimal lower cut-off point was determined (Youden 1950).

3. Results

3.1. Model performance by FIA stratification

3.1.1. Periodic FIA and forest spectral variability

Stratifying Landsat-FIA model calibrations by forest spectral variability proximal to periodic FIA plot centers, irrespective of inventory date, revealed no significant differences (all p-values > 0.588) in overall performance (based on R^2_{adj} values) between the three sets of models (ALL FIA, ALL HET FIA, and ALL HOM FIA). However, of the three periodic FIA stratifications above, those using spectrally heterogeneous (HET) FIA plots benefited five of the eight RBA models for conifer species and groups (Table 4). Conversely, for total BA and RBA models for broadleaved species and groups, all but one (birch) performed better when spectrally homogeneous (HOM) FIA plots were used (Table 4).

In each instance of periodic FIA data stratification listed above, the Y-block of dependent forest variables (i.e., TOTBA and all RBAs) was calibrated simultaneously as a set of models within one overall xPLS regression run. Stratification by periodic FIA date showed the 1990 HOM FIA model set had a lower PRESS value (Table 4, 78.59%) than all other model sets, where six of the thirteen forest structure models had superior overall R^2_{adj} values. In terms of structure model sets, five of the eight conifer models calibrated using ALL HET FIA plots had the best overall R^2_{adj} values (Table 4). We did not produce maps from this latter set of models due to concerns over small sample size (e.g., 29 for big-pines; 19 for both jack pine and tamarack; and 22 for cedar).

3.1.2. Forest spectral homogeneity and periodic FIA date (1977 vs 1990)

Stratification of spectrally homogeneous periodic FIA plots (HOM FIA) by inventory year (1977 and 1990) for structure model calibrations is shown in Table 4. Models calibrated using just the 1990 HOM FIA data and 1985 Landsat sensor data (ca. 5-year lag) performed significantly better overall in terms of R^2_{adj} values (p-value = 0.0096) than those calibrated using the 1977 HOM FIA inventory with these same Landsat sensor data (ca. 8-year lag). The two 1977 HOM FIA exceptions were big-pines and cedar, despite the greater FIA-to-image time-lag. Moreover, the set of structure models calibrated using the 1990 HOM FIA ground data produced a better PRESS value (78.59%) than did the 1977 HOM FIA model set (83.42%), while the number of image predictor variables in each case was similar at ca. 35. Further analysis of 122 common sample plots, we found that the TOTBA variables between 1977 and 1990 were significantly different (p-value < 0.0001).

Table 4. Forest structure calibration results (R^2_{adj} values and associated sample size (n)) stratified by FIA plot spectral variability proximal to plot centre and periodic FIA date (ca.1985) for WRS-2 P27.

	All FIA		All HET FIA		ALL HOM FIA		1977 HOM FIA		1990 HOM FIA	
Stratification	R^2_{adj}	n	R^2_{adj}	n	R^2_{adj}	n	R^2_{adj}	n	R^2_{adj}	n
TOTBA	0.26	632	0.15	101	0.31	550	0.25	323	**0.36**	371
Ash	0.56	179	0.32	7	0.59	172	0.62	97	**0.65**	122
Aspen	0.59	450	0.43	58	0.63	404	0.56	239	**0.67**	267
Maple	0.50	230	0.19	17	0.52	214	0.53	101	**0.58**	159
Birch	0.36	435	**0.40**	48	0.35	395	0.32	224	0.37	265
Pine	0.39	147	**0.65**	36	0.34	117	0.35	62	0.39	81
Big-pines	0.39	111	**0.59**	29	0.31	88	0.37	47	0.35	59
Jack pine	0.14	68	**0.30**	19	0.13	51	0.17	25	0.19	39
Balsam fir	0.28	422	0.21	62	0.33	373	0.37	207	**0.40**	258
Spruce	0.75	306	**0.82**	73	0.72	248	0.69	127	0.79	184
Black spruce	0.77	196	**0.82**	62	0.73	148	0.70	75	0.79	112
Cedar	0.40	146	0.47	22	0.42	131	**0.49**	84	0.38	87
Tamarack	0.44	73	0.09	19	0.50	56	0.47	29	**0.61**	45
PRESS	79.37		85.88		79.72		83.42		**78.59**	
No. of Variables	39		17		38		35		36	

Note: Overall model PRESS (predicted residual error sum of squares) statistics range from zero to hundred, where lower is better. Model R^2_{adj} values shown in bold represent the best structure calibrations across the five FIA stratifications listed.

3.1.3. Periodic FIA stratification and overall model performance

With the exception of the 1990 HOM FIA model set, all other model sets calibrated using 1985 Landsat variables and periodic FIA ground data stratified in various ways (Table 4) did not yield significant differences in overall model performance (all p-values > 0.588). Of these, the All HET FIA set of models used the fewest image predictors (17) compared to 35–39 for the other sets, but this was also coupled with the highest overall model PRESS value and the poorest RBA estimate for SBW host species (R^2_{adj} = 0.21, balsam fir). Moreover, the All HET FIA set of species models also produced the least accurate estimates of RBA for three out of the four broadleaved species (Table 4).

Test stratifications of these periodic FIA inventory data revealed superior calibration results for certain species or groups of species here and there at different levels. However, the best overall model calibration set between periodic FIA inventory data and the 1985 Landsat sensor data resulted from the use of the 1990 HOM FIA ground data. Moreover, this model set also produced the best estimates for SBW host species (balsam fir, white spruce and black spruce) (Table 4), which is germane to our specific research needs. Hence, the 1990 HOM FIA model set was selected for mapping forest species abundance and distribution in the ca. 1985-time step.

3.1.4. Periodic vs. annual FIA plot stratification

Separate forest structure calibration tests using just the spectrally homogeneous (HOM) FIA plots, but stratified by the two plot designs (periodic [1990] vs. annual [ca. 2005], Figure 1) for the ca. 1985 and 2005 Landsat image time frames, respectively, are shown in Tables 5 and 6. Two of the three 1985 forest structure model sets (WRS-2 P26, P27) calibrated using 1990 HOM periodic FIA data performed significantly better (p-values = 0.0202 and < 0.0001, respectively) than their 2005 FIA companion. However, performance between respective WRS-2 P28 model sets calibrated with respective versions of FIA was not significantly different (p-value > 0.2381). Closer examination of common sample plots (n = 58) to annual (2003) and periodic (1990), we found TOTBA variable was not significantly different (p-value = 0.249). When compared by FIA

Table 5. Structure models calibrated using only spectrally homogenous (HO) 1990 periodic FIA plots (*n*) and ca. 1985 Landsat imagery. Accuracy metrics (R^2_{adj}, LOOCV, Test-RMSE) for ca. 1985 forest structure models are presented for each WRS-2 paths.

	P26			P27			P28		
Species	R^2_{adj} (*n*)	LOOCV	Test-RMSE	R^2_{adj} (*n*)	LOOCV	Test-RMSE	R^2_{adj} (*n*)	LOOCV	Test-RMSE
TOTBA	0.34 (204)	3.00	0.70	**0.37** (371)	3.45	0.42	0.31 (507)	3.18	0.48
Ash	0.19 (27)	3.28	1.45	**0.64** (122)	7.25	1.06	0.57 (205)	7.88	0.91
Aspen	0.54 (135)	12.02	1.31	0.63 (267)	13.92	0.37	**0.64** (368)	15.56	1.35
Maple	**0.82** (65)	8.25	1.02	0.55 (159)	7.62	1.03	0.44 (184)	5.79	1.14
Birch	**0.52** (152)	11.13	1.29	0.34 (265)	8.28	0.26	0.21 (305)	5.31	1.42
Pine	**0.43** (45)	8.31	1.91	0.37 (81)	7.94	0.74	0.41 (88)	8.91	0.01
Big-pines	**0.41** (30)	5.50	0.03	0.32 (59)	6.05	1.46	0.32 (70)	5.61	0.15
Jack pine	0.21 (22)	5.18	1.88	0.17 (39)	3.18	0.73	**0.31** (41)	5.74	0.16
Balsam fir	**0.43** (146)	7.03	0.64	0.36 (258)	6.66	0.14	0.42 (260)	6.84	3.08
Spruce	0.77 (131)	12.92	2.10	**0.78** (184)	10.84	2.56	0.75 (194)	11.46	0.29
Black spruce	0.78 (83)	13.44	1.65	**0.79** (112)	10.65	1.36	0.74 (125)	10.50	0.02
Cedar	0.32 (45)	8.56	2.77	0.36 (87)	9.12	0.20	**0.44** (111)	10.73	0.87
Tamarack	0.45 (32)	5.32	0.77	0.54 (45)	7.19	0.51	**0.64** (80)	8.31	1.00
BLF	**0.90** (172)	10.56	0.96	0.82 (238)	13.43	0.52	0.86 (360)	13.55	0.68
CON	**0.91** (177)	10.38	0.95	0.82 (277)	13.45	0.44	0.86 (344)	13.50	0.96
SBW host	**0.67** (169)	14.67	2.74	0.66 (277)	12.98	2.70	**0.67** (299)	13.13	2.79

Note: Metrics LOOCV (leave-one-out cross-validation) and test-RMSE (Root Mean Sum of Squares) were computed during model calibration and model testing, respectively. Units for LOOCV and test-RMSE are same for TOTBA (m^2 ha^{-1}) and RBAs of species (%) models. Model R^2_{adj} values shown in bold represent the best overall calibration across three WRS-2 paths.

Table 6. Structure models calibrated using annual FIA plots (*n*) inventoried from 2003–2006 and ca. 2005 Landsat imagery. Accuracy metrics (R^2_{adj}, LOOCV, Test-RMSE) for ca. 2005 forest structure models are presented for each WRS-2 paths.

