Remote Sensing of Environment 115 (2011) 671–691

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Multi-sensor data fusion for estimating forest species composition and abundance in northern Minnesota

Peter T. Wolter *, Philip A. Townsend [1]

Department of Forest and Wildlife Ecology, University of Wisconsin—Madison, 1630 Linden Drive (Suite 226), Madison, WI 53706 USA

ARTICLE INFO

Article history:
Received 4 August 2010
Received in revised form 27 October 2010
Accepted 28 October 2010

Keywords:
Forest composition
Forest structure
Data fusion
Multi-sensor
SAR
Radarsat
PALSAR
SPOT
Landsat
Basal area
PLS regression
Minnesota
Great Lakes

ABSTRACT

The magnitude, duration, and frequency of forest disturbance caused by the spruce budworm and forest tent caterpillar in northern Minnesota and neighboring Ontario, Canada have increased over the last century due to a shift in forest species composition linked to historical fire suppression, forest management, and pesticide application that has fostered increased dominance of host tree species. Modeling approaches are currently being used to understand and forecast potential management effects in changing insect disturbance trends. However, detailed forest composition data needed for these efforts is often lacking. We used partial least squares (PLS) regression to integrate different combinations of satellite sensor data including Landsat, Radarsat-1, and PALSAR, as well as pixel-wise forest structure information derived from SPOT-5 sensor data (Wolter et al., 2009), to determine the best combination of sensor data for estimating near species-level proportional forest composition (12 types: 10 species and 2 genera). Single-sensor and various multi-sensor PLS models showed distinct species-dependent sensitivities to relative basal area (BA), with Landsat variables showing greatest overall sensitivity. However, best results were achieved using a combination of data from all these sensors, with several C-band (Radarsat-1) and L-band (PALSAR) variables showing sensitivity to the composition and abundance of specific species. Pixel-level forest structure estimates derived from SPOT-5 data were generally more sensitive to conifer species abundance (especially white pine) than to hardwood species composition. Relative BA models accounted for 68% (jack pine) to 98% (maple spp.) of the variation in ground data with RMSE values between 2.46% and 5.65% relative BA, respectively. Receiver operating characteristic (ROC) curves were used to determine the effective lower limits of usefulness of species relative BA estimates which ranged from 5.94% (jack pine) to 39.41% (black ash). These estimates were then used to produce a dominant forest species map for the study region with an overall accuracy of 78%. Most notably, this approach facilitated discrimination of aspen from paper birch as well as spruce and fir from other conifer species which is crucial for the study of forest tent caterpillar and spruce budworm dynamics in the Upper Midwest. We also demonstrate that PLS regression is an effective data fusion strategy for mapping composition of heterogeneous forests using satellite sensor data.

1. Introduction

Spatially explicit information about forest tree species composition, stand structure, landscape configuration, and changes in each of these over time are essential for understanding and forecasting ecosystem function (Pastor & Post, 1986; Aber et al., 1995; Sturtevant & Gustafson, 2004; Bonan, 2008). This information may be used to track the effects of climate change on biological diversity (Peters & Darling, 1985; Betts et al., 1997; He et al., 2002; Percy et al., 2002), improve our knowledge of fauna habitat preference and suitability (MacArthur & MacArthur, 1961; Pastor et al., 1999; Hanowski et al., 2005; Hyde et al., 2006; Moen et al., 2008), or assess the impact of past and present forest management on tree species distribution and abundance (Wolter & White, 2002; Friedman & Reich, 2005). Data on host tree species also provides insight into the patterns, cycles, and severity of insect disturbance, as well as the likely consequences of outbreaks on a regional scale (Blais, 1983; Batzer et al., 1995; Radeloff et al., 2000; Scheller et al., 2005; Bouchard et al., 2006; Ellenwood & Krist, 2007).

In recent years, the estimation of forest biophysical parameters using satellite data has become increasingly routine (Franco-Lopez et al., 2001; Cohen et al., 2003; Pulliainen et al., 2003; Rauste, 2005; Hyde et al., 2006; Chopping et al., 2008; McRoberts et al., 2007, and many more). However, improvements in the level of detail and precision among many estimated parameters continue to be a challenge (Couturier et al., 2009; Moisen et al., 2006; Wolter et al., 1995, 2008, 2009), especially when the aim is to augment or supplant traditional, ground-based forest inventory data (Holmgren & Thuresson, 1998; Meng et al., 2009; McRoberts, 2009).

* Corresponding author. Tel.: +1 608 263 9107(office); fax: +1 608 262 9922.
E-mail addresses: ptwolter@wisc.edu (P.T. Wolter), ptownsend@wisc.edu (P.A. Townsend).
[1] Tel.: +1 608 262 1669(office); fax: +1 608 262 9922.

0034-4257/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2010.10.010

Pixel-wise estimation of forest tree species composition using multi-spectral satellite sensor data (e.g., Landsat or SPOT) poses unique challenges in regions where species diversity and structural complexity are high (Wolter et al., 1995, 2008; Franco-Lopez et al., 2001; Reese et al., 2002; Leckie et al., 2005; McRoberts, 2008; Stabach et al., 2009; Bauer et al., 2009). For example, in the upper Midwest, especially in northern Minnesota, multi-spectral forest classifications have rarely achieved accuracies greater than 65–85% (Franco-Lopez et al., 2001). By and large, this is because it is difficult to adequately account for the full range of spectral variability of forests using ground-based training data (Holmgren & Thuresson, 1998; Franco-Lopez et al., 2001), due to differences in density, age, understory vegetation, soil, slope, etc. If spectrally confounding factors such as terrain position (Bolstad & Lillesand, 1992), tree density (Huang et al., 2001), tree size and crown closure (Wolter et al., 2009), understory vegetation (Shen et al., 1985) or other landscape characteristics (Vogelmann et al., 1998) are accounted for—or at least their affects are reduced in some way—accuracies among image-derived forest species composition and abundance estimates may be improved beyond past levels.

1.1. Study objective

The objective of this research was to investigate the combined use of optical (Landsat and SPOT-derived forest structure estimates) and synthetic aperture radar (SAR) satellite image data for estimating the distribution and abundance of four dominant deciduous hardwood types (two genera and two species: aspens, *Populus tremuloides* and *P. grandidentata* combined; maples, *Acer saccharum* and *A. rubrum* combined; paper birch, *Betula papyrifera*; and black ash, *Fraxinus nigra*) and eight coniferous tree species (red pine, *Pinus resinosa*; white pine, *P. strobes*; jack pine, *P. banksiana*; black spruce, *Picea mariana*; white spruce, *P. glauca*; balsam fir, *Abies balsamea*; eastern larch (tamarack), *Larix laricina*; and northern white cedar, *Thuja occidentalis*) in northern Minnesota to support ongoing efforts to understand insect disturbance dynamics in the upper Midwest. We extend the work of Wolter et al. (2008), where partial least squares (PLS) regression was applied to estimate and map the distribution and abundance of host tree species (balsam fir and spruces) for spruce budworm (*Choristoneura fumiferana*), to include all major forest dominants, including host species (aspens) for forest tent caterpillar (*Malacosoma disstria*). Tree species rarely found (*Quercus* spp.) or found only intermittently as sparse associates (e.g. *B. alleghaniensis, P. balsamifera*, *F. americana*, *F. pennsylvanica,* and *Prunus serotina*) with dominant species were not focal points of this mapping effort. The aim is to accurately discriminate host tree species from other optically similar non-host species to improve insect disturbance modeling.

A secondary objective was to determine whether C-band (5.30 GHz) and L-band (1.27 GHz) synthetic aperture radar (SAR) amplitude data can be used in concert with SPOT structure estimates, optical measures of forest phenology (Wolter et al., 1995), and other indices to distinguish aspens from paper birch (*Betula papyrifera*), as these two forest types are optically similar (Shen et al., 1985; Franco-Lopez et al., 2001). In the past, it was common practice to combine aspen and birch for Landsat image classifications (e.g., Moore & Bauer, 1990; Wolter & White, 2002), as efforts to separate these species have been unsuccessful (Franco-Lopez et al., 2001) or merely satisfactory when performed under ideal conditions (Wolter et al., 1995). Northeast Minnesota provides an excellent test landscape for this data fusion experiment because forests in this region are heterogeneous in terms of both spatial configuration and species composition (Heinselman, 1973).

1.2. Background

1.2.1. Forest phenology and satellite image data

The discrimination of forest species using optical imagery can be improved by the optimal timing of image acquisition with respect to phenological timing (Sayn-Wittgenstein, 1961, Wolter et al., 1995) (Table 1). Imagery from late spring is important to discriminate aspen (*Populus* spp.) from birch (*Betula* spp.) because leaves of the understory woody vegetation (e.g., *Corylus cornuta, Acer spicatum, or Cornus* spp.) are not fully developed at that time, and hence less likely to dampen spectral differences observed between overstory species (Shen et al., 1985; Badhwar et al., 1986). Species-level forest classifications with overall accuracies as high as 80% have been developed in Wisconsin (Wolter et al., 1995) and Minnesota (Wolter & White, 2002) using Landsat imagery that captures specific phenological events, with results outperforming efforts that did not use such information (Bolstad & Lillesand, 1992; Schriever & Congalton, 1995; Mickelson et al., 1998; Franco-Lopez et al., 2001).

1.2.2. Combined use of optical and synthetic aperture radar (SAR) imagery

SAR data have been used extensively to estimate forest structural variables (Sader, 1987; Ranson et al., 1997; Koskinen et al., 2001; Fransson et al., 2001; Santoro et al., 2002; Pulliainen et al., 2003; Engdahl et al., 2004) and to classify forest cover types (Lim et al., 1989; Ranson & Sun, 1994; Dobson et al., 1995; Touzi et al., 2004). However, fewer examples of research exist in which SAR data have been used alone or in combination with optical data to map or estimate forest composition (e.g., Leckie, 1990; Dobson et al., 1995; Townsend, 2002).