	P26			P27			P28		
Species	R^2_{adj} (*n*)	LOOCV	Test-RMSE	R^2_{adj} (*n*)	LOOCV	Test-RMSE	R^2_{adj} (*n*)	LOOCV	Test-RMSE
TOTBA	0.31 (182)	3.71	2.23	0.29 (428)	3.34	1.02	**0.39** (443)	3.98	1.09
Ash	0.28 (41)	11.81	1.16	0.50 (300)	9.31	3.60	**0.59** (190)	13.80	0.86
Aspen	0.51 (124)	22.08	5.44	0.48 (139)	15.52	0.32	**0.57** (302)	24.31	3.01
Maple	**0.76** (74)	12.91	1.12	0.43 (179)	7.39	1.51	0.45 (150)	4.98	1.90
Birch	**0.44** (131)	18.58	2.93	0.22 (266)	14.83	0.59	0.15 (232)	3.55	2.30
Pine	0.29 (24)	13.83	0.24	0.27 (84)	15.60	1.29	**0.46** (68)	9.76	4.55
Big-pines	0.29 (13)	14.90	0.96	0.26 (60)	5.63	1.47	**0.42** (61)	7.59	2.97
Jack pine	**0.31** (15)	9.53	0.72	0.14 (42)	9.78	0.18	0.20 (27)	3.85	1.57
Balsam fir	**0.48** (149)	8.18	0.06	0.29(304)	6.12	1.85	0.28 (208)	5.57	0.14
Spruce	0.61 (105)	10.26	13.83	0.62 (210)	12.19	3.05	**0.72** (159)	9.84	1.97
Black spruce	0.31 (72)	11.54	1.37	0.57 (150)	11.95	0.50	**0.77** (107)	9.69	0.72
Cedar	0.32 (43)	22.66	6.48	0.32 (94)	19.28	0.68	**0.49** (109)	13.27	1.45
Tamarack	0.28 (15)	7.05	1.52	**0.51** (52)	7.26	0.36	0.30 (64)	11.46	4.76
BLF	**0.83** (165)	11.92	1.93	0.72 (323)	15.82	4.60	0.80 (379)	15.02	2.80
CON	0.84 (166)	11.89	0.32	0.76 (319)	15.66	3.61	**0.85** (308)	14.23	3.11
SBW host	**0.64** (165)	11.82	13.77	0.56 (307)	13.18	1.20	0.63 (269)	11.67	2.10

Note: Metrics LOOCV (leave-one-out cross-validation) and test-RMSE (Root Mean Sum of Squares) were computed during model calibration and model testing, respectively. Units for LOOCV and Test-RMSE are same for TOTBA (m^2 ha^{-1}) and RBAs of species (%) models. Model R^2_{adj} values shown in bold represent the best structure calibrations across three WRS-2 paths.

protocol (periodic vs annual) according to two general functional groups (i.e., broadleaved and conifer species and groups) across the three WRS-2 paths, the periodic 1990 FIA to 1985 Landsat model calibrations (Table 5) were significantly better for both broadleaved (five models) and conifer models (10 models) in P27 (p-values each < 0.0008) than their P27 complements using both 2005 Landsat and annual FIA data (Table 6). In P26, the 1990 periodic FIA-calibrated set of conifer RBA models performed significantly better (p-value = 0.0372)

compared to the 2005 annual FIA-calibrated set of conifer RBA models, while the broadleaved species RBA models were similar between dates (p-value = 0.7483). Moreover, all individual P27 models and all but three P26 models (ash, jack pine, and balsam fir) calibrated using 1990 periodic FIA outperformed the 2005 FIA-Landsat calibrations. The westernmost path (P28) was more evenly split between the two FIA plot designs, with the 1990 periodic FIA calibrations appearing to perform slightly better for broadleaved species (p-value = 0.0867) compared to conifers (p-value = 0.1519). For the three WRS-2 paths and two FIA plot designs the number of FIA calibration plots for jack pine (average n = 31) and ash (average n = 63) were relatively low compared to the number of plots used for most other model calibrations, especially for P26 (Tables 5 and 6).

3.2. Structure model performance by WRS-2 path and Landsat image dates

The amount of variation in the 1990 periodic FIA data explained by the 1985 image predictor variables for each respective structure variable across the three WRS-2 paths ranged from low (e.g., R^2_{adj} = 0.19 and 0.17 for black ash and jack pine, respectively) to high (e.g., R^2_{adj} = 0.79 and 0.82 for black spruce and maple, respectively) (Table 5). While the magnitude of R^2_{adj} and RMSE values for individual Landsat-FIA structure model calibrations for species and group RBA varied substantially within a particular WRS-2 path, this path-wise variability was similar across the three WRS-2 paths (Table 5), with a few notable exceptions in P26 (e.g., maple and birch), P27 (ash) and P28 (e.g., tamarack, cedar, and jack pine). Thus, accuracy for modelled estimates of SBW host remained consistent with R^2_{adj} values of ca. 0.67 across the three WRS-2 paths for 1985. Separate R^2_{adj} values for the spruce (white spruce and black spruce) and balsam fir RBA estimates ranged from 0.75 to 0.78 and 0.36 to 0.43, respectively (Table 5).

With respect to the ca. 2005 structure model calibrations using concurrent annual FIA data (Table 6), path-wise performance among the individual structure models also varied substantially, with R^2_{adj} and RMSE values ranging from 0.14 to 0.85 and 0.06–13.8, respectively. However, as was observed in 1985, this path-wise variability was similar among the three WRS-2 paths. Nevertheless, the 1990 models for P28 were significantly better (p-value = 0.0076) as a set than the other two WRS-2 paths, where nine out of 16 P28 models were superior to those from the other two Landsat paths. Moreover, three P28 exceptions (tamarack, BLF, and SBW) were only marginally outperformed by their P26 counterparts. Coupled with the superior performance of the P28 models is the fact that this path had significantly more FIA plots available for calibration in 1990 than P26 (p-value < 0.0001), while sample sizes between P28 and P27 were not significantly different (p-value = 0.4782).

Model results for SBW host species RBA (spruce, black spruce, and balsam fir) were mixed within a given WRS-2 path but generally consistent by path across the two mapping periods (Tables 5 and 6). The variation accounted for by the spruce models (spruce and black spruce) across both dates ranged from moderate to high (R^2_{adj} range 0.61 to 0.78, RMSE range 0.02–13.83%). While the balsam fir RBA model results were generally poor across all WRS-2 paths and image dates, results within P26 for both image dates were higher than the other paths (e.g., R^2_{adj} = 0.43 and 0.48 for 1985 and 2005, respectively) irrespective of FIA plot-design differences. Balsam fir RBA model cross-validation results yielded RMSE values ranging from 5.57% to 8.18% of their observed TOTBA values. Calibration of models improved substantially when we combined FIA data

into more generalized dependent variable groups: broadleaved [BLF], conifers [CON], and SBW host (i.e., R^2_{adj} ranged 0.72–0.83 and 0.76–0.85, for BLF and CON, respectively). Lastly, accuracy among the combined SBW host species RBA models was moderate across the three paths and two calibration dates (average R^2_{adj} = 0.64 and RMSE = 12.91%), which is of particular importance to this study (Tables 5 and 6).

3.3. Salient Landsat predictors for 1985 and 2005 mapping periods

The relative importance of respective image predictors varied more by date than they did by WRS-2 path. In some instances, known forest phenology windows used to target Landsat data acquisition for each WRS-2 path and date were not perfectly captured. This was due to annual variation in phenology (see Ahlgren 1957), TM orbital phase mismatches (16-day revisit), or excessive cloud cover. Therefore, we present frequency of important Landsat predictors for all model sets (Figure 3, scaled loadings ≥ |0.8|) and the relative importance of Landsat predictors for 1985 (Figure 4) and 2005 (Figure 5) for P27 only.

In general, Landsat indices such as NDVI, NDMI, NBR, MSI, SVR, and Tasselled Cap transformations (Table 3) were more frequently retained by the xPLS routine during model calibration than were individual sensor bands (Figure 3). There were also clear differences in the number of predictors and magnitude of predictor importance between the two calibration periods (Figures 4 and 5). For the 1985 (Figure 4), the aspen and ash species RBA models show strong (scaled loadings ≥ |0.8|) dependence on indices from two seasons (April [P] and September [S2], respectively), while the spruce models (black spruce and spruce) and tamarack models had a wider profile of predictor importance, but with less emphasis in September (Table 2). Conversely, the 2005 calibration period involved nearly twice the number of image predictors with a more even distribution of strong variable loadings among the various models (Figure 5).

3.4. Receiver operator curve (ROC) analysis: gauging model sensitivity

Receiver operator curve (ROC) analysis has been used to identify the optimal lower threshold value at which the true-positive rate of identification is maximized and false-positive rate is minimized so as to mitigate the chances of over mapping erroneously low RBA values (Wolter and Townsend 2011). Here, optimal RBA threshold values for each of the respective forest structure models for 1985 and 2005 are shown in Table 7. In general, optimal RBA cut-off value for all forest structure models was not significantly different between the two time-periods (p-value = 0.6478). Broadleaved species had higher RBA threshold values than conifer species, with the exception of ash. Aspen had the highest threshold value at 27.97% and 26.16% RBA for 1985 and 2005, respectively. Optimal lower RBA values for the pine groups (i.e., pine, big-pine, and jack pine) range from 5.3% to 12.12%, while the balsam fir RBA threshold was about 11.50% for both 1985 and 2005 periods. The two spruce categories (spruce and black spruce) and cedar had relatively high RBA threshold values (12.23–17.63%) among conifer species RBA models. Optimal threshold values for the remaining three combined species models (BLF, CON, and SBW host) were somewhat higher than individual species and genus-level RBA models (Table 7). We used the optimal lower RBA threshold values during the production of final forest composition and abundance maps for

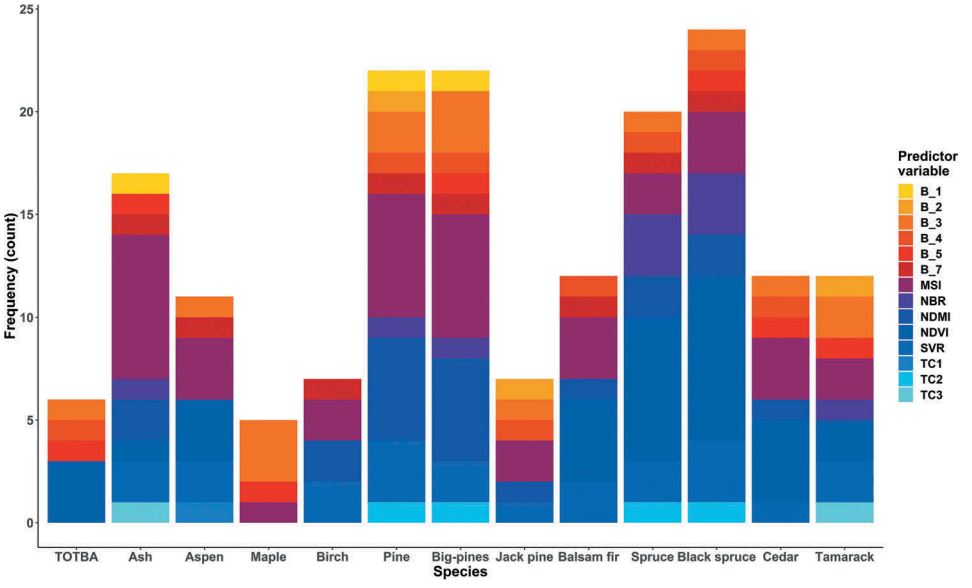

Figure 3. Frequency distribution of the salient (≥ |0.8|) centre-scaled image predictor variables retained by xPLS regression during Landsat-FIA model calibrations for total basal area (TOTBA) and the relative basal areas (RBA) of species and groups of species.

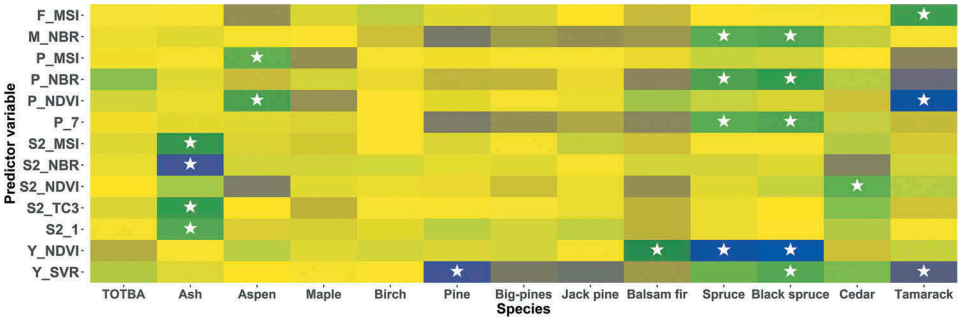

Figure 4. Landsat TM image predictor variables for WRS-2 P27 ca. 1985. The loading of each predictor is displayed in a continuous range of colour from blue (high negative importance) through intermediate yellow (low importance) to green (high positive importance). Stars indicate image predictor variables that had absolute loadings greater than or equal to 0.8.

each of the two periods. However, due to the sheer volume of graphic data, we chose to display only the spatial distribution maps of SBW host for each mapping period (Figure 6).