Theoretical SAR backscatter models have shown sensitivity to numerous vegetation attributes, including dielectric properties, structure, and vegetation orientation (Dobson et al., 1992a; Imhoff, 1995). Each forest type has a unique set of structural attributes (e.g., bole shape, crown shape, branch arrangement, and leaf orientation) that may produce unique backscatter at different radar wavelengths (Sader, 1987; Leckie, 1990; Ahern et al., 1993; Ranson et al., 1997; Imhoff, 1995). As a result, C-band (5.6 cm wavelength) and L-band (23.6 cm wavelength) SAR image data have been used with knowledge-based classification procedures to map broad forest categories characterized by differences in leaf type (needle-leaf or broadleaf) and canopy shape (excurrent or decurrent) with high accuracy (Dobson et al., 1995). Airborne SAR data were used in Canada to classify jack pine (*Pinus banksiana*), black spruce (*Picea mariana*), quaking aspen, and three classes of mixed forest with 90% accuracy, and 98% accuracy in the separation of pure jack pine from pure black spruce (Saatchi & Rignot, 1997). Use of both leaf-on and leaf-off C- and L-band SAR data has been suggested for improving forest classification results (Ranson et al., 1997), but has not yet been widely explored in heterogeneous forests.

It should be noted that detection of specific hardwood phenology events (e.g., leaf-flush/drop) using C-band SAR data is possible (Sharma et al., 2005). However, confounding variation in C-band backscatter due to unstable environmental conditions during spring and autumn limits reliable application (Ranson & Sun, 1994, 2000; Verbyla, 2001).

Table 1
Unique phenology windows that aid forest type determination using optical imagery.

Date	Species	Event
15—20 May	Quaking and bigtooth aspen *(P. tremuloides, P. grandidentata,)*	First hardwoods to flush leaves
1—5 June	Black ash *(Fraxinus nigra)*	Last hardwood to flush leaves
5—10 September	Black ash	First hardwood to shed leaves
20—25 September	Sugar and red maple *(Acer saccharum, A. rubrum)*	Unique autumn leaf color
15—20 October	Eastern larch *(Larix laricina)*	Unique autumn leaf color
Late February–March	Eastern larch	Leaves off, snow covered ground

By gathering SAR imagery during environmentally stable seasons (summer and winter), we aim to maximize the strengths of SAR imagery through fusion with optical sensor data (Pohl & Van Genderen, 1998; Nielsen, 2002; Treuhaft et al., 2004). The goal of image fusion is to employ different types of sensor data to obtain more information from their union than is possible from separate analyses. Nielsen (2002) described the use of canonical correlation analysis (CCA) and partial least squares (PLS) regression as effective means of achieving data fusion, since both are used to transform the original data into new orthogonal variables or latent components. Fusion of Landsat and Radarsat-1(C-band) data is purely a multi-sensor union of disparate data with no difference in spatial resolution (Pohl & Van Genderen, 1998). Thus, these sensors provide complementary information (Pohl & Van Genderen, 1998). On the other hand, the combination of Radarsat-1(30 m pixels, C-band) and Phased Array L-band Synthetic Aperture Radar (PALSAR) (12.5 m pixels, L-band) data represents a multi-resolution, multi-wavelength integration that may provide additional power for tree species discrimination. For example, white spruce and red pine or aspen and paper birch are spectrally similar using summer Landsat data, but are very different in terms of branch, foliage, and bole arrangement (e.g. multi-stemmed paper birch, Cooper, 1913; Buell & Niering, 1957). Therefore, these species differ considerably in radar reflectivity (Hallikainen et al., 1990; Ranson & Saatchi, 1992).

Both single polarizations (HH and VV) and cross polarizations (HV and VH) of L-band SAR data are sensitive to forest biomass and other forest parameters due to strong L-band interaction (double bounce) with tree boles (Richards et al., 1987; Durden et al., 1989; Wang et al., 1993, 1995; Dobson et al., 1995; Santoro et al., 2009) and large branches (Sader, 1987; Dobson et al., 1992b; Pulliainen et al., 1999). In addition, correlations between backscatter and forest basal area and above ground biomass at L-band HV and VV polarizations suggest that L-band SAR response is weakly dependent on species type (Sader, 1987; Watanabe et al., 2006). While data from the Japanese Earth Resources Satellite (JERS-1, 18 m pixels, L-band) have been combined with C-band SAR data to map forest composition (Kellndorfer et al., 1997), only a few published studies to date have used data from the new L-band PALSAR sensor (launched in January 2006) to map general land cover (Turkar & Rao, 2008), make assessments of hurricane damage (Watanabe et al., 2007), or to estimate tree height (Shimada, 2007). We demonstrate one of the first assessments of the use of PALSAR data to characterize forest composition in detail.

1.2.3. Landsat, SPOT, and partial least squares regression

PLS regression (Wold, 1966, 1975) is similar to the canonical variate substitution (CVS) fusion method described by Campbell (1993). In CVS, linear combinations of bands (factors) are generated to maximize the difference between input image variables with reference to variation in field training data (Pohl & Van Genderen, 1998). Similarly, PLS is a linear regression method that also forms components (factors or latent variables) as new independent variables (explanatory variables or predictors) in a regression model. However, the unique assumption with PLS is that response and predictor variable blocks are manifestations of the same set of underlying latent variables (Wold et al., 2009). As with CVS, components in PLS are determined by both the response and predictor variables (Wold et al., 2009). Consequently, a PLS regression model can be expected to have a smaller number of components without an appreciably smaller R^2 value.

Recently, PLS regression was successfully used with multi-seasonal Landsat sensor data to estimate the distribution and relative basal area (BA) of host tree species (white spruce, black spruce, and balsam fir) for spruce budworm within a 6.4 million hectare area covering the Border Lakes region of northern Minnesota, U.S.A. and northwestern Ontario, Canada (Wolter et al., 2008). Their analyses yielded estimates for overall forest BA (R^2 of 0.62, RMSE 4.67 $m^2\ ha^{-1}$ or $\leq$20% of measured BA), spruce relative BA (R^2 0.88, RMSE 12.57 $m^2\ ha^{-1}$), and balsam fir relative BA (R^2 0.64, RMSE 6.08 $m^2\ ha^{-1}$). The use of PLS regression was appropriate in this case as the technique does not assume zero error in the predictor/image variables, which many linear methods do (Curran & Hay, 1986); and PLS regression uniquely handles multi-collinearity among image predictors making it more

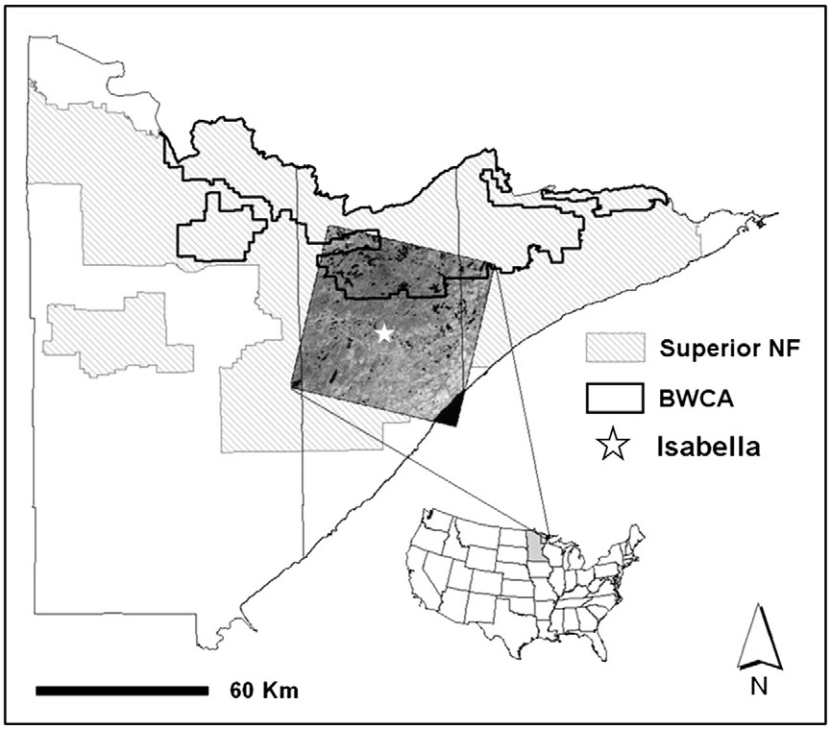

Fig. 1. The study area in northeast Minnesota which consists of a SPOT-5 image footprint (K587 and J253) within the Superior National Forest and part of the Boundary Waters Canoe Area (BWCA) wilderness.

robust than traditional multiple linear regression (Wolter et al., 2008). This work builds on Wolter et al. (2008) by including ten additional tree species to the predictive models.

Wolter et al. (2009) demonstrated use of PLS regression to estimate forest structural parameters in northeastern Minnesota using 5 and 10 m SPOT-5 data. We developed an automated variable selection procedure for use with PLS regression (Wolter et al., 2008) that produced models for pixel-wise estimation and mapping of mean values, respectively, for deciduous hardwood and coniferous crown diameter ($R^2 = 0.82$ and 0.93, RMSE 0.62 and 0.47 m), bole diameter ($R^2 = 0.82$ and 0.90, RMSE 2.92 and 3.75 cm), tree height ($R^2 = 0.69$ and 0.92, RMSE 1.27 and 1.59 m), crown closure ($R^2 = 0.52$ and 0.68, RMSE 5.49 and 6.02%), vertical length of live crown ($R^2 = 0.58$ and 0.81, RMSE 0.96 and 1.25 m), and basal area ($R^2 = 0.71$ and 0.74, RMSE 2.47 and 4.58 $m^2\ ha^{-1}$) for a 3600 km^2 area in northeast Minnesota (see Fig. 1). Predictive variables included neighborhood statistics (standard deviation, variance, semivariogram sill variance, and ratios of these metrics at 5 and 10 m) calculated from multi-spectral (10 m) and panchromatic (5 m) SPOT-5 sensor data and derivatives.

2. Methods

2.1. Study region

The study region consists of a single 60 × 60 km Systeme pour l'Observation de la Terre (SPOT-5) image footprint (K587, J253 centered at N 47.681° latitude and W 91.345° longitude) located within the Superior National Forest (Fig. 1). This region is composed of diverse tree species (five conifer genera and seven broadleaved, deciduous tree genera) and is considered transitional between the sub-boreal, Great Lakes–St. Lawrence forests and boreal forest (Heinselman, 1973; Baker, 1989). These forests have been intensively managed for wood fiber for over 100 years. Through time, forest management practices and fire suppression have altered forest composition resulting in the current dominance of aspen, paper birch, white spruce, and balsam fir (Heinselman, 1973; Reich et al., 2001; Wolter & White, 2002). However, the northern portion of the region is largely protected (Fig. 1) and has an extensive fire history that supports large stands of pioneer forest dominated by jack pine. In addition, the study area contains remnants of old-growth white pine (*Pinus strobus*) and red pine forests (Heinselman, 1973; Frelich & Reich, 1995).