4. Discussion

4.1. *Influences of FIA stratification on model calibration*

4.1.1. *Spectral heterogeneity*

Overall, stratification of FIA plots according to forest neighbourhood spectral variability in the SWIR did not have a significant impact on ground-to-satellite model calibration

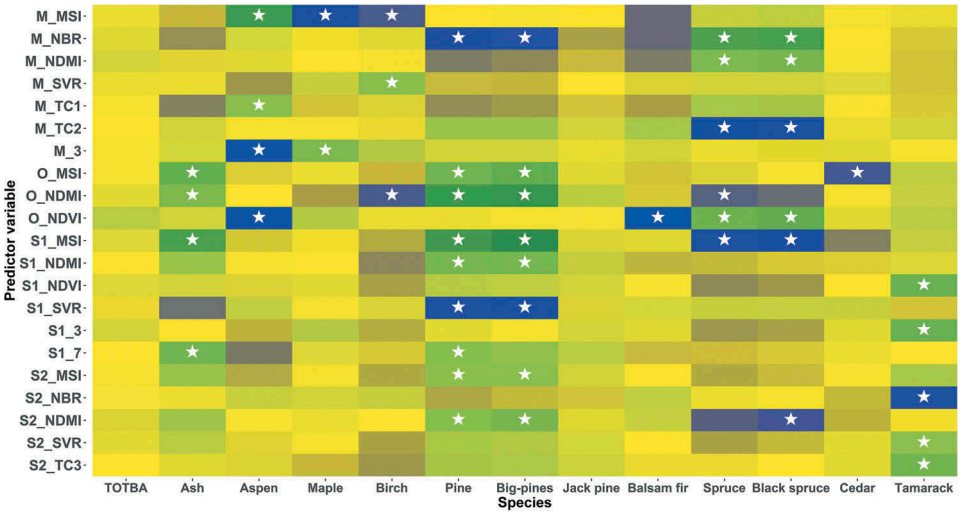

Figure 5. Landsat TM image predictor variables for WRS-2 P27 ca. 2005. The loading of each predictor is displayed in a continuous range of colour from blue (high negative importance) through intermediate yellow (low importance) to green (high positive importance). Stars indicate image predictor variables that had absolute loadings greater than or equal to 0.8.

Table 7. Minimum relative basal area (RBA, %) threshold values used for mapping respective species and species groups for ca. 1985 and ca. 2005 forest structure layers. Minimum values were calculated using receiver operating characteristic (ROC) curve analyses and assessed via the Youden index (Youden 1950).

	Minimum relative basal area (RBA, %)	
Species	ca.1985	ca.2005
Ash	6.85	10.45
Aspen	27.97	26.16
Maple	16.08	11.97
Birch	14.51	14.51
Pine	9.13	9.13
Big-pines	8.09	12.12
Jack pine	7.05	5.30
Balsam fir	11.38	12.01
Spruce	15.16	12.20
Black spruce	17.63	17.63
Cedar	16.66	16.79
Tamarack	11.68	7.51
BLF	23.21	28.99
CON	21.20	15.88
SBW host	22.16	22.42

performance. However, HOM stratification is likely optimal for ash, maple, and aspen because they frequently grow in relatively pure stands in this region (Wolter and Townsend 2011), while birch diverges from this generalization. Birch can be found in pure stands and in mixtures with aspen adjacent to Lake Superior, but farther inland birch is more commonly found in mixtures with conifers and a wider variety of other broad-leaved species (Wolter and Townsend 2011). Notably, while there were many HOM birch

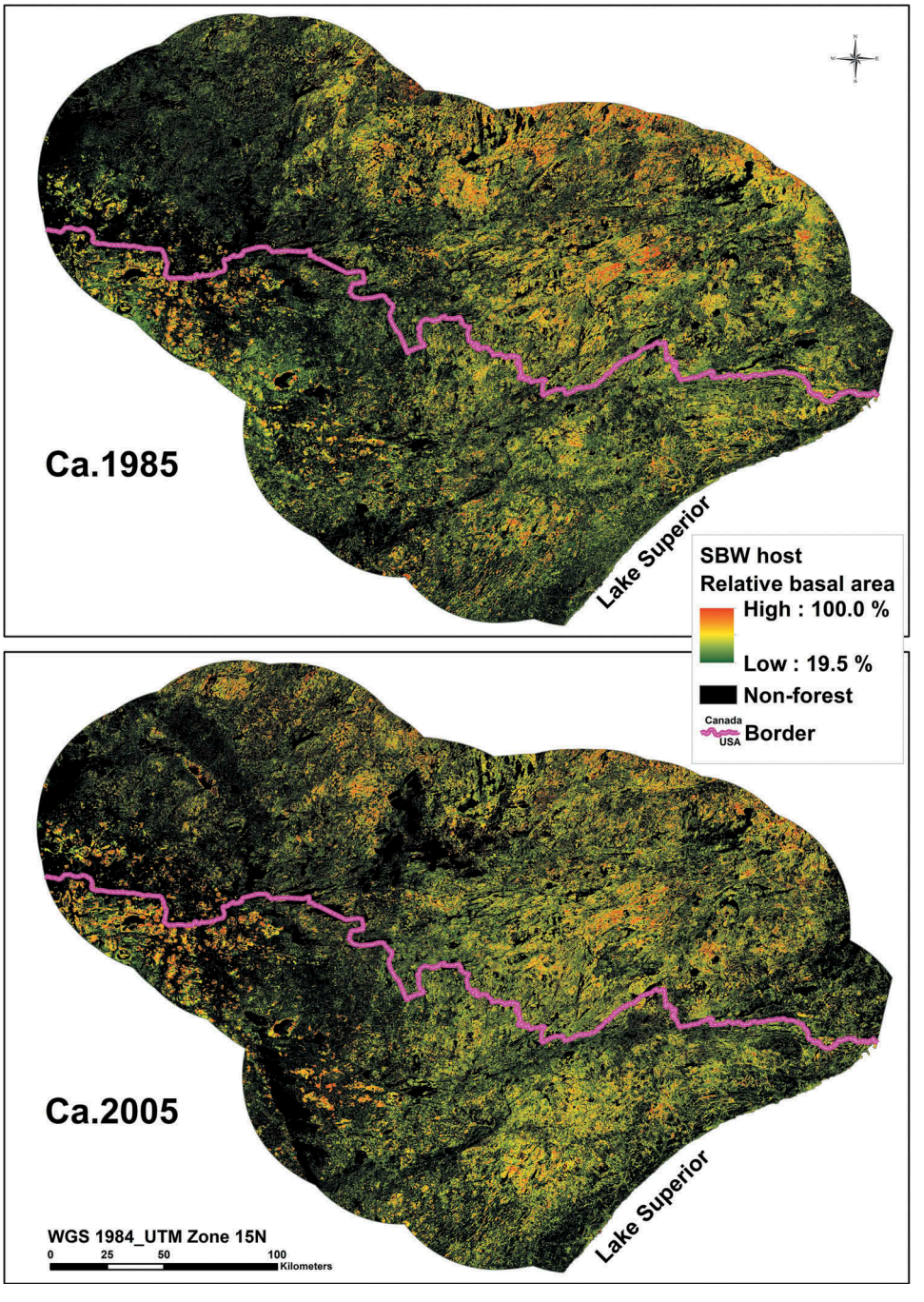

Figure 6. Spatial distribution of spruce budworm (SBW) host species: balsam fir (*Abies balsamea*) and spruce (*Picea spp.*) abundance for ca.1985 and ca. 2005 mapping periods.

plots (n = 395) compared to HET birch plots (n = 48) based on SWIR-5 forest reflectance, only 25 plots contained birch RBA greater than 50%, of which, only three HOM birch plots

may be considered pure birch stands (RBA of birch >70%). Simply put, plots with birch tended to be spectrally homogenous, regardless of associated species composition.

In contrast, among the conifers, balsam fir, white spruce, upland black spruce, and the three pine species do not generally grow in pure stands of similar patch sizes as broadleaved species. Hence, while spectral variability in Landsat SWIR-5 reflectance may be low (i.e., either pure conifer or pure broadleaved), conifer species diversity could be high. Tamarack (almost exclusively a wetland species), lowland black spruce, and pure plantations of the three pine species are exceptions. For example, considerable areas of tamarack are abundant within the study area (Wolter et al. 2012b), which we believe explains the higher performance of HOM over HET tamarack model calibrations. This also explains the opposite scenario for HET black spruce models (Table 4), as inter-mixing of black spruce with jack pine, white spruce, balsam fir, and broadleaved species in uplands of the northern half of the study area is common (Dimitrov, Bhatti, and Grant 2014).

Stratifying FIA ground data by Landsat-measured surface SWIR reflectance prior to structure model calibration with Landsat sensor variables shed light on the effects of unique species compositions on species mapping across this landscape. While the HOM spectral neighbourhood stratification produced calibration accuracy results higher than those where all FIA data were used (p-value = 0.0035), we do not recommend stratifying FIA ground data by Landsat surface reflectance in the future. Though not attempted here, we also do not recommend stratification of FIA plots by purity within the FIA database to streamline or improve similar calibration efforts using Landsat sensor data due to substantial spatial uncertainty in both data sources. Spatial uncertainties include suboptimal FIA subplot spatial orientation and size with respect to 30-meter Landsat pixel (Figure 1), spatial location accuracy of FIA plot centers (Hoppus and Lister 2007; Miles, Chen, and Leatherberry 1995), and geo-registration accuracy of Landsat pixels (Irons, Dwyer, and Barsi 2012).

4.1.2. Periodic FIA (1977 vs 1990)

Clear date-wise differences in model performance raised questions of possible effects related to variability in either changes in FIA sampling protocol between these dates (Goeking 2015; Miles, Chen, and Leatherberry 1995; Miles et al. 2007) or possible Landsat-FIA temporal lag effects, when all other factors (FIA plot locations, relative sample sizes, image dates, and FIA plot design) were held constant. The ability of the Landsat sensor to detect sub-pixel forest change diminishes with increasing time (Wilson and Sader 2002; Jin and Sader 2005; Coppin and Bauer 1996). Here, however, the apparent relationship between ground inventory and space-based forest reflectance appears to have changed substantially with only a marginal increase in time lag (e.g., 1977–1985 and 1990–1985) between ground and space-based observations. Contrary to our results, Nelson et al. (2011) found an insignificant difference in biomass and forest classification accuracy when Landsat images from 2000 and 2007 were used with 2003 field plot data. The small overall relative difference in time-lag between FIA-image combinations (5 versus 8 years) led us, rather, to suspect variability in the periodic FIA sampling protocols between 1977 and 1990 as a more likely explanation for skewed model performance in favour of the 1990 FIA-Image combination (Table 4). While we cannot say with certainty that this is the case, we did find a significant difference in TOTBA values between the two sampling periods for undisturbed forest (n = 122, p-values < 0.0001). As such, Goeking (2015) warns FIA data users of the pitfalls of failing

to distinguish between actual and apparent change resulting from different inventory designs.

There are several potentially significant factors with respect to the 1977 and 1990 periodic FIA sampling protocols that could have contributed to differences in model calibration results shown in Table 4. First, to keep costs down during the 1990 inventory period, two-thirds of all forested FIA plots in 1977 that were classified as undisturbed by 1990 in northern Minnesota – via interpretation of aerial photography – were 'modeled' rather than actually revisited in the field to update forest biophysical parameters for the 1990 FIA period (Miles, Chen, and Leatherberry 1995). The remaining one-third of undisturbed forest plots were both re-measured and modelled as a calibration/validation step for this modelling process (Miles, Chen, and Leatherberry 1995). Modelling was performed using the 1977 FIA field data and individual tree growth models to simulate plot-wise re-measurement, which then became the basis for the 1990 periodic FIA for undisturbed forest plots (Hansen et al. 1990; Miles, Chen, and Leatherberry 1995). However, data used in this study from the two dates (1977 and 1990) within the northeastern and north-central regions of Minnesota are recorded as not being modelled, but the reliability of these data entries is unknown (M.D. Nelson USFS pers. comm.). Second, while the 1977 and 1990 FIA plot designs each consisted of 10 variable-radius subplots (Figure 1), there are some clues that during the 1977 inventory there were more instances where fewer than 10 subplots were actually measured per full plot compared to the 1990 sample period (Miles, Chen, and Leatherberry 1995). Third, many of the 1977 FIA plots that were selected for field revisits in 1990 inventory period could not be located. Hence, the remedy was to replace the older plot with a new plot at the 'approximate location' of the 1977 plot centre, but neither the absolute or relative positional accuracy with reference to the old FIA plot centres was documented (Miles, Chen, and Leatherberry 1995).

4.1.3. Periodic vs annual plot design

The comparatively better performance of our periodic FIA-based models over those using annual FIA data (Table 5, 6) is not as easily explained by differences between FIA sampling protocols, since different complements of image predictor variables were used (Table 2). However, with the exception of P26 images were more closely matched in time with their respective sets of FIA ground data (10.4-, 2.3-, and 6.3-year ranges for P26-P28, respectively). Though this is usually preferable in terms of minimizing potential time-lag issues, finding the full complement of species-specific phenology images known to be important in this region for identifying forest species (Wolter et al. 1995; Wolter and Townsend 2011; Wilson, Lister, and Riemann 2012) was not always possible. For instance, early September is a key date for identifying stands of black ash in this region, which tend to drop their leaves in early autumn long before other broadleaved species (Eder 1989; Wolter et al. 1995). From Table 2, 5, and 6, in each instance where an early September image is missing (i.e., P26 in 1990 and 2005), the associated ash RBA model performance suffered (R^2_{adj} 0.19 and 0.28, respectively, compared to an average R^2_{adj} of 0.58 otherwise for ash). In other cases, where Landsat images from within general phenology timeframes were acquired, slight shifts in peak phenology due to inter-annual variability (Ahlgren 1957) likely dampened the effectiveness of these predictors. Hence, not all differences in Landsat-FIA model calibration performance can be attributed to changes in ground sampling

protocol (Figure 1) that took place upon adoption of the annual FIA strategy (Bechtold and Patterson 2005; Vandendriesche and Haugen 2008).

Despite this, we suspect the more numerous, more tightly spaced, variable-radius subplots around plot center under the older, periodic FIA design better represents ground reflectance in the neighbourhood of a single Landsat pixel (Figure 1). The closest six variable-radius periodic FIA subplots are 21.3 m from plot center. The three annual FIA fixed-radius (7.32 m) subplots are 36.6 m from plot center, which represents an additional 15.3 meters (Figure 1). Under the periodic FIA subplot design, if a Landsat pixel is registered to the ground in error by up to one half a pixel (15 m) the surface reflectance recorded by the sensor would still be represented, in part, by ground information from as many as 4–7 of the 10 subplots depending on the direction of error. Moreover, substantial error inherent in autonomous code-phase GPS positioning of FIA plot centers only adds spatial uncertainty to this scenario (McRoberts 2010). While both FIA plot designs clearly include information from outside this theoretical spatial error zone of a single Landsat pixel, the salient difference is that the annual FIA plot design leaves the majority of this error zone unsampled (Figure 1). Therefore, a Landsat pixel with modest spatial registration error is represented by just one of the four subplots under the annual FIA design, which represents 18.7% of a 30-meter pixel. As mentioned earlier, under idyllic homogenous forest conditions such spatial sampling differences may be of little consequence, but become increasingly relevant as forest spatial heterogeneity increases. One could simply select the center annual FIA subplot (or any such subplot that appears to coincide with a pixel) and ignore data from the remaining subplots, but uncertainty in spatial accuracy among both FIA plot centers and Landsat pixels may bring into question the wisdom of such a decision (Franco-Lopez, Ek, and Bauer 2001; McRoberts, Tomppo, and Næsset 2010).

There are several examples in the literature where ground inventory data were paired with satellite sensor data for quantifying forest structure with little or no discussion of the potential impacts of sampling design/protocol or time-lags on modelled forest structure results (Franco-Lopez, Ek, and Bauer 2001; McRoberts et al. 2007; Ohmann and Gregory 2002; Wilson, Lister, and Riemann 2012). However, performance between our FIA-calibrated forest TOTBA and RBA models and those from similar studies in this region supports the conjecture that the difference among ground sampling designs is a salient factor. Working in the same area, Wolter et al. (2008) achieved substantially better calibration accuracies for total BA and balsam fir RBA (average R^2_{adj} 0.65 and 0.69, respectively) by using ground subplot sampling designed specifically for integration with 30-meter Landsat sensor data: four variable-radius subplots arranged orthogonally around, and 30 m from, a central subplot. From Figure 1, one can see that neither the periodic or annual FIA full plot designs match well with a single 30-meter Landsat sensor pixel, especially given the positional uncertainties in periodic FIA discussed above. By comparison, Franklin (1986) used various circular plot sizes (100 m^2, 250 m^2, or 314 m^2) depending on forest stand purity to collect field data in order to improve the model accuracy. Our results support their assertion that matching ground plot sampling designs with either sensor spatial resolution characteristics or forest composition characteristics, or both, are important factors to consider when upscaling plot-level data to landscapes using satellite sensor data.

Other factors, as discussed, may certainly confound the accuracy potential of satellite-based forest structure models calibrated using ground plot data, such as ground plot positional accuracy (Hoppus and Lister 2007; McRoberts 2010), annual variation in forest

phenology aids, critical phenology gaps in the satellite archive (Wolter and Townsend 2011), or variability in species abundance across the landscape discussed below. However, based on comparisons to past research in this heterogeneous forest region (Wolter et al. 2008; Wolter and Townsend 2011), we believe FIA subplot configuration was a key factor affecting our ability to produce more accurate, continuous map estimates of past forest structure conditions (Figure 1, Tables 4, 5 and 6).

4.2. Species abundance and modelling accuracy

Previous studies mention the inherent variability in species abundance across the three Landsat WRS-2 paths that cover this region of northeastern Minnesota (Franco-Lopez, Ek, and Bauer 2001; Wolter et al. 2008). Franco-Lopez, Ek, and Bauer (2001) reported classification accuracies of 68%, 18%, and 0% accuracy for species from high to low abundance in their study area (aspen, balsam fir, and white pine, respectively). The abundance-to-calibration accuracy conjecture may hold in cases where tree species are 1) well represented in upper canopy positions and 2) have RBA levels within an FIA plot high enough to be detected by the satellite sensor (Figure 7–9). For example, spruce species are generally considered common throughout this region of Minnesota (Wolter et al. 2008; Wolter and Townsend 2011), but typically occur in lower to moderate abundance (Figure 7), with the exception of more sparsely distributed lowland black spruce stands that frequently achieve 100% RBA in this region. Analysis of our FIA ground data (i.e., 338 and 325 FIA plots contained either white or black spruce in 1985 and 2005, respectively) shows that the RBA of 'spruce' is frequently less than 35% (Figure 7). However, since spruces typically share over-story position with associated tree species, calibration results for RBA of the spruce group were among the most accurately modelled species or genus group in this study, with an average R^2_{adj} across the two time periods of 0.71 (Tables 5 and 6). Balsam fir, on the other hand, is also ubiquitous on this landscape ($n = 422$, Table 3), with typical FIA plot-wise abundance below 40% RBA [Figure 7]). The key difference is that balsam fir is primarily an understory species (Wolter and Townsend 2011) and, thus, is far less visible to optical satellite sensors in spite of relatively high spatial abundance. As such, balsam fir's unique growth habit explains the poor calibration performance (average $R^2_{adj} = 0.38$ for two time-periods) compared to the spruce model (Tables 5 and 6). In agreement with this conjecture, Wolter et al. (2008) also noted differences in model performance between balsam fir and spruce RBA within this landscape ($R^2_{adj} = 0.69$ and 0.88, respectively).

In other cases, where species appear to have more or less similar abundance in the FIA data, we see clear disparities in model calibration accuracy (Tables 5 and 6) due to species phenology. For instance, tamarack ($n = 187$) and cedar ($n = 202$) often occur in truly pure stands and each has relatively high abundance, as indicated by the number of FIA plots. However, calibration accuracy is, on average, higher for tamarack than cedar. Specifically, tamarack is a deciduous conifer and, hence, our combined use of both summer and winter Landsat sensor data augments detectability of tamarack compared to cedar (see Wolter et al. 1995, 2008). Conversely, mixed results for maple ($n = 667$) and aspen ($n = 355$) across the three WRS-2 paths are more a function of spatial distribution rather than specific phenological differences (Figure 8). For instance, both aspen and maple have strong phenological differences that we took advantage of in our selection of Landsat sensor data (Table 2) following the example of Wolter et al. (1995) and other studies in this region

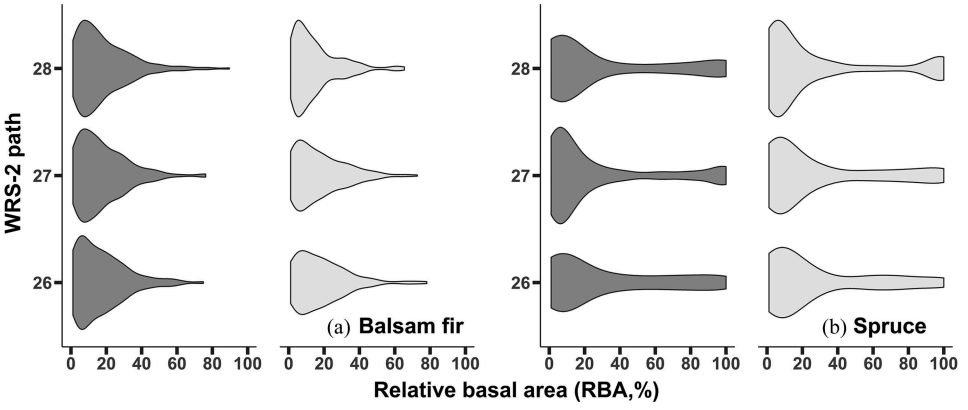

Figure 7. Distribution of FIA plot data for spruce budworm (SBW) host species: (a) balsam fir and (b) spruce used in xPLS regression model calibrations for the ca.1985 (dark grey) and ca. 2005 (light grey) mapping periods by the three WRS-2 paths.