2.2. Field data

Field data included plots sampled in 2003–2004 (n = 67) and 2006–2007 (n = 120) by the authors as well as additional plots sampled in 2006 by the Natural Resources Research Institute (NRRI) at the University of Minnesota—Duluth to support forest bird habitat modeling (Fig. 2). Field plot data sampled by the authors consisted of a total five variable-radius (Grosenbaugh, 1952) subplots: one located at the intersection and one at each of the four end points of two crossing 60 m transect lines placed near center of large (≥4.4 ha), homogenous stands (Fig. 2A). Sufficient stand size and homogeneity assured that stand edge effects were minimized during analysis, and that image misregistration errors were inconsequential. Basal area (BA) by species was measured at each subplot using a metric basal area factor (BAF) 2 prism. In addition, percent cover of forest trees by species was estimated visually in one of 10 cover classes following Peet et al. (1998) for the canopy and subcanopy. Total crown closure for the plot was estimated using 121 densitometer measurements made in one meter increments along the two 60 m plot axes. Data from NRRI consisted of three variable-radius subplots per plot sampled using English BAF 10 prisms (Fig. 2B). Species composition data were then used as dependent variables (total BA and the relative BA for 12 forest types) in the PLS regression models.

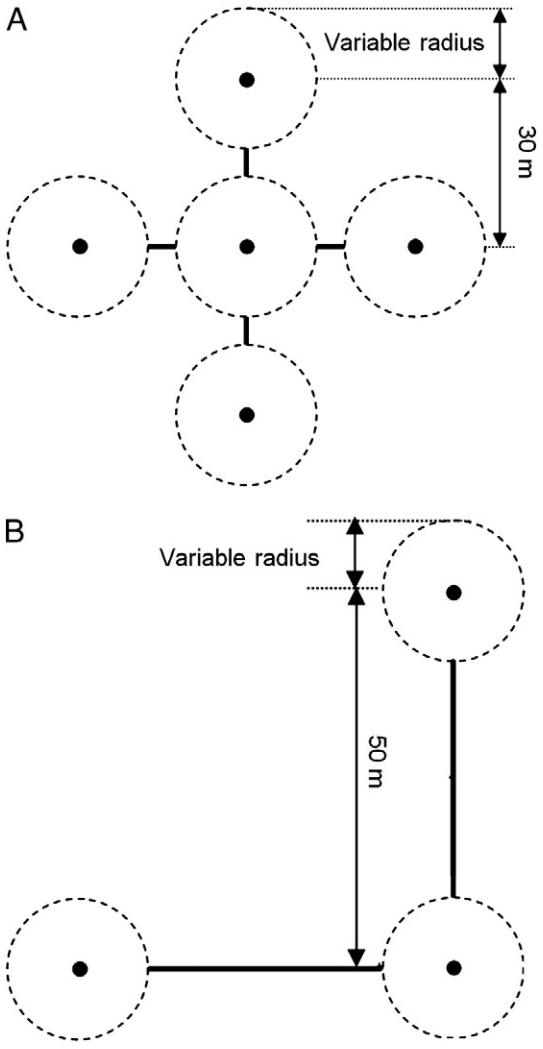

Fig. 2. (A) layout of field plots for this study (collected 2006–2007) and an earlier study (2003–2004, Townsend and Sturtevant unpublished data) designed for integration with 30 m satellite data, and (B) plot layout for data collected by the Natural Resources Research Institute (NRRI) of the University of Minnesota—Duluth for forest bird habitat modeling. Estimates of basal area (BA) by species for a plot were collected at each subplot using angle count sampling with a metric basal area factor (BAF) two prism (A) or using an English factor 10 prism (B). NRRI subplot measurements of basal area (ft^2ac^{-1}) were averaged by plot and converted to $m^2\ ha^{-1}$ to account for differences in prism BAF.

2.3. Satellite remote sensing data

Landsat's six reflective bands (30 m), mapped estimates of forest canopy structure derived from 5 and 10 m SPOT-5 sensor data (31 August and 15 July 2006, respectively) reported by Wolter et al. (2009), and amplitude data from two SAR sensors (Radarsat and PALSAR, Table 2) were used to calibrate PLS regression models for estimating relative BA for 12 forest species types. Wolter et al. (2009) provides a detailed description of the SPOT-5 sensor data (bands and processing) used to produce the forest canopy structure estimates used in this study.

2.3.1. Synthetic aperture radar

Six Radarsat-1 amplitude images and one PALSAR amplitude image (fine beam, 34.3° incidence angle, Table 2) were acquired and registered to UTM zone 15 coordinates. Radarsat-1 images collected during summer and winter (Table 2) were selected to minimize effects due to environmental variability (Ranson & Sun, 1994). Radarsat-1 is a C-band SAR sensor (5.3 GHz, 5.6 cm wavelength) with horizontal send and receive (HH) polarization that is capable of

Table 2
Satellite image data used in this study. Radarsat-1data (HH, C-band, 5.3 GHz) have a pixel resolution of 30 m and were collected as standard beam 2 which have an earth incidence angle of 27.7°. PALSAR data (HH and HV, L-band, 1.27 GHz) have a 12.5 m pixel resolution and an earth incidence angle of 34.3°. Temperatures on the ground at the time of SAR data collection are listed.

Landsat-5 date	Radarsat date	Radarsat temp. (°C)	PALSAR date	PALSAR temp. (°C)
3/3/2008	7/27/2006	17.8	10/19/2008	8.9
5/4/2007	8/20/2006	7.8		
6/9/1997	12/18/2006	−11.1		
8/26/2008	1/11/2007	−2.2		
9/24/2001	2/4/2007	−30		
9/27/2002	2/28/2007	−3.9		
10/14/1985				

collecting data at several incidence angles. Radarsat-1 data used in this study are all standard mode 2 (S2) products, which have an average incidence angle of 27.5°. Standard mode 2 provided the best repeat coverage of our study area and a good compromise between undesirable sensitivity to low local slope effects at steeper angles and terrain shadowing effects at shallower angles (see Singhroy & Saint-Jean, 1999). PALSAR is an L-band SAR sensor (1.27 GHz, 23.6 cm wavelength) which has several operational modes. The dual polarization PALSAR data used in this study were collected as part of the Japan Aerospace Exploration Agency's (JAXA) comprehensive acquisition strategy (see Rosenqvist et al., 2006), where incidence angle and other sensor parameters were fixed in advance to streamline the acquisition of spatially and temporally consistent global coverage. Hence, both HH and horizontal send and vertical receive (HV) energy are recorded at a 12.5 m pixel resolution with an incidence angle of 34.3°. Because SAR data are sensitive to the dielectric properties of vegetation and soil, temperature at the times of data acquisition were noted for later interpretation (Table 2). Both the Radarsat-1 and PALSAR image data were processed (radiometric, geo-location, and terrain correction) using Alaskan Satellite Facility (ASF) software MapReady 2.1 (ASF, University of Alaska—Fairbanks, Fairbanks, AK).

SAR image data are inherently noisy due to signal-dependent speckling (Lee, 1981), and are often filtered (sacrificing spatial information) to dampen these effects (Sader, 1987; Wu & Sader, 1987; Lopes et al., 1993; Rauste, 2005; Amini & Sumantyo, 2009). However, neighborhoods surrounding any given SAR pixel (e.g., mean or standard deviation, Sader, 1987) may contain far more valuable information than any one pixel (Dell'Acqua et al., 2006). To extract such information and avoid edge effects, pixel-wise local SAR statistics were collected by first using a 10 m, multi-spectral (red, green, near-infrared, and shortwave infrared), SPOT-5 image (15 July 2006) to define spectral Euclidean neighborhoods (distance ≤ 10 digital numbers) within a 97.5 m radius from the center pixel location (following Wolter et al., 2009). After a SPOT-5 neighborhood was defined for a pixel location, the mean and standard deviation of all the corresponding SAR image pixels in the neighborhood were calculated and written to separate image files. Finally, two shaded relief images were produced from 10 m digital elevation models (DEM) (source: www.usgsquads.com/elevationdata.htm) using the respective satellite's ephemeris information to model earth surface illumination by microwave energy. The two relief images were derived specifically to be used during PLS regression to account for differences in observed SAR backscatter related to topography (see Ranson et al., 2001).

2.3.2. Landsat image data

Seven Landsat Thematic Mapper (TM) images were acquired for this study based on temporal proximity to field data collection (2006–2007) or whether they captured specific forest phenology events (Tables 1, 2). All seven Landsat images were downloaded from the USGS Earth Resources Observation and Science Center (EROS) web site (source: http://glovis.usgs.gov/) in UTM zone 15 coordinates. These freely available 30 m Landsat images are convenient as they are precision-orthorectified and geo-corrected using Global Land Cover Facility (GLCF) GeoCover data (source: www.landcover.org). The seven Landsat images used in this study exhibited excellent pixel coregistration with each other. However, these data should be used with some caution as cubic convolution rather than nearest neighbor resampling is employed, which is known to negatively affect some analyses (Jensen, 2005). Each Landsat image was converted to top of atmosphere reflectance using published sensor calibration coefficients (Thome et al., 2004) and then topographically corrected (see

Table 3
List of predictor variables used with PLS regression and ground plot data to estimate species-wise relative basal area. Dates for Landsat bands and indices are indicated using the letters R (March), M (May), J (June), A (August), S (24 Sept.), T (27 Sept.), and O (October). For example, J2 and SVR_J represent June TM2 and SVR, respectively.