(Wolter et al. 2008). We posit that the reason for the strong divergence in model performance between maple ($R^2_{adj} > 0.76$) and aspen ($R^2_{adj} < 0.51$) in P26 compared to the other WRS-2 paths is linked to the fact that our 'maple' dependent variable includes both sugar and red maple. Sugar maple forms nearly pure stands that exist almost exclusively within the North Shore Highlands ecological subsection that parallels the Lake Superior shore. This lakeshore sugar maple belt is completely captured in P26, partially captured in P27, and completely absent from P28. Red maple, on the other hand, is more evenly distributed across the whole study area but at relatively low RBA levels in association with other hardwood and conifer species. Though each of these maple species has distinct autumn leaf colouration (Eder 1989; Wolter et al. 1995), the generally low RBA of red maple across the landscape serves to largely dampen its Landsat-scale multi-spectral advantage, compared to that of sugar maple, especially within P26 (Tables 5 and 6).

Similarly, the big-pines (white and red combined) and jack pine both have comparable relative abundance (average $n = 49$ and 31, respectively) and distribution (Figure 9). However, the big-pines models were typically better than the jack pine models (Tables 5 and 6). Similar differences in model calibration performance were noted in two previous studies (Wolter et al. 2008; Wolter and Townsend 2011), but logical explanations for observed differences in model calibration performance were not provided. Whether this is due to spectral advantage, growth character, or field plot design with respect to varying forest heterogeneity (Riemann et al. 2010; Franklin 1986; Ohmann, Gregory, and Roberts 2014; Ruefenacht et al. 2008) remains a vexing problem. Nevertheless, our study focuses on the utility of FIA data as a source of ground data to calibrate continuous forest composition models. Whether the various FIA subplot sizes and orientations between the two dates are differentially better for some species assemblages over others was beyond the scope of this study. Given the nature of the discussion above, it is clear why our highly generalized forest structure variables, BLF and CON, were consistently higher compared to the SBW host model or any of the species RBA models (Tables 5 and 6), a trend that is well known (Franklin 1986; Franco-Lopez, Ek, and Bauer 2001; Ohmann and Gregory 2002; Wolter et al. 2008).

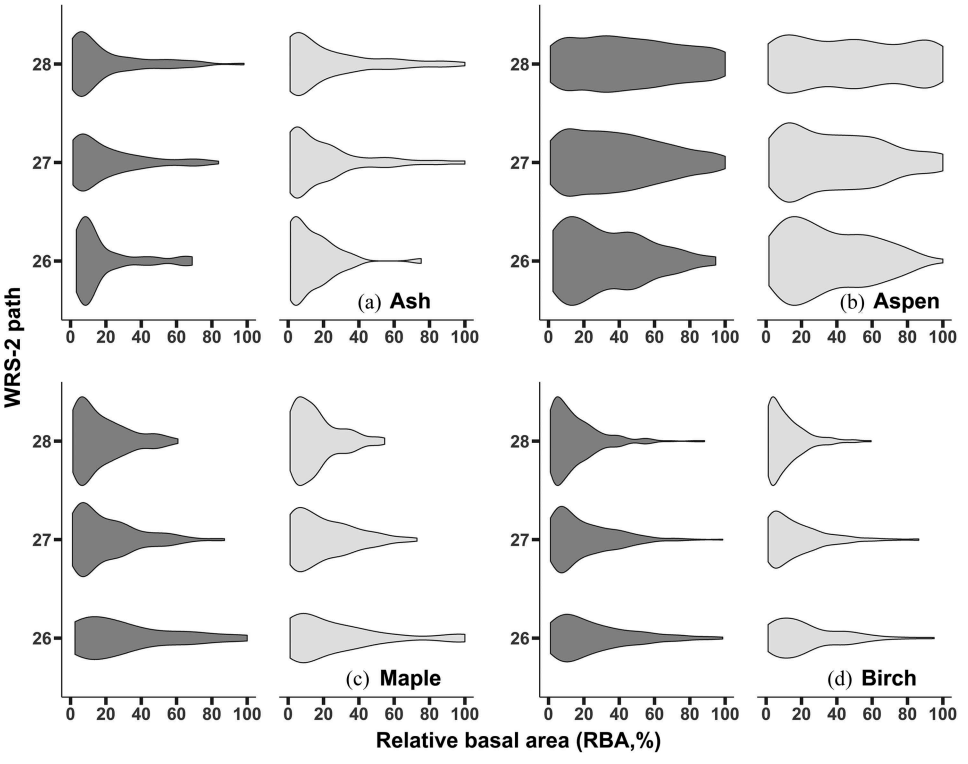

Figure 8. Distribution of FIA plot data for broadleaved species: (a) ash, (b) aspen, (c) maple and (d) birch used in xPLS regression model calibrations for the ca.1985 (dark grey) and ca. 2005 (light grey) mapping periods by the three WRS-2 paths.

We believe forest species abundance estimates derived via the integration of FIA and Landsat sensor data will continue to be poor to only moderately effective for regions of the US where species and spatial heterogeneity are relatively high, as in northern Minnesota (Moore and Bauer 1990). However, the effectiveness of the Landsat-FIA union will likely improve where forest heterogeneities are naturally less complex, such as in the Black Hills of South Dakota, USA (Brown and Cook 2006), or where forest management has had a simplifying effect through time (Ishii, Tanabe, and Hiura 2004). In both cases, the lack of spatial parity between FIA subplot arrangement and satellite sensor spatial resolution should become less relevant as forest homogeneity increases.

Lastly, though comparisons to various modelling approaches in the literature were not an objective of this research, two studies conducted in Minnesota provide insightful comparisons (Franco-Lopez, Ek, and Bauer 2001; Wilson, Knight, and McRoberts 2018), as each combined Landsat imagery and ground data for estimation and mapping purposes. Among three different models: xPLS regression (our method, Wolter et al. 2012), K-NN (Franco-Lopez, Ek, and Bauer 2001), and Random Forest with fourier transformation of Landsat time series data (Wilson, Knight, and McRoberts 2018), precision for TOTBA estimates is better using Random Forest (average RMSE 3.5, 9.1, and 0.33 m^2 ha^{-1}, respectively). In case of the generalized broadleaved group (BLF), xPLS models had explained more variance in the ground data than random forest regression (Wilson, Knight, and McRoberts 2018) (model R^2_{adj} 0.77 vs 0.49).

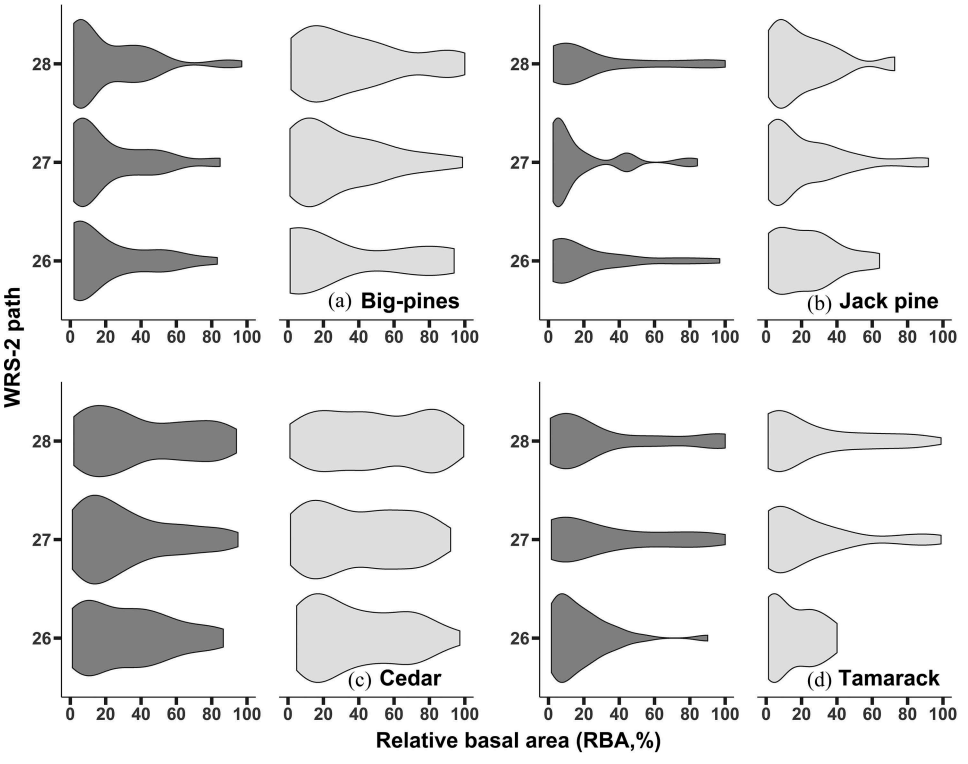

Figure 9. Distribution of FIA plot data for conifer species: (a) big-pines, (b) jack pine, (c) cedar, and (d) tamarack used in xPLS model calibrations for the ca.1985 (dark grey) and ca. 2005 (light grey) mapping periods by the three WRS-2 paths.

However, these metrics of model performance (R^2_{adj} and RMSE) for corresponding species were somewhat better in Wolter et al. (2008) where a sensor-specific sample plot-design was used, and was further improved when integrated with multi-sensor data (C-band SAR (5.3 GHz, Radarsat-1), L-band SAR (1.27 GHz, Palsar-1), Landsat and SPOT-5) (Wolter and Townsend 2011). In particular, results for balsam fir RBA (R^2_{adj} = 0.78, RMSE = 2.26%) were clearly better than any of our balsam fir RBA models calibrated via FIA ground data and Landsat sensor data alone (Tables 5 and 6). Hence, while combination of Landsat-specific ground plot design with ancillary satellite sensor data, such as SAR, has improved forest structure model calibrations, results may be further improved via the use of more elegant modelling approaches.

4.3. Future recommendations for integration of FIA and satellite sensor data

Stratification of FIA plots by Landsat-based spectral variability proximal to FIA plot centre explained unique, species-wise modelling performance issues related to the degree of conifer and hardwood mixture. However, stratifying by crown position and heterogeneity among species or group of species within the FIA database itself may shed more light on how to best use FIA data for mapping forest structure using Landsat sensor data; above and beyond issues related to changes in sampling protocol across years. Also, our decision to calculate plot-wise averages of FIA forest structure information across all subplots

(periodic and annual) rather than selectively picking subplots that appeared to match 30-meter pixel positions was deliberate. This was done expressly to understand differences in effectiveness for modelling forest structure between versions of FIA and to compare results in studies where Landsat-optimized plot designs were used (Wolter et al. 2008; Wolter and Townsend 2011). However, because of the uncertainty in both pixel-to-ground registration and documented uncertainty in FIA plot locations (McRoberts 2010; Miles, Chen, and Leatherberry 1995), we suggest avoiding the temptation to single out individual subplots/pixel combinations for these purposes. The reader should note that new approaches are being developed to improve the positional accuracy of annual FIA plot locations, but periodic FIA locations not coincident with annual FIA will remain unchanged (M.D. Nelson, pers. comm.). Evidence from this study clearly shows, based on the nearly identical study conducted by Wolter et al. (2008) where Landsat-tailored, custom ground plot sampling was used, that neither past nor current FIA ground sampling strategies are optimal for use with 30-metre Landsat sensor data. Although, the more numerous and tighter spatial arrangement of subplots under the older periodic FIA sampling design appears to offer some benefits over the current FIA design (Figure 1), especially under heterogeneous forest conditions. An important similarity between this study and the Wolter et al. (2008) work is that the relative species- and group-specific differences in calibration precision reported here were also evident in Wolter et al. (2008) study, which points to the limitations of the Landsat sensor mentioned above and not FIA. However, the overall lower magnitude of model calibration accuracy compared to that of the Wolter et al. (2008) study, we believe, specially reflects differences in ground plot design. Hence, with reference to the use of periodic FIA, using just the six subplots that are immediately adjacent to the centre subplot (Figure 1) will likely improve FIA-Landsat model calibration results, especially among increasingly heterogeneous forest stands.