Predictor image	Description
R1, R2, R3, R4, R5, R6, R7, R8, R9	3 March 2008 Landsat bands (blue, green, red, NIR, SWIR5, SWIR7, respectively) and tasseled cap components (Crist & Kauth, 1986) brightness (7), greeness (8), and wetness (9).
M1, M2, M3, M4, M5, M6, M7, M8, M9	4 May 2007 Landsat bands and tasseled cap components.
J1, J2, J3, J4, J5, J6, J7, J8, J9	9 June 1997 Landsat bands and tasseled cap components.
A1, A2, A3, A4, A5, A6, A7, A8, A9	26 August 2008 Landsat bands and tasseled cap components.
S1, S2, S3, S4, S5, S6, 57, 58, S9	24 September 2001 Landsat bands and tasseled cap components.
T1, T2, T3, T4, T5, T6, T7, T8, T9	27 September 2002 Landsat bands and tasseled cap components.
O1, O2, O3, O4, O5, O6, O7, O8, O9	14 October 1985 Landsat bands and tasseled cap components.
SAVI	Soil-adjusted vegetation index (Huete, 1988).
SR	Simple NIR/RED ratio (Jordan, 1969).
MSI	Moisture stress index (Rock et al., 1986).
GEMI	Global environmental monitoring index (Pinty & Verstraete, 1992).
SVR, SVR7	Shortwave infrared to visible ratio (Wolter et al., 2008): SWIR/visible & SWIR7/visible.
ASH_J, ASH_S, ASH_T	SR from leaf-off black ash (J: 6/9/1997, S: 9/24/2001, T: 9/27/2002) minus SR from 8/26/2008.
TAM_R, TAM_M	SR from 3/3/08 and 5/4/07 (eastern larch leaf-off) minus SR from 8/26/08; respectively.
AI_S, AI_T, AI_O	Autumn indices (TM3/TM1) for 9/24101, 9/27/02, and 10/14/85 (J. Vogelmann pers. comm.)
R1M, R2M, R3M, R4M, R5M, R6M	Radarsat neighborhood means (M), where 1 = 7/27/06, 2 = 8/20/06, 3 = 12/18/06, 4 = 1/11/07, 5 = 2/4/07, and 6 = 2/28/07.
R1S, R2S, R3S, R4S, R5S, R6S	Radarsat neighborhood standard deviation (S), where 1 = 7/27/06, 2 = 8/20/06, 3 = 12/18/06, 4 = 1/11/07, 5 = 2/4/07, and 6 = 2/28/07.
R13, R14, R15, R16, R23, R24, R25, R26	Ratios of the means for the specified Radarsat images (e.g., R13 = R1M / R3M and so forth).
RLF, PLF	Shaded relief using Radarsat and PALSAR ephemeris, respectively.
PHM, PVM	PALSAR neighborhood means for HH and HV, respectively.
PHS, PVS	PALSAR neighborhood Std. Dev. for HH and HV, respectively.
PAL	PALSAR ratio PHM / PVM.
CDIA, DBH, HT	SPOT-based canopy diameter, bole diameter, tree height (Wolter et al., 2009).
CC, LC, BA	SPOT-based crown closure, live crown length; basal area (Wolter et al., 2009).

Riaño et al., 2003) using a 30 m digital elevation model (source: http://ned.usgs.gov/) and scene-specific sun position information. However, Landsat scene-to-scene radiometric normalizations were not performed.

Wolter et al. (2008) identified key Landsat image bands, dates, ratios, and indices that were import for estimating forest basal area in Minnesota (Table 3). In particular, the shortwave infrared (SWIR) to visible ratio (SVR), which is the average of both Landsat SWIR bands divided by the average of the three visible bands, was the strongest predictor of fir and spruce relative BA. Here, we added the SVR7 ratio which uses only the second Landsat SWIR band (TM7, 2.09–2.35 μm) in the calculation of the ratio because TM7 has been shown to be more sensitive to forest basal area than TM5 (Brockhaus & Khorram, 1992). We also employed the autumn index (AI = TM3/TM1, J. Vogelmann pers. comm.) for the September and October images because AI specifically enhances elevated red reflectance due to autumn foliar senescence (Table 3). The late September images were selected from the Landsat archive as these dates fall approximately within the timeframe of maximum maple and paper birch senescence for this region (Ahlgren, 1957; Eder, 1989; Wolter et al., 1995).

2.3.3. SPOT-5 structure information

Wolter et al. (2009) successfully estimated six forest structural parameters for the current study area using pixel-wise neighborhood statistics (standard deviation, variance, semivariogram sill variance, and ratios of these metrics at 5 and 10 m) calculated from multi-spectral and panchromatic SPOT-5 sensor data and derivatives. Estimates were derived using PLS regression and used to map tree canopy diameter (CDIA), bole diameter at breast height (DBH), tree height (HT), crown closure (CC), vertical length of live crown (LC), and basal area (BA). To make use of this structure information for estimating forest species relative BA, the original 10 m estimates were first smoothed using a 3×3 median filter then transformed to 30 m using nearest neighbor resampling (Table 3). For details on SPOT data processing see Wolter et al. (2009).

2.4. PLS regression model development and forest composition estimation

Nine PLS regression models were developed to explore the sensitivity of 147 image variables (Table 3) derived from four satellite sensors (Landsat TM, SPOT-5, Radarsat-1, and PALSAR), including different sensors combinations (Table 4), for simultaneously estimating total basal BA and relative BA for 12 forest types (dependent variables Table 5) using plot data collected in 2006–2007 (n = 120). A recursive backward elimination band selection approach was developed for use with the PLS regression procedure (see Wolter et al., 2008) to pre-select the most relevant image predictor variables. Once the best model was identified in each case, band-wise PLS regression coefficients for respective relative BA estimates by species were used to produce species-specific maps of relative BA.

Table 4
Nine PLS regression models were tested for estimating forest composition. The acronym SAR indicates combined use of Radarsat-1with PALSAR. A lower predicted residual sum of squares (PRESS) statistic indicates a better overall PLS regression model.

Sensor combinations studied	Number of images	Number of variables	Variables selected	Number of components	Model PRESS
Landsat TM + SAR + SPOT	16	147	63	55	0.43
Landsat TM + SAR	14	141	72	52	0.48
Landsat TM + SPOT	9	119	57	48	0.46
SAR + SPOT	9	34	31	29	0.70
SAR	7	28	27	26	0.82
LandsatTM	7	113	56	48	0.52
PALSAR	1	6	6	6	0.91
Radarsat	6	22	19	18	0.87
SPOT	2	6	6	6	0.86

Table 5
Dependent forest composition variables modeled using PLS regression. Relative basal area (RBA) values for separate species or types represent percentages of the total basal area (i.e. species BA / total BA).

Dependent variable	Description
TBA	Total basal area (m^2 ha^{-1})
MPL	Relative BA sugar and red maple
ASP	Relative BA quaking and bigtooth aspen
BIR	Relative BA paper birch
ASH	Relative BA black ash
JP	Relative BA jack pine
WP	Relative BA white pine
RP	Relative BA red pine
BS	Relative BA black spruce
WS	Relative BA white spruce
FIR	Relative BA balsam fir
WC	Relative BA white cedar
EL	Relative BA eastern larchl

The relative BA models were validated during the PLS regression procedure using leave-one-out cross validation (Gong, 1986), while the resulting relative BA estimates were assessed using Receiver Operating Characteristic (ROC) curves (Hanley & McNeil, 1982; Kerekes, 2008). ROC curves were used to specifically determine the optimal lower limit of detection for a given forest type at a given pixel location. The goal of ROC curve analysis is to find a balance between prediction sensitivity or the true positive rate of detection (the benefits) and the false positive rate (the cost). Area under the curve (AUC) and Youden's index (Youden, 1950) both provide single metrics that summarize the accuracy of this test, where

$$\text{Youden's Index} = 1 - (\text{false positive rate} + \text{false negative rate}). \quad (1)$$

2.5. Mapping relative basal area

The best relative BA cut-off values resulting from ROC curve analysis were used to exclude problematic predictions from being mapped. Once this was done for each forest species type, mapped estimates were then used to construct a forest type map. While these spatially explicit relative BA estimates for the 12 forest types can be combined in a number of different ways to produce forest cover maps, here we constructed a map of the single dominant forest species type occurring at each pixel location. The resulting forest cover map was validated using field data collected in summer of 2003–2004 and field data collected by NRRI (N. Danz, unpublished data, University of Minnesota—Duluth) in 2006 (n = 138) for a total validation set of n = 203.

Prior to accuracy assessment, the forest cover type map was smoothed using a 3×3 majority filter. Then, ground data points were used to extract pixel information, which was used to produce a confusion matrix for determining the user's, producer's, and overall accuracy (Congalton & Green, 1999) of this dominant forest type classification. Note, however, that all data on relative dominance by species type are retained and can be used to map any user-defined combination of species.

3. Results

3.1. Single and multi-sensor model comparisons

Multi-temporal Landsat image data (see Tables 2, 3) yielded the best single-sensor PLS regression models for estimating species type composition (Fig. 3), with one exception (SPOT for white pine). Structural variance in plot data explained by the Landsat-based forest

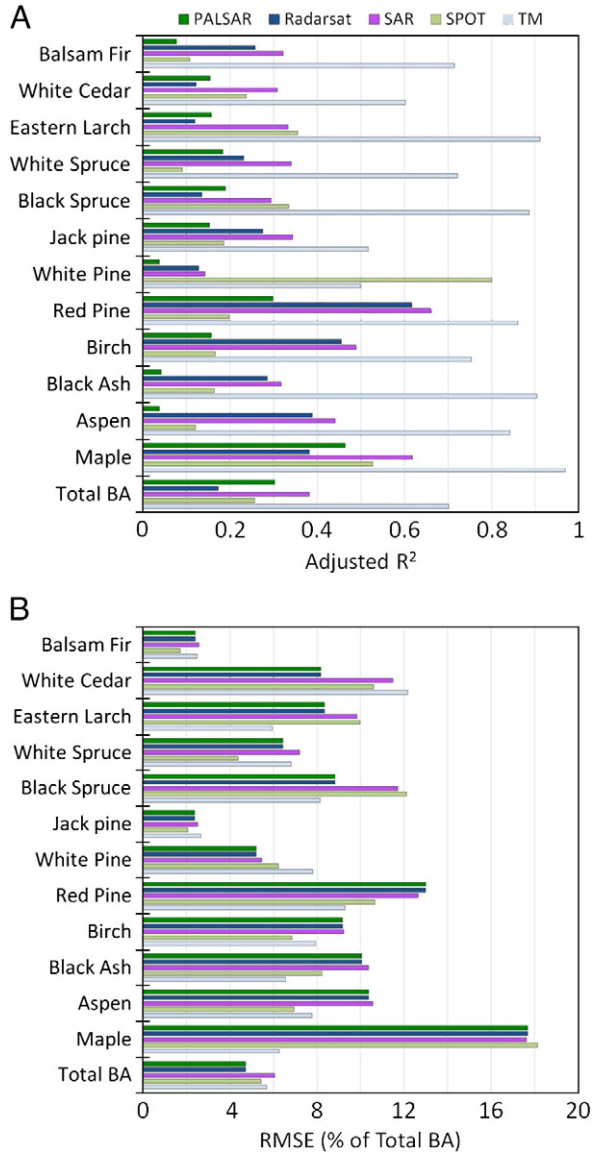

Fig. 3. PLS model results for species relative basal area and total basal area using data from individual satellite sensors and Radarsat-1and PALSAR combined (SAR). Landsat was superior to the other sensor data (except for white pine) with the highest adjusted R^2 (A) and lowest root mean squared error (B) values. Note total basal area RMSE units are $m^2\ ha^{-1}$.

composition models (R^2 range 0.52–0.97) was often twice that explained using SPOT structure data (R^2 range 0.12–0.80), Radarsat-1data (R^2 range 0.12–0.62), or PALSAR data (R^2 range 0.08–0.47) alone (Fig. 3A). Moreover, 10 out of the 13 Landsat-based composition estimates showed similar or lower RMSE values compared to the other single-sensor models (Fig. 3B). Estimates of composition based on multi-temporal Radarsat-1 amplitude data and derivatives (see Tables 2, 3) ranked second behind Landsat in seven out of the 13 cases, ahead of SPOT and PALSAR, respectively. In every instance, the combined use of Radarsat-1and PALSAR (hereafter referred to simply as SAR) to estimate total BA and species relative BA yielded coefficients of determination substantially higher (R^2 range 0.14–0.66) than either sensor's individual results (Fig. 3A). However, RMSE values for the combined SAR model estimates were also generally higher (except for maple and red pine) than estimates produced using either Radarsat-1or PALSAR sensor data alone (Fig. 3B). SPOT-based structure information was second only to Landsat data for explaining variation in eastern larch ($R^2 = 0.35$, RMSE = 9.98%) and black spruce ($R^2 = 0.33$, RMSE = 12.13%) relative BA, while the SPOT-based white pine relative BA estimates ($R^2 = 0.80$, RMSE = 6.24%) eclipsed all other sensor-specific models including Landsat ($R^2 = 0.50$, RMSE = 7.81%) for characterizing this species (Fig. 3).

Models that employed data from all the sensors (Landsat, SPOT, and SAR) outperformed all other sensor combinations (Table 4) in terms of proportion of ground variation in forest composition explained (Fig. 4A) and, with the exception of white cedar (RMSE = 7.53%, Landsat-SPOT), error among the respective estimates (Fig. 4B). Total variation in ground data explained by the Landsat-SAR-SPOT model was 86.3% (Table 6). Estimates produced using either Landsat with SPOT or Landsat with SAR explained nearly as much variance in species structure among plot data with similar or slightly higher error in most cases (Fig. 4A, B). Structure models produced using a combination of SPOT and SAR data explained the least amount of variation in plot data and had the highest error among the respective estimates (Fig. 4), except for white pine relative BA ($R^2 = 0.85$, RMSE = 5.54%), which rivaled the TM-SPOT ($R^2 = 0.90$, RMSE = 4.73%) and TM-SPOT-SAR model results ($R^2 = 0.91$, RMSE = 4.53%).

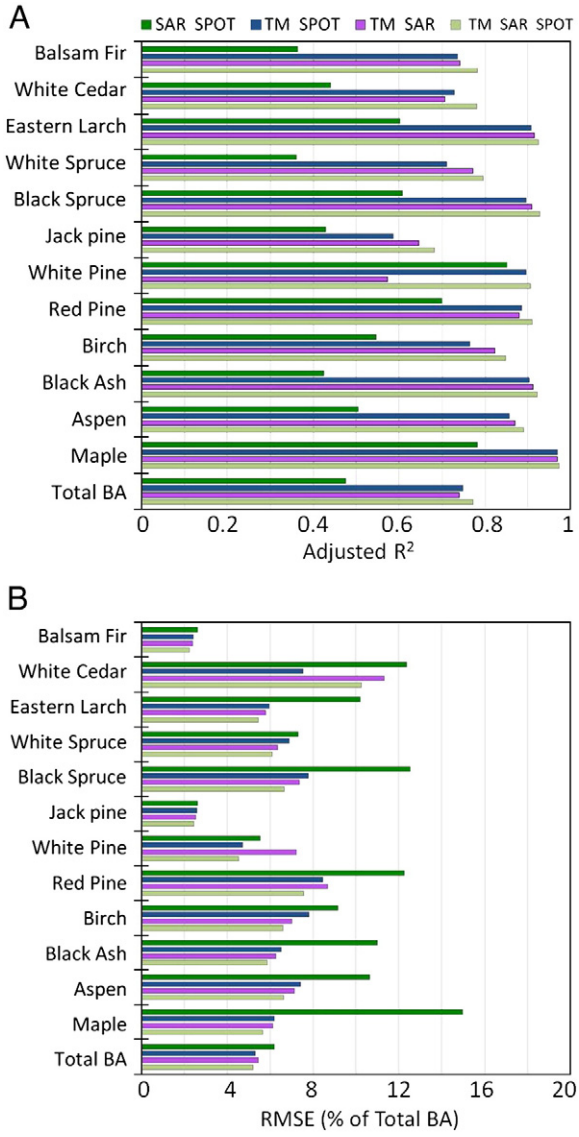

Fig. 4. PLS model results for species relative basal area and total basal area using combinations of different satellite sensor data, where SAR refers to use of Radarsat-1with PALSAR data. While collective use of data from all the sensors studied produced the highest adjusted R^2 values (A) and the lowest root mean squared errors (B) in most cases, other sensor combinations that included Landsat rivaled the best model results. Note total basal area RMSE units are $m^2\ ha^{-1}$.

Table 6
PLS regression model cross validation results for total (m^2 ha^{-1}) and relative (%) basal area (relative BA = total BA/species BA). PLS regression models were calibrated using ground data collected specifically for this study. Note total basal area RMSE is in m^2 ha^{-1} (*).

Image variables used		63
PRESS		0.43
Components		55
Model variation explained (%)		99.9
Variation in field data explained (%)		86.3
Model Pr>F		0.0001
Forest parameter statistics	Adj. R^2	RMSE (%)
Total basal area (m^2 ha^{-1})	0.77	5.20*
Maple (sugar and red)	0.98	5.65
Aspen (quaking and bigtooth)	0.89	6.64
Paper birch	0.85	6.60
Black ash	0.92	5.89
Jack pine	0.68	2.46
White pine	0.91	4.53
Red pine	0.91	7.57
Black spruce	0.93	6.57
White spruce	0.80	6.11
Balsam fir	0.78	2.26
Northern white cedar	0.78	10.26
Eastern larch	0.93	5.44

3.2. PLS estimates and component loadings

In general, Landsat-SAR-SPOT model estimates for species relative BA (except maple) slightly underestimated observed relative BA, with jack pine, white cedar, and white spruce showing the greatest deviations from unity (Figs. 5, 6). High R^2 values among relative BA estimates are inflated in cases where stands consistently lacked intermediate abundance values—where species abundance was either very high or low. This was especially true for sugar maple and eastern larch (Figs. 5, 6) which occur almost exclusively as pure stands within the study area. Species that typically occur across a more complete distribution of relative BA values (such as aspen and birch) reflect a more realistic portrayal of PLS model performance.

In any event, iterative variable selection based on PLS regression (Wolter et al., 2008) was used to optimize four sensor-specific and five multi-sensor models for estimating total forest BA and species relative BA (Table 4). In each case, resulting component loading information relates the reduced set of predictor bands (e.g., 63 bands used in the Landsat-SAR-SPOT model) to each of the 13 forest composition parameters. However, due to the volume of data, we report only component loading results for the best PLS model (Fig. 7, Landsat-SAR-SPOT). Positive and negative component loadings results are indicated below using '+' and '−' symbols, respectively, preceding the image predictor variable name.

3.2.1. Hardwood component loadings

Overall, Landsat predictor bands and derivatives exhibited greater influence than SAR and SPOT variables for estimating hardwood relative BA (Fig. 7, Table 3). The most heavily weighted bands for determining maple relative BA were greenness indices from 9 June 1997 and 26 August 2008 (+J8, −GEMI_A, and −SAVI_J), followed closely by 3 March 2008 NIR (−R4), 4 May 2007 brightness (+M7), and the 24 September 2001 autumn index (+AI_S). The strongest non-Landsat loading for maple was SPOT canopy diameter (+CDIA), followed by SPOT canopy closure (−CC), −R26 (leaf-on/off ratio of Radarsat-1neighborhood means 20 August 2006 / 28 February 2007), SPOT length of live crown (−LC), and −R23 (leaf-on/off ratio of Radarsat-1neighborhood means 20 August 2006 / 18 December 2006). All other SPOT and SAR variables had lesser loading weights for maple relative BA.