Though the Landsat sensor certainly exhibits limitations in the context of this forest mapping study, it does combine relatively high spatial resolution (30-meter, optimized for forest monitoring, Short 1982) with a relatively large footprint size (170 km x183 km) that is considered optimal for national mapping efforts (Vogelmann et al. 2001). Other optical satellite sensors with higher spatial resolution have been shown to incrementally improve the accuracy of forest mapping efforts (Salajanu and Olson 2001), but at the expense of smaller coverage areas and often substantially increased cost. Moreover, the specific issues related to hidden forest understory components discussed above would likely persist.

Alternatively, airborne Lidar data combined with Landsat sensor data are often touted as the solution to canopy penetration issues and have proven highly effective for mapping forest structure in this region (Deo et al. 2017; Engelstad et al. 2019). In other instances, however, such a union was largely ineffective for quantifying understory balsam fir abundance (Engelstad et al. 2019). Hence, we believe further exploration of the fusion of repeat-pass, polar-orbiting SAR sensor data, especially at L-band frequencies (Blomberg et al. 2018; Wolter and Townsend 2011), with multi-temporal Landsat sensor data is warranted, especially where the detection of understory forest layers, such as balsam fir, is critical. With that said, systematic, spatial mismatches between FIA plot configuration and sensor spatial resolution, especially at 30-meter, will likely continue to confound forest structure mapping results.

5. Conclusion

In this study, we demonstrated the combined use of archived FIA and Landsat sensor data to produce past forest species distribution and abundance maps of northern Minnesota to support forest ecosystem studies, including spruce budworm outbreak dynamics (Robert et al. 2018). Our findings suggest that calibrations between FIA ground data and Landsat sensor data are influenced by three salient factors, although the relative contribution of each was not specifically quantified in this study. Factors include variation in FIA subplot sampling protocols (Figure 1), variability and heterogeneity in species abundance captured by FIA sampling across three adjacent Landsat orbital paths (Figure 7–9), and the ability of the Landsat sensor to detect species that are either partially hidden from view (e.g., balsam fir) or that lack some phenology-based, multi-spectral/multi-temporal advantage for detection. It is important to note that the last factor is not a limitation related to FIA, but rather limitations of optical, polar-orbiting sensors such as Landsat (e.g., repeat cycle, cloud issues, and canopy penetration).

Evidence from this study clearly shows, based on the nearly identical study conducted by Wolter et al. (2008) where Landsat-scale, custom ground plot sampling was used, that neither past nor current FIA ground sampling strategies are optimal for use with 30-meter Landsat sensor data in this region of the country. However, the more numerous and tighter spatial arrangement of subplots under the older periodic FIA sampling design offers some benefits over the current FIA design, especially in heterogeneous forest conditions. Hence, given our national commitment to the future of the Landsat mission (Irons, Dwyer, and Barsi 2012), we recommend that the FIA sampling protocol should eventually be augmented to include measurements of forest structure that are more easily integrated with 30-meter Landsat sensor data, as the majority of US forests area not conveniently homogeneous. In tandem with such FIA augmentation, we recommend the fusion of multi-temporal Landsat sensor data and L-band SAR sensor data as a solution to afford more complete detection of forest structure, as past studies have yielded encouraging results in this region (Wolter and Townsend 2011).

Acknowledgements

We would like to thank the U.S. Forest Service's Forest Inventory and Analysis program granting access to these data via memorandum of understanding (3-MU-11242305-004). We would like to extend our sincere gratitude to Mark D. Nelson (Northern Research Station, USDA Forest Service, Minnesota) for valuable information and suggestions regrading use of FIA data and improvement of this manuscript. Last, special thanks to the anonymous reviewers that served to improve the quality of this manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was funded by the U.S. Forest Service (grant 15-JV-11242313-089), a USDA McIntire Stennis Capacity Grant (project IOW5348), and Iowa State University's Department of Natural Resource Ecology and Management.

References

Ahlgren, C. E. 1957. "Phenological Observations of Nineteen Native Tree Species in Northeastern Minnesota." *Ecology* 38 (4): 622–628. doi:10.2307/1943128.

Allen, D. M. 1971. *The Prediction Sum of Squares as a Criterion for Selecting Predictor Variables*. University of Kentucky.

Baker, W. L. 1989. "Landscape Ecology and Nature Reserve Design in the Boundary Waters Canoe Area, Minnesota." *Ecology* 70 (1): 23–35. doi:10.2307/1938409.

Baker, W. L. 1992. "Effects of Settlement and Fire Suppression on Landscape Structure." *Ecology* 73 (5): 1879–1887. doi:10.2307/1940039.

Bechtold, W. A., and P. L. Patterson. 2005. "The Enhanced Forest Inventory and Analysis Program-national Sampling Design and Estimation Procedures." *Gen. Tech. Rep*, SRS-80. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station, 85.

Blackard, J. A., M. V. Finco, E. H. Helmer, G. R. Holden, M. L. Hoppus, D. M. Jacobs, A. J. Lister et al. 2008. "Mapping US Forest Biomass Using Nationwide Forest Inventory Data and Moderate Resolution Information." *Remote Sensing of Environment* 112 (4): 1658–1677. DOI:10.1016/j.rse.2007.08.021.

Blais, J. R. 1983. "Trends in the Frequency, Extent, and Severity of Spruce Budworm Outbreaks in Eastern Canada." *Canadian Journal of Forest Research* 13 (4): 539–547. doi:10.1139/x83-079.

Blomberg, E., L. Ferro-Famil, M. J. Soja, L. M. H. Ulander, and S. Tebaldini. 2018. "Forest Biomass Retrieval from L-band Sar Using Tomographic Ground Backscatter Removal." *IEEE Geoscience and Remote Sensing Letters* 15 (7): 1030–1034. doi:10.1109/LGRS.2018.2819884.

Brown, P. M., and B. Cook. 2006. "Early Settlement Forest Structure in Black Hills Ponderosa Pine Forests." *Forest Ecology and Management* 223 (1–3): 284–290. doi:10.1016/j.foreco.2005.11.008.

Cappuccino, N., D. Lavertu, Y. Bergeron, and R. Jacques. 1998. "Spruce Budworm Impact, Abundance and Parasitism Rate in a Patchy Landscape." *Oecologia* 114 (2): 236–242. doi:10.1007/s004420050441.

Cohen, W. B., and T. A. Spies. 1992. "Estimating Structural Attributes of Douglas-fir/western Hemlock Forest Stands from Landsat and SPOT Imagery." *Remote Sensing of Environment* 41 (1): 1–17. doi:10.1016/0034-4257(92)90056-P.

Coppin, P. R., and M. E. Bauer. 1996. "Digital Change Detection in Forest Ecosystems with Remote Sensing Imagery." *Remote Sensing Reviews* 13 (3–4): 207–234. doi:10.1080/02757259609532305.

Crocker, S. J., G. C. Liknes, F. R. McKee, J. S. Albers, and B. H. Aukema. 2016. "Stand-level Factors Associated with Resurging Mortality from Eastern Larch Beetle (Dendroctonus Simplex LeConte)." *Forest Ecology and Management* 375: 27–34. doi:10.1016/j.foreco.2016.05.016.

Dale, V. H., L. A. Joyce, S. McNulty, R. P. Neilson, M. P. Ayres, M. D. Flannigan, P. J. Hanson et al. 2001. "Climate Change and Forest Disturbances." *BioScience* 51 (9): 723–734. DOI:10.1641/0006-3568-(2001)051[0723:CCAFD]2.0.CO;2.

Deo, R. K., M. B. Russell, G. M. Domke, C. W. Woodall, M. J. Falkowski, and W. B. Cohen. 2017. "Using Landsat Time-series and LiDAR to Inform Aboveground Forest Biomass Baselines in Northern Minnesota, USA." *Canadian Journal of Remote Sensing* 43 (1): 28–47. doi:10.1080/07038992.2017.1259556.

Dimitrov, D. D., J. S. Bhatti, and R. F. Grant. 2014. "The Transition Zones (Ecotone) between Boreal Forests and Peatlands: Ecological Controls on Ecosystem Productivity along a Transition Zone between Upland Black Spruce Forest and a Poor Forested Fen in Central Saskatchewan." *Ecological Modelling* 291: 96–108. doi:10.1016/j.ecolmodel.2014.07.020.

Dymond, C. C., E. T. Neilson, G. Stinson, K. Porter, D. A. MacLean, D. R. Gray, M. Campagna, and W. A. Kurz. 2010. "Future Spruce Budworm Outbreak May Create a Carbon Source in Eastern Canadian Forests." *Ecosystems* 13 (6): 917–931. doi:10.1007/s10021-010-9364-z.

Eder, J. J. 1989. "Don't Shoot unless It's Autumn." *Journal of Forestry* 87 (6): 50–51.

Engelstad, P. S., M. Falkowski, P. Wolter, A. Poznanovic, and P. Johnson. 2019. "Estimating Canopy Fuel Attributes from Low-Density LiDAR." *Fire* 2 (3): 38. doi:10.3390/fire2030038.

Franco-Lopez, H., A. R. Ek, and M. E. Bauer. 2001. "Estimation and Mapping of Forest Stand Density, Volume, and Cover Type Using the K-nearest Neighbors' Method." *Remote Sensing of Environment* 77 (3): 251–274. doi:10.1016/S0034-4257(01)00209-7.

Franklin, J. 1986. "Thematic Mapper Analysis of Coniferous Forest Structure and Composition." *International Journal of Remote Sensing* 7 (10): 1287–1301. doi:10.1080/01431168608948931.

Franklin, J. F., T. A. Spies, R. Van Pelt, A. B. Carey, D. A. Thornburgh, D. R. Berg, D. B. Lindenmayer et al. 2002. "Disturbances and Structural Development of Natural Forest Ecosystems with Silvicultural Implications, Using Douglas-fir Forests as an Example." *Forest Ecology and Management* 155 (1–3): 399–423. DOI:10.1016/S0378-1127(01)00575-8.

Frelich, L. E., and P. B. Reich. 1995. "Spatial Patterns and Succession in a Minnesota Southern-Boreal Forest." *Ecological Monographs* 65 (3): 325–346. doi:10.2307/2937063.

Geladi, P., and B. R. Kowalski. 1986. "Partial Least-squares Regression: A Tutorial." *Analytica Chimica Acta* 185: 1–17. doi:10.1016/0003-2670(86)80028-9.

Goeking, S. A. 2015. "Disentangling Forest Change from Forest Inventory Change: A Case Study from the US Interior West." *Journal of Forestry* 113 (5): 475–483. doi:10.5849/jof.14-088.

Goerndt, M. E., B. T. Wilson, and F. X. Aguilar. 2019. "Comparison of Small Area Estimation Methods Applied to Biopower Feedstock Supply in the Northern US Region." *Biomass and Bioenergy* 121: 64–77. doi:10.1016/j.biombioe.2018.12.008.