The strongest component loadings for aspen relative BA came from a mix of Landsat greenness and SWIR-related bands and indices: 24 September TM4/TM3 simple ratio (+SR_S), June SWIR7/visible ratio (+SVR7_J), May moisture stress index (+MSI_M), 24 September SAVI (−SAVI_S), as well as June Tasseled Cap wetness, SWIR7, and Tasseled Cap greenness (+J9, −J6, and +J8, respectively). The

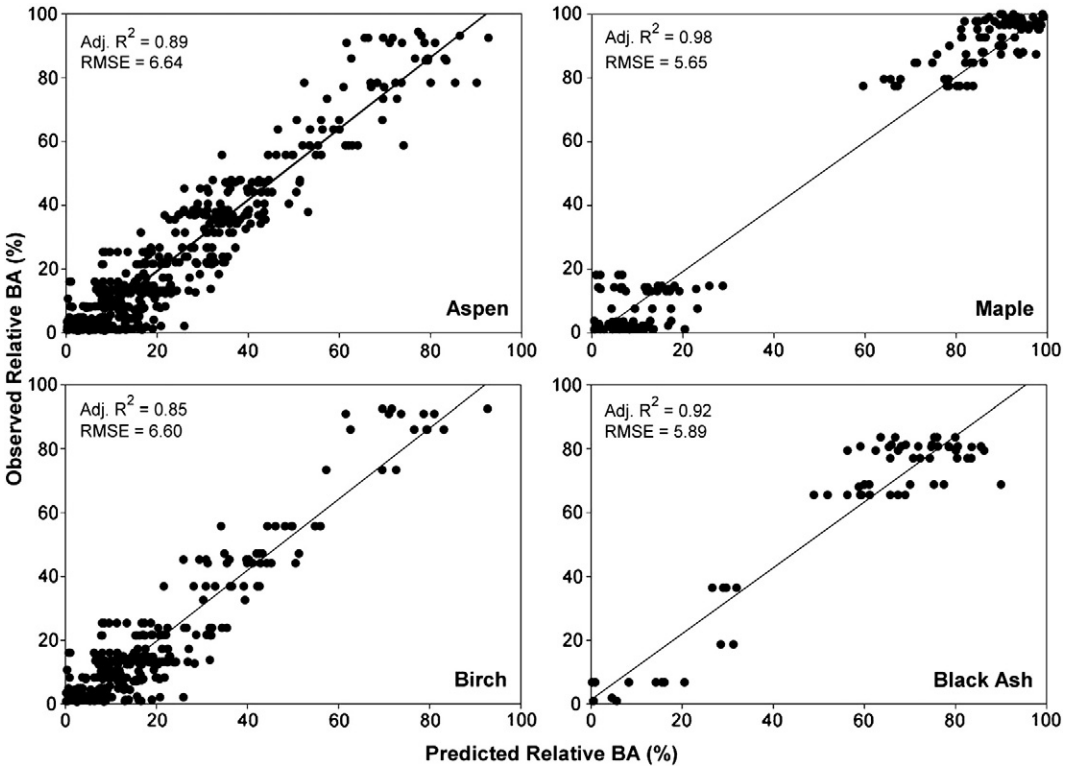

Fig. 5. PLS regression results for relative basal area of deciduous hardwood species. Estimates for maple are close to the 1:1 line while the aspen, birch, and black ash models slightly underestimate observed relative basal area.

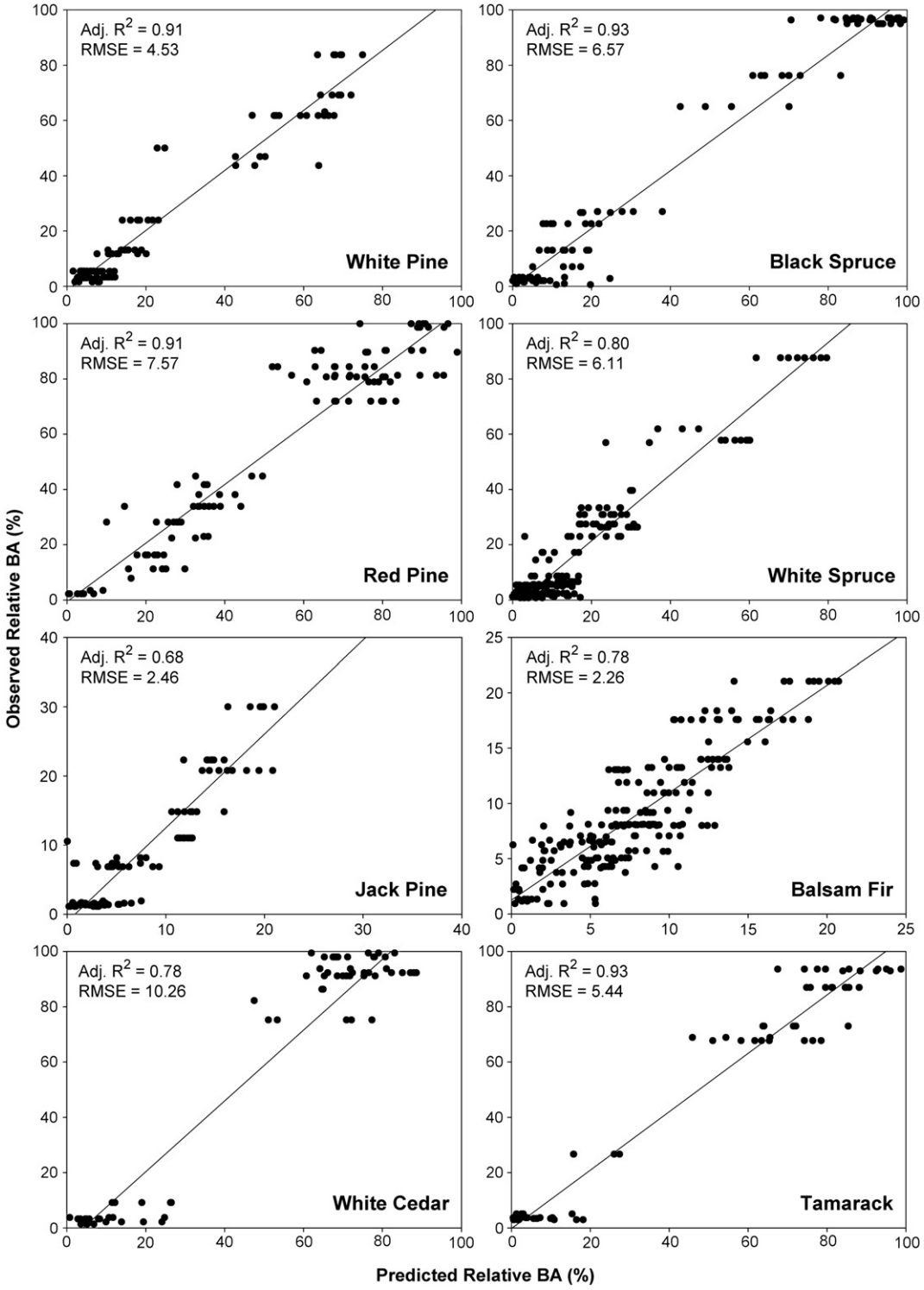

Fig. 6. PLS regression results for relative basal area of conifer species. In general models are close to the 1:1 line. However, jack pine, white cedar, and white spruce underestimate relative basal area at high levels of abundance.

strongest non-Landsat predictor variables for aspen relative BA were −R26, +R4S (leaf-off Radarsat-1neighborhood SD for 11 January 2007), SPOT canopy closure and diameter (−CC and −CDIA), and three other Radarsat-1variables: −R2S (leaf-on SD, 20 August 2006), −R6M (leaf-off mean, 28 February 2007), and +R23 (leaf-on/off ratio).

Top ranked component loading for paper birch relative BA were also dominated by Landsat variables with emphasis on 27 September 2002 wetness (−T9), March SWIR5 and NIR (−R5 and +R4), and June SWIR5 (−J6). Leading SPOT and SAR variables were −LC, +R2S, +CDIA, +R1S (leaf-on Radarsat-1neighborhood SD 27 July 2006), PALSAR HV and HH neighborhood means (−PVM and −PHM, respectively), and SPOT tree height (+HT) in order of decreasing magnitude.

For black ash relative BA, Landsat variables ranked ahead of SAR and SPOT variables, but there were no clearly prominent loadings among Landsat-based variables. March Tasseled Cap greenness (−R8) had the greatest magnitude but the next four ranking variables had similar weights (−SAVI_O, −J8, −M7, and −J9). Among the SPOT and SAR variables, canopy closure (+CC) was clearly the dominant

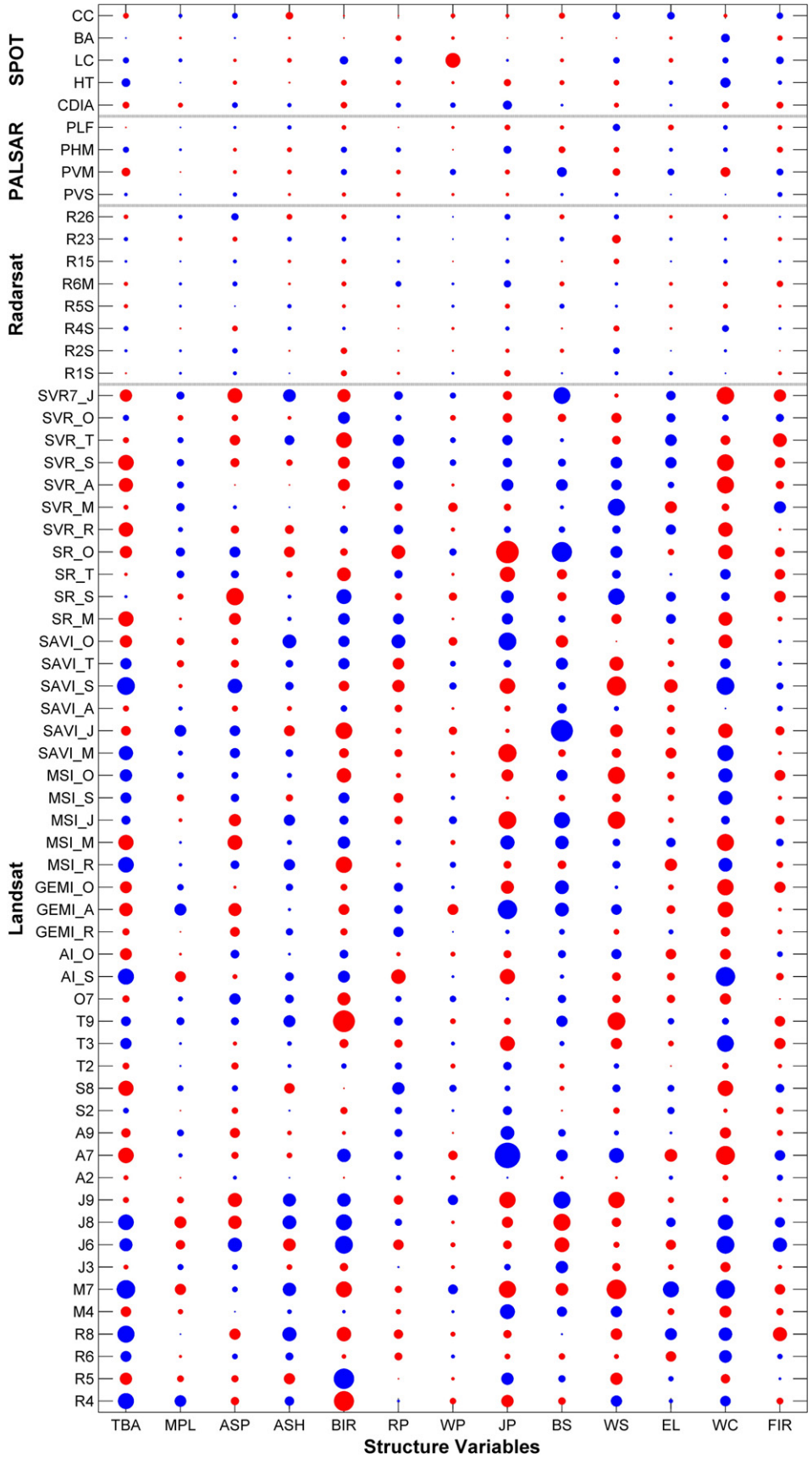

Fig. 7. Scaled PLS component loadings for the reduced set of 63 bands used by the best combination of sensors (Landsat-SAR-SPOT) to model forest composition. Positive and negative loadings are depicted as red and blue filled circles, respectively, while circle size indicates loading magnitude. Overall, Landsat variables and derivatives are loaded more heavily than either SAR or SPOT-based variables for describing forest composition.

predictor over +R26, +PHM, +LC, +PVM, and −R23, respectively (Fig. 7).

3.2.2. Conifer component loadings

Landsat variables also generally outweighed SAR and SPOT variables for estimating conifer relative BA. However, PALSAR and SPOT-based structure variables (Wolter et al., 2009) were loaded more strongly for conifer than for hardwoods, while Radarsat-1 loading magnitudes were similar for hardwoods and conifers (Fig. 7). Among the Landsat variables, greater emphasis was placed on SVR, MSI, and greenness-related indices compared to the hardwood component loadings for relative BA (Fig. 7). Landsat variables used to estimate red and white pine relative BA had a lower, more homogenous range of loading magnitudes compared to other conifer species. Overall, top ranked loadings for red pine were +AI_S, −SAVI_O, and +SR_O, while white pine's top loadings were +LC, +GEMI_A, and June wetness (−J9). The white pine model was unique among the 12 forest types in that it was the only case where a SPOT-based variable (+LC) was strongest overall for estimating relative BA. Moreover, +LC was a full order of magnitude greater than any other SPOT or SAR variable for estimating white pine relative BA (Fig. 7). For red pine, SPOT-based basal area (+BA), −R6M (leaf-off), −PHM, +HT, and −CDIA were all approximately one half the magnitude shown for −LC. For jack pine, the two dominant variables overall were −A7 and SR_O, while −CDIA, −PHM, −R6M and +HT were dominant among the SAR and SPOT variables.

Top ranked component loadings for white spruce (+M7, +SAVI_S, +T9, and +MSI_J), black spruce (−SAVI_J, −SR_O, +J8, and −J9), and balsam fir (−J6, +R8, +SVR_T, and +SVR7_J) were dominated by Landsat greenness-related and SWIR-based indices and bands, with balsam fir showing lower overall component magnitudes (Fig. 7). Moreover, black spruce component loadings were generally negative while white spruce and balsam fir were positive. Among the SAR and SPOT loadings, PALSAR PVM was among the top two variables for estimating balsam fir (+LC, −PVM, +CDIA, and −CC), black spruce (−PVM, +PHM, +CC, and +R6M), and white spruce (+R23, +PVM, −CC, −PLF) relative BA.

Lastly, the strongest four component loadings for estimating white cedar (−AI_S, −M7, +A7, and −SAVI_S) and eastern larch (−M7, +SAVI_S, +A7, +MSI_R) relative BA were similar in that they shared three variables. SAR and SPOT-based variables −HT, +PVM, and −BA were ranked highest for white cedar, while −CC, −PVM, and +PLF were salient for eastern larch (Fig. 7), where PLF is a shaded relief image variable derived using a 30 m digital elevation model and PALSAR's orbital position and overpass time parameters (Table 3).

3.3. Receiver operating characteristic (ROC) curve analysis

We used ROC curves to identify the best lower limit of relative BA per species type for successful detection using PLS regression. In general, the ROC curves suggested higher limits for detection of hardwood species than conifers (Table 7). Paper birch was the exception in this regard with a best lower limit at 8.73% relative BA. As a result, paper birch had the lowest over all Youden's Index (YI) and AUC values indicating that birch is more difficult to accurately distinguish than other forest species. In contrast, black ash had the highest cut-off value at 39.41% RBA, which is consistent with the ground data in which very few plots had intermediate abundance of black ash (Fig. 5). As a result, the AUC and YI values were among the highest observed for all species. Maple was similar to black ash in this regard with a best cut-off value of 24.65%. However, the error rate for maple detection was slightly higher than black ash with a YI value of 0.86. On the other hand, estimates of aspen were more evenly distributed between the extremes. The best cut-off value of 22.05% for aspen indicates that values below this level are unreliable.

Table 7
Receiver operating characteristic (ROC) curve analysis, performed prior to assembling cover type maps, for selecting optimal species presence/absence cut-off values. Area under curve (AUC) (Hanley & McNeil, 1982) and Youden's index (Youden, 1950) are metrics for assessing test accuracy. Estimated relative basal area (RBA) by species and study-specific plot information were used as input data for the ROC tests.

Cover type	Best RBA cut-off (%)	AUC 0.5–1	AUC 95 % C.I		Youden's index (0–1)
			Lower	Upper	
Quaking/bigtooth aspen	22.05	0.97	0.88	1.01	0.84
Paper birch	8.73	0.89	0.81	0.97	0.64
Sugar/red maple	24.65	0.98	0.94	1.03	0.86
Black ash	39.41	1.00	0.98	1.02	0.98
White pine	8.49	0.97	0.91	1.04	0.85
Red pine	27.48	0.99	0.94	1.03	0.89
Jack pine	5.94	0.95	0.85	1.05	0.89
Black spruce	24.10	0.96	0.86	1.05	0.87
White spruce	11.76	0.96	0.89	1.04	0.86
Balsam fir	8.47	0.92	0.82	1.03	0.68
White cedar	26.16	0.95	0.85	1.05	0.89
Eastern larch	23.88	1.00	0.97	1.03	0.99

Like black ash and maple, white cedar and eastern larch occur in stands in which they are often the sole canopy species present. The disconnect between low and high estimates of relative BA for these species was closely linked to their optimal cut-off values from the ROC relative BA curve (Figs. 5, 6, 7, Table 7). This was not the case for the other conifer species. Cut-off values for balsam fir and white pine were 8.47 and 8.49%, respectively. However, the YI value for balsam fir was the lowest of the conifers at 0.68 which indicates detection error rates remain high. YI values among the remaining conifer species were similar with values between 0.85 and 0.89 indicating good detection ability. Although the ability to detect jack pine stands was good, relative BA estimates fell below observed values at higher jack pine abundance levels (Fig. 6).

Application of cut-off values derived from ROC curves prior to mapping effectively excluded estimates that would otherwise be problematic. The resulting maps of relative BA for hardwoods (Fig. 8) and conifers (Figs. 9, 10) give an indication of the pattern and abundance of species occurrence across the study area.

3.4. Forest cover type accuracy assessment

Based on independent field data, the overall accuracy of the map of dominant forest cover types was 78.0% (Khat 0.75, Table 7). Aspen and birch separated fairly well with user's accuracy of approximately 81%, but a lower producer's accuracy for birch (64%). Pine species were distinguished from spruce and fir with moderate to high user's accuracy (53–92%), while producer's accuracies were between 69% and 83%. Cedar and larch were distinguished from spruces and fir fairly well with user's accuracies of 80% and 58% while producer's accuracies were 89 and 88%, respectively. The accuracies for black ash and maple were among the highest reported with user's accuracies of 92% and 87% and producer's accuracies of 79% and 93%, respectively.

Much of the error reported above is linked to confusion between the two spruce species and between white spruce and white pine. Also, while only a limited number of dominant balsam fir plots were available, the only confusion with fir was white pine.

4. Discussion

4.1. Important image predictor variables

The use of PLS regression analysis to integrate data from multiple large-footprint sensors to estimate forest composition and structure is a novel approach that has been successfully tested in two other studies (Wolter et al., 2008, 2009). In this study, fusion of C-band and L-band SAR data with Landsat optical and SPOT-based structure data was unique. The potential of SAR backscatter data for characterizing