Gong, G. 1986. "Cross-Validation, the Jackknife, and the Bootstrap: Excess Error Estimation in Forward Logistic Regression." *Journal of the American Statistical Association* 81 (393): 108–113. doi:10.1080/01621459.1986.10478245.

Hall, F. G., D. B. Botkin, D. E. Strebel, K. D. Woods, and S. J. Goetz. 1991. "Large-scale Patterns of Forest Succession as Determined by Remote Sensing." *Ecology* 72 (2): 628–640. doi:10.2307/2937203.

Hanley, J. A., and B. J. McNeil. 1982. "The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve." *Radiology* 143 (1): 29–36. doi:10.1148/radiology.143.1.7063747.

Hansen, M. H., T. Matsumoto, R. Nakamura, R. Suzuki, T. Shimizu, and Y. Yanagihara 1990. "A Comprehensive Sampling System for Forest Inventory Based on an Individual Tree Growth Model, St. Paul, MN." Ph.D diss., University of Minnesota, College of Natural Resources.

Hansen, M. J., S. E. Franklin, C. Woudsma, and M. Peterson. 2001. "Forest Structure Classification in the North Columbia Mountains Using the Landsat TM Tasseled Cap Wetness Component." *Canadian Journal of Remote Sensing* 27 (1): 20–32. doi:10.1080/07038992.2001.10854916.

Hardisky, M. A., V. Klemas, and R.M. Smart. 1983. The influence of soil salinity, growth form, and leaf moisture on the spectral radiance of Spartina alterniflora canopies. Photogrammetric Engineering and Remote Sensing, 49: 77–83

Haynes, K. J., A. M. Liebhold, and D. M. Johnson. 2009. "Spatial Analysis of Harmonic Oscillation of Gypsy Moth Outbreak Intensity." *Oecologia* 159 (2): 249–256. doi:10.1007/s00442-008-1207-7.

Healey, S. P., W. B. Cohen, Y. Zhiqiang, and O. N. Krankina. 2005. "Comparison Of Tasseled Cap-based Landsat Data Structures for Use in Forest Disturbance Detection." *Remote Sensing Of Environment* 97 (3): 301-310. .

Heinselman, M. L. 1973. "Fire in The Virgin Forests Of The Boundary Waters Canoe Area, Minnesota." *Quaternary Research* 3 (3): 329–382. doi:10.1016/0033-5894(73)90003-3.

Hopkins, P. F., A. L. Maclean, and T. M. Lillesand. 1988. "Assessment of Thematic Mapper Imagery for Forestry Applications under Lake States Conditions." *Photogrammetric Engineering and Remote Sensing* 54 (1): 61–68.

Hoppus, M., and A. Lister. 2007. "The Status of Accurately Locating Forest Inventory and Analysis Plots Using the Global Positioning System." In: R. E. McRoberts, G. A. Reams, P. C. Van Deusen, and W. H. McWilliams, eds. Proceedings of the seventh annual forest inventory and analysis symposium; October 3-6, 2005; Portland, ME. Gen. Tech. Rep. WO-77. Washington, DC: US Department of Agriculture, Forest Service: 179–184. (Vol. 77).

Irons, J. R., J. L. Dwyer, and J. A. Barsi. 2012. "The Next Landsat Satellite: The Landsat Data Continuity Mission." *Remote Sensing Environment* 122: 11–21. doi:10.1016/j.rse.2011.08.026.

Ishii, H. T., S. I. Tanabe, and T. Hiura. 2004. "Exploring the Relationships among Canopy Structure, Stand Productivity, and Biodiversity of Temperate Forest Ecosystems." *Forest Science* 50 (3): 342–355.

Jactel, H., and E. G. Brockerhoff. 2007. "Tree Diversity Reduces Herbivory by Forest Insects." *Ecology Letters* 10 (9): 835–848. doi:10.1111/j.1461-0248.2007.01073.x.

James, P. M., B. R. Sturtevant, P. Townsend, P. Wolter, and M.-J. Fortin. 2011. "Two-dimensional Wavelet Analysis of Spruce Budworm Host Basal Area in the Border Lakes Landscape." *Ecological Applications* 21 (6): 2197–2209. doi:10.1890/09-1876.1.

Jin, S., and S. A. Sader. 2005. "MODIS Time-series Imagery for Forest Disturbance Detection and Quantification of Patch Size Effects." *Remote Sensing of Environment* 99 (4): 462–470. doi:10.1016/j.rse.2005.09.017.

Key, C. H., and N. C. Benson. 1999. The Normalized Burn Ratio (NBR): a Landsat TM radiometric index of burn severity. United States Geological Survey, Northern Rocky Mountain Science Center, Bozeman, MT.

Koricheva, J., M. J. Klapwijk, and C. Björkman. 2012. "Life History Traits and Host Plant Use in Defoliators and Bark Beetles: Implications for Population Dynamics." In *Insect Outbreaks Revisited*, 175–196. Wiley-Blackwell

Liu, W., C. Song, T. A. Schroeder, W. B. Cohen, and W. B. Schroeder. 2008. "Predicting Forest Successional Stages Using Multitemporal Landsat Imagery with Forest Inventory and Analysis Data." *International Journal of Remote Sensing* 29 (13): 3855–3872. doi:10.1080/01431160701840166.

Masek, J. G., E. F. Vermote, N. E. Saleous, F. G. Robert Wolfe, K. Hall, F. Huemmrich, F. Gao, J. Kutler, and T.-K. Lim. 2006. "A Landsat Surface Reflectance Dataset for North America, 1990-2000." *IEEE Geoscience and Remote Sensing Letters* 3 (1): 68–72. doi:10.1109/LGRS.2005.857030.

McRoberts, R. E. 2009. "A Two-step Nearest Neighbors Algorithm Using Satellite Imagery for Predicting Forest Structure within Species Composition Classes." *Remote Sensing of Environment* 113 (3): 532–545. doi:10.1016/j.rse.2008.10.001.

McRoberts, R. E. 2010. "The Effects of Rectification and Global Positioning System Errors on Satellite Image-based Estimates of Forest Area." *Remote Sensing of Environment* 114 (8): 1710–1717. doi:10.1016/j.rse.2010.03.001.

McRoberts, R. E., E. Tomppo, K. Schadauer, C. Vidal, G. Ståhl, G. Chirici, A. Lanz, E. Cienciala, S. Winter, and W. Brad Smith. 2009. "Harmonizing National Forest Inventories." *Journal of Forestry* 107 (4): 179–187.

McRoberts, R. E., E. O. Tomppo, A. O. Finley, and J. Heikkinen. 2007. "Estimating Areal Means and Variances of Forest Attributes Using the K-nearest Neighbors' Technique and Satellite Imagery." *Remote Sensing of Environment* 111 (4): 466–480. doi:10.1016/j.rse.2007.04.002.

McRoberts, R. E., E. O. Tomppo, and E. Næsset. 2010. "Advances and Emerging Issues in National Forest Inventories." *Scandinavian Journal of Forest Research* 25 (4): 368–381. doi:10.1080/02827581.2010.496739.

McRoberts, R. E., G. R. Holden, M. D. Nelson, G. C. Liknes, and D. D. Gormanson. 2005. "Using Satellite Imagery as Ancillary Data for Increasing the Precision of Estimates for the Forest Inventory and Analysis Program of the USDA Forest Service." *Canadian Journal of Forest Research* 35 (12): 2968–2980. doi:10.1139/x05-222.

McRoberts, R. E., M. D. Nelson, and D. G. Wendt. 2002. "Stratified Estimation of Forest Area Using Satellite Imagery, Inventory Data, and the k-Nearest Neighbors Technique." *Remote Sensing of Environment* 82 (2–3): 457–468. doi:10.1016/S0034-4257(02)00064-0.

Miles, P., C. M. Chen, and E. C. Leatherberry. 1995. "Minnesota Forest Statistics, 1990, Revised." Resource Bulletin NC-158. St. Paul, MN: U.S. Dept. of Agriculture, Forest Service, North Central Forest Experiment Station.

Miles, P., G. J. Keith Jacobson, E. J. Brand, D. Meneguzzo, M. E. Mielke, C. Olson. 2007. "Minnesota's Forests 1999-2003 (Part A)."*Resour. Bull.* NRS-12A. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station. 92.

Moeur, M., and A. R. Stage. 1995. "Most Similar Neighbor: An Improved Sampling Inference Procedure for Natural Resource Planning." *Forest Science* 41 (2): 337–359.

Moore, M. M., and M. E. Bauer. 1990. "Classification of Forest Vegetation in North-central Minnesota Using Landsat Multispectral Scanner and Thematic Mapper Data." *Forest Science* 36 (2): 330–342.

Nelson, M. D., S. Healey, W. Keith Moser, J. G. Masek, and W. Cohen. 2011. "Consistency of Forest Presence and Biomass Predictions Modeled across Overlapping Spatial and Temporal Extents." *Mathematical and Computational Forestry and Natural-Resource Sciences* 3 (2): 102–113.

Nelson, M. D., S. P. Healey, W. Keith Moser, and M. H. Hansen. 2009. "Combining Satellite Imagery with Forest Inventory Data to Assess Damage Severity following a Major Blowdown Event in Northern Minnesota, USA." *International Journal of Remote Sensing* 30 (19): 5089–5108. doi:10.1080/01431160903022951.

Ohmann, J. L., and M. J. Gregory. 2002. "Predictive Mapping of Forest Composition and Structure with Direct Gradient Analysis and Nearest-neighbor Imputation in Coastal Oregon, USA." *Canadian Journal of Forest Research* 32 (4): 725–741. doi:10.1139/x02-011.

Ohmann, J. L., M. J. Gregory, and H. M. Roberts. 2014. "Scale Considerations for Integrating Forest Inventory Plot Data and Satellite Image Data for Regional Forest Mapping." *Remote Sensing of Environment* 151: 3–15. doi:10.1016/j.rse.2013.08.048.

Pastor, J., A. Sharp, and P. Wolter. 2005. "An Application of Markov Models to the Dynamics of Minnesota's Forests." *Canadian Journal of Forest Research* 35 (12): 3011–3019. doi:10.1139/x05-198.

Raffa, K. F., E. N. Powell, and P. A. Townsend. 2013. "Temperature-driven Range Expansion of an Irruptive Insect Heightened by Weakly Coevolved Plant Defenses." *Proceedings of the National Academy of Sciences* 110 (6): 2193–2198. doi:10.1073/pnas.1216666110.

Reineke, L. H. 1933. "Perfecting a Stand-density Index for Even-aged ForestS." Journal of Agricultural Research, 46 (10), 627–638.

Riemann, R., B. T. Wilson, A. Lister, and S. Parks. 2010. "An Effective Assessment Protocol for Continuous Geospatial Datasets of Forest Characteristics Using USFS Forest Inventory and Analysis (FIA) Data." *Remote Sensing of Environment* 114 (10): 2337–2352. doi:10.1016/j.rse.2010.05.010.

Riemann,R, Wilson, B. T., Lister, A., and Parks, S. 2010. "An Effective Assessment Protocol for Continuous Geospatial Datasets of Forest Characteristics Using USFS Forest Inventory and Analysis (FIA) Data." *Remote Sensing of Environment* 114 (10): 2337–2352.

Robert, L. E., D. Kneeshaw, and B. R. Sturtevant. 2012. "Effects of Forest Management Legacies on Spruce Budworm (*Choristoneura Fumiferana*) Outbreaks." *Canadian Journal of Forest Research* 42 (3): 463–475. doi:10.1139/x2012-005.

Robert, L.-E., B. R. Sturtevant, B. J. Cooke, P. M. A. James, M.-J. Fortin, P. A. Townsend, P. T. Wolter, and D. Kneeshaw. 2018. "Landscape Effects of Forest Legacies on Spruce Budworm Outbreaks: Unintended Impacts of Large-scale Forest Management." *Ecography* 41: 1556–1571. doi:10.1111/ecog.03553.

Rock, B. N., J. E. Vogelmann, D. L. Williams, A. F. Vogelmann, and T. Hoshizaki. 1986. Remote Detection of Forest Damage: Plant responses to stress may have spectral "signatures" that could be used to map, monitor, and measure forest damage. Bioscience 36 (7): 439-445.

Roland, J. 2005. "Are The" Seeds" of Spatial Variation in Cyclic Dynamics Apparent in Spatially-replicated Short Time-series? an Example from the Forest Tent Caterpillar." In *Annales Zoologici Fennici*, 397–407). Finnish Zoological and Botanical Publishing Board.

Rouse Jr, J.W., R.H. Haas, J.A. Schell, and D.W. Deering. 1974. Monitoring vegetation systems in the Great Plains with ERTS. Proceedings of the Third Earth Resources Technology Satellite-1 Symposium. Greenbelt, MD: NASA: 301-317.

Ruefenacht, B., M. V. Finco, M. D. Nelson, R. Czaplewski, E. H. Helmer, J. A. Blackard, G. R. Holden et al. 2008. "Conterminous US and Alaska Forest Type Mapping Using Forest Inventory and Analysis Data." *Photogrammetric Engineering and Remote Sensing* 74 (11): 1379–1388. DOI:10.14358/PERS.74.11.1379.

Salajanu, D., and C. E. Olson. 2001. "The Significance of Spatial Resolution: Identifying Forest Cover from Satellite Data." *Journal of Forestry* 99 (6): 32–38.

Scheller, R. M., D. J. Mladenoff, T. R. Crow, and T. A. Sickley. 2005. "Simulating the Effects of Fire Reintroduction versus Continued Fire Absence on Forest Composition and Landscape Structure in the Boundary Waters Canoe Area, Northern Minnesota, USA." *Ecosystems* 8 (4): 396–411. doi:10.1007/s10021-003-0087-2.

Short, N. M. 1982. *The Landsat Tutorial Workbook, NASA RP-1078*, 35. Washington, D.C.: Government Printing Office.

Sommerfeld, A., C. Senf, B. Buma, A. W. D'Amato, T. Després, I. Díaz-Hormazábal, S. Fraver et al. 2018. "Patterns and Drivers of Recent Disturbances across the Temperate Forest Biome." *Nature Communications* 9 (1): 4355. DOI:10.1038/s41467-018-06788-9.

Song, C., T. A. Schroeder, and W. B. Cohen. 2007. "Predicting Temperate Conifer Forest Successional Stage Distributions with Multitemporal Landsat Thematic Mapper Imagery." *Remote Sensing of Environment* 106 (2): 228–237. doi:10.1016/j.rse.2006.08.008.

Sturtevant, B. R., B. J. Cooke, D. D. Kneeshaw, and D. A. MacLean. 2015. "Modeling Insect Disturbance across Forested Landscapes: Insights from the Spruce Budworm." In *Simulation Modeling of Forest Landscape Disturbances*, 93–134. Cham: Springer.

Sturtevant, B. R., B. R. Miranda, P. T. Wolter, P. M. A. James, M.-J. Fortin, and P. A. Townsend. 2014. "Forest Recovery Patterns in Response to Divergent Disturbance Regimes in the Border Lakes Region of Minnesota (USA) and Ontario (Canada)." *Forest Ecology and Management* 313: 199–211. doi:10.1016/j.foreco.2013.10.039.

Sturtevant, B. R., E. J. Gustafson, L. Wei, and H. S. He. 2004. "Modeling Biological Disturbances in LANDIS: A Module Description and Demonstration Using Spruce Budworm." *Ecological Modelling* 180 (1): 153–174. doi:10.1016/j.ecolmodel.2004.01.021.

Su, Q., T. D. Needham, and D. A. MacLean. 1996. "The Influence of Hardwood Content on Balsam Fir Defoliation by Spruce Budworm." *Canadian Journal of Forest Research* 26 (9): 1620–1628. doi:10.1139/x26-182.

Thomas, N. E., S. N. Chengquan Huang, S. P. Goward, K. Rishmawi, K. Schleeweis, and A. Hinds. 2011. "Validation of North American Forest Disturbance Dynamics Derived from Landsat Time Series Stacks." *Remote Sensing of Environment* 115 (1): 19–32. doi:10.1016/j.rse.2010.07.009.

Tucker, C. J. 1978. "A Comparison of Satellite Sensor Bands for Vegetation Monitoring." *Photogrammetric Engineering and Remote Sensing* 44 (Nov): 1369–1380.

Turner, D. P., W. B. Cohen, R. E. Kennedy, K. S. Fassnacht, and J. M. Briggs. 1999. "Relationships between Leaf Area Index and Landsat TM Spectral Vegetation Indices across Three Temperate Zone Sites." *Remote Sensing of Environment* 70 (1): 52–68. doi:10.1016/S0034-4257(99)00057-7.

Turner, M. G. 2010. "Disturbance and Landscape Dynamics in a Changing World1." *Ecology* 91 (10): 2833–2849. doi:10.1890/10-0097.1.

Vandendriesche, D., and L. Haugen. 2008. "Comparison of FVS Projection of Oak Decline on the Mark Twain National Forest to Actual Growth and Mortality as Measured over Three FIA Inventory Cycles." In: R. N. Havis and N. L. Crookston., edscomps. 2008. *Third Forest Vegetation Simulator Conference*; 2007 February 13–15;Fort Collins, CO. Proceedings RMRS-P-54. Fort Collins, CO: US Department of Agriculture, Forest Service, Rocky Mountain Research Station. p. 68–80 (Vol. 54).

Vermote, E. F., N. El Saleous, C. O. Justice, Y. J. Kaufman, J. L. Privette, L. Remer, J. C. Roger, and D. Tanre. 1997. "Atmospheric Correction of Visible to Middle-infrared EOS-MODIS Data over Land Surfaces: Background, Operational Algorithm and Validation. Journal of Geophysical Research." *Atmospheres* 102 (D14): 17131–17141. doi:10.1029/97JD00201.

Vogelmann, J. E., S. M. Howard, L. Yang, C. R. Larson, B. K. Wylie, and N. V. Driel. 2001. "Completion of the 1990s National Land Cover Data Set for the Conterminous United States from Landsat Thematic Mapper Data and Ancillary Data Sources." *Photogrammetric Engineering and Remote Sensing* 67: 6.

Volney, W. J. A., and D. G. McCullough. 1994. "Jack Pine Budworm Population Behaviour in Northwestern Wisconsin." *Canadian Journal of Forest Research* 24 (3): 502–510. doi:10.1139/x94-067.

Wesołowski, T., and R. Patryk. 2006. "Tree Defoliation by Winter Moth Operophtera Brumata L. During an Outbreak Affected by Structure of Forest Landscape." *Forest Ecology and Management* 221 (1–3): 299–305. doi:10.1016/j.foreco.2005.10.023.

Wilson, B. T., A. J. Lister, and R. I. Riemann. 2012. "A Nearest-neighbor Imputation Approach to Mapping Tree Species over Large Areas Using Forest Inventory Plots and Moderate Resolution Raster Data." *Forest Ecology and Management* 271: 182–198. doi:10.1016/j.foreco.2012.02.002.

Wilson, B. T., C. W. Woodall, and D. M. Griffith. 2013. "Imputing Forest Carbon Stock Estimates from Inventory Plots to a Nationally Continuous Coverage." *Carbon Balance and Management* 8 (1): 1. doi:10.1186/1750-0680-8-1.

Wilson, B. T., J. F. Knight, and R. E. McRoberts. 2018. "Harmonic Regression of Landsat Time Series for Modeling Attributes from National Forest Inventory Data." *ISPRS Journal of Photogrammetry and Remote Sensing* 137: 29–46. doi:10.1016/j.isprsjprs.2018.01.006.

Wilson, E. H., and S. A. Sader. 2002. "Detection of Forest Harvest Type Using Multiple Dates of Landsat TM Imagery." *Remote Sensing of Environment* 80 (3): 385–396. doi:10.1016/S0034-4257(01)00318-2.

Wold, S. 1994. "PLS for Multivariate Linear Modeling QSAR: In *Chemometric Methods in Molecular Design*." *Methods and Principles in Medicinal Chemistry Van De Waterbeemd H (Editor) Verlag-Chemie* 2008: 195–218.

Wolter, P. T., B. R. Sturtevant, B. R. Miranda, S. M. Lietz, P. A. Townsend, and J. Pastor. 2012b. *Greater Border Lakes Region Land Cover Classification and Change Detection*. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station. doi:10.2737/RDS-2012-0007.

Wolter, P. T., C. A. Johnston, and G. J. Niemi. 2006. "Land Use Land Cover Change in the US Great Lakes Basin 1992 to 2001." *Journal of Great Lakes Research* 32 (3): 607–628. doi:10.3394/0380-1330(2006)32[607:LULCCI]2.0.CO;2.

Wolter, P. T., D. J. Mladenoff, G. E. Host, and T. R. Crow. 1995. "Improved Forest Classification in the Northern Lake States Using Multi-temporal Landsat Imagery." *Photogrammetric Engineering and Remote Sensing* 61 (9): 1129–1144.

Wolter, P. T., E. A. Berkley, S. D. Peckham, A. Singh, and P. A. Townsend. 2012. "Exploiting Tree Shadows on Snow for Estimating Forest Basal Area Using Landsat Data." *Remote Sensing Environment* 121: 69–79. doi:10.1016/j.rse.2012.01.008.

Wolter, P. T., and M. A. White. 2002. "Recent Forest Cover Type Transitions and Landscape Structural Changes in Northeast Minnesota, USA." *Landscape Ecology* 17 (2): 133–155. doi:10.1023/A:1016522509857.

Wolter, P. T., and P. A. Townsend. 2011. "Multi-sensor Data Fusion for Estimating Forest Species Composition and Abundance in Northern Minnesota." *Remote Sensing of Environment* 115 (2): 671–691. doi:10.1016/j.rse.2010.10.010.

Wolter, P. T., P. A. Townsend, and B. R. Sturtevant. 2009. "Estimation of Forest Structural Parameters Using 5 and 10 Meter SPOT-5 Satellite Data." *Remote Sensing of Environment* 113 (9): 2019–2036. doi:10.1016/j.rse.2009.05.009.

Wolter, P. T., P. A. Townsend, B. R. Sturtevant, and C. C. Kingdon. 2008. "Remote Sensing of the Distribution and Abundance of Host Species for Spruce Budworm in Northern Minnesota and Ontario." *Remote Sensing of Environment* 112 (10): 3971–3982. doi:10.1016/j.rse.2008.07.005.

Youden, W. J. 1950. "Index for Rating Diagnostic Tests." *Cancer* 3 (1): 32–35. doi:10.1002/(ISSN)1097-0142.

Zhu, Z., and D. L. Evans. 1994. "US Forest Types and Predicted Percent Forest Cover from AVHRR Data." *Photogrammetric Engineering and Remote Sensing* 60: 5.