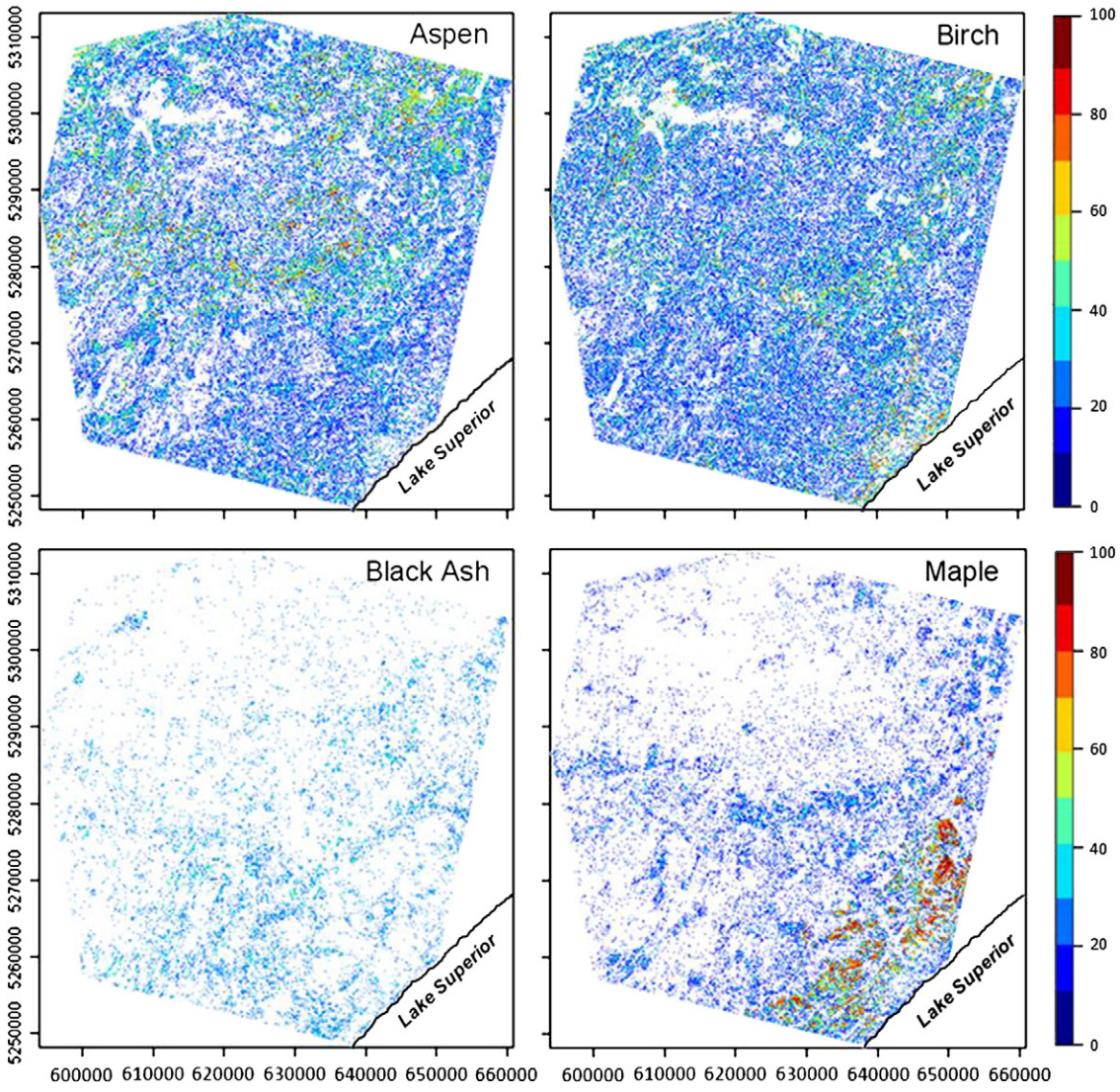

Fig. 8. Study-wide estimates for relative basal area (RBA) for four hardwood species. Aspen and birch show distinct patterns of high abundance, but at very low RBA the two species are ubiquitous on the landscape. Maple shows high abundance in the southeast along Lake Superior, while black ash abundance and distribution are more limited. Coordinates on the map axes are UTM zone 15 in meters.

forest structure in the U.S. Lake States is well known (Ahern et al., 1991, 1993, 1995; Dobson et al., 1992a,b, 1995). However, integration of fine scale, spatially explicit forest structure information (Wolter et al., 2009) with optical and SAR sensor data—specifically PALSAR—has not been demonstrated for mapping forest species composition.

4.1.1. Synthetic aperture radar—hardwoods

Of the 63 out of 147 variables selected under the best sensor combination for estimating forest composition, 11 were SAR variables (Fig. 7, eight Radarsat-1 and three PALSAR). In general, multi-season Radarsat-1amplitude data were moderately important for estimating relative BA for most species, while PALSAR variables were more strongly loaded for four of the conifer species (black spruce, white spruce, jack pine, and white cedar). Most of the species' relative BA models reported relatively strong loadings on SAR variables, with maple being the exception (Fig. 7). SAR variables were generally of less importance compared to the multi-temporal Landsat variables. Nevertheless, it may be worthwhile adding C-band data from spring when branch dielectric constants and radar interactions are high to improve species-level discrimination (Rignot et al., 1994).

C-band SAR data R3M and R3S from the coldest winter date (Table 2, −11.1 °C) were almost entirely excluded (used only in one leaf-on/off ratio, R23) from the best combined sensor models. Several factors could have contributed to the low overall relative sensitivity of the December Radarsat-1 data to forest composition differences. Extreme cold conditions, for one, are known to decreased the dielectric constant of canopy constituents enough, compared to partially frozen branches from warmer days (≥−5° to −7 °C), to retard C-band interactions with small branches (Lin, 1967). Such diminished C-band interaction can enhance penetration of the signal through the forest canopies (Kwok et al., 1994) whereby decreasing sensitivity to species-level differences. On warmer dates where temperatures were only a few degrees below freezing (Table 2), the forest floor could have remained fully frozen while branches were only partially frozen (see Lin, 1967). Moreover, thoroughly frozen snow cover behaves like dry soil (Pulliainen et al., 2003), which has little effect on microwave backscatter (Wegmuller, 1990). Under such conditions, the forest floor produces a lower overall contribution to backscatter than the stem and canopy (Koskinen et al., 2001), which may explain why Radarsat-1 imagery from warmer sub-zero dates were stronger factors than the much colder 12/18/2006 image date for discriminating forest species (Fig. 7).

It was encouraging that several Radarsat-1variables (leaf-on, leaf-off, and leaf-on/off ratios) were ranked relatively strongly for mapping aspen and paper birch abundance (often with opposite magnitude). Maple and black ash showed far less dependence on Radarsat-1variables

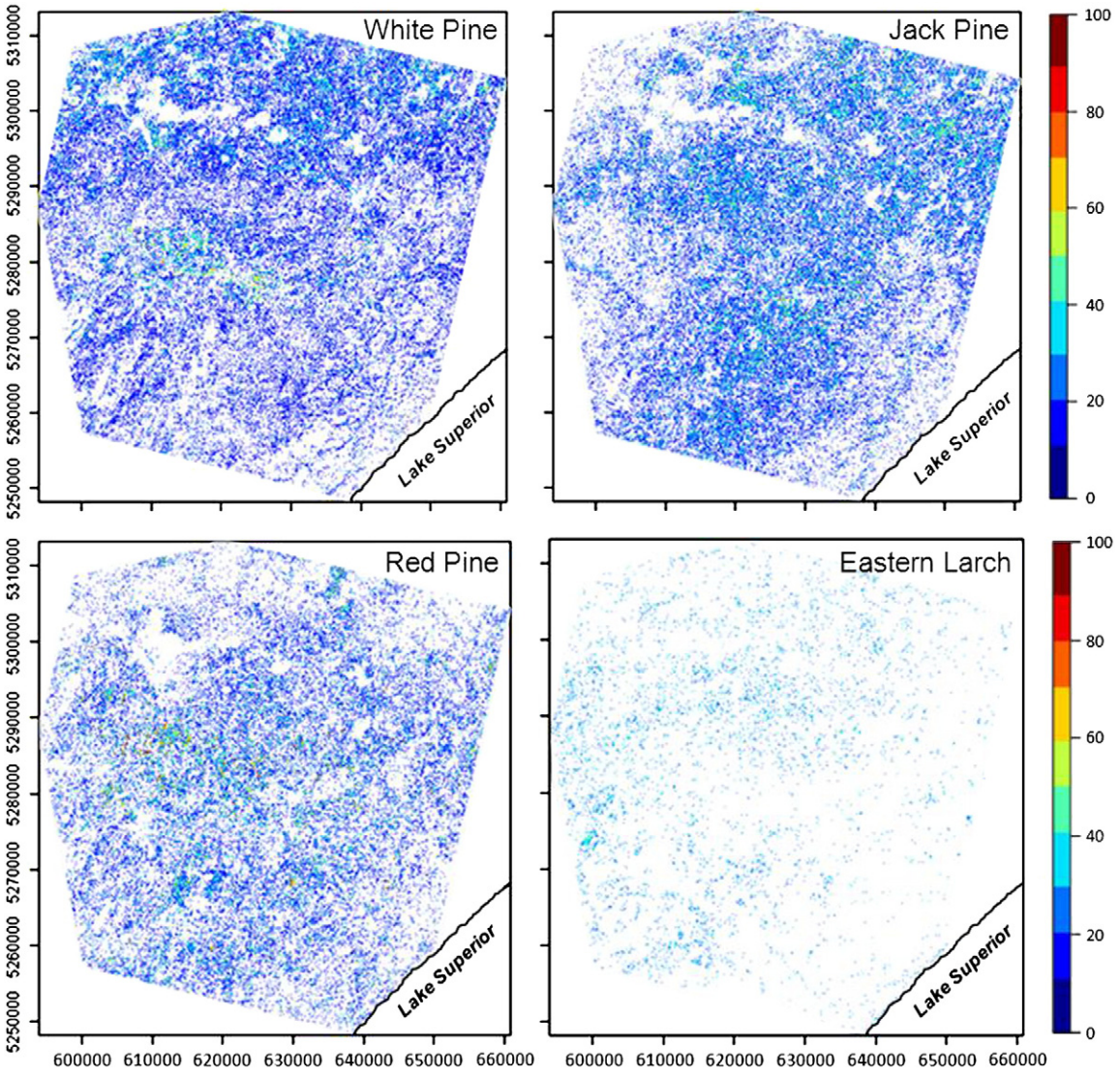

Fig. 9. Study-wide relative basal area (RBA) estimates for pines and eastern larch. White pine extent increases from south to north while highest RBA is west of the center of the study area and to the northwest. Red pine is more homogenous across the landscape with highest levels near the center of the study area. Jack pine distribution is distinct from the other pines showing highest RBA from south central to the north and northeast. Coordinates on the map axes are UTM zone 15 in meters.

(Fig. 7). Of note, −R23 and +R26 (both leaf-on/off ratios) were equally important for distinguishing black ash and paper birch. Each of these species has a unique variation in the allocation of biomass to stems, branches (very thick versus very fine twigs, respectively), and leaves that differs substantially from either aspen or maple. Even if total above ground biomass is similar, the known differences in biomass allocation between these hardwood species (Perala & Alban, 1994) can affect C-band backscatter (Imhoff, 1995), while branch-bole intersection geometry (i.e., corner reflectors, Sader, 1987) and leaf clumping difference (e.g., Chen et al., 1997) may also be diagnostic.

Predictor variables from seasonal Radarsat-1 imagery (Table 3) were employed specifically to examine whether C-band SAR imagery of hardwood foliage plus branching structure (leaf-on) and/or branching structure alone (leaf-off) might lend predictive power for distinguishing hardwood species. The elevated loadings on C- and L-band SAR variables for aspen (quaking and bigtooth) and paper birch are notable as these species have been difficult to discriminate using optical imagery (Shen et al., 1985; Wolter et al., 1995; Moore & Bauer, 1990; Franco-Lopez et al., 2001). Our results indicate that canopy-level differences such as leaf and fine branch structure (Perala & Alban, 1994; Ranson & Sun, 1994; Dobson et al., 1995) as well as bole arrangement and orientation (Cooper, 1913; Buell & Niering1957) are likely characteristics that facilitate discrimination of these species (e.g., Sader, 1987; Townsend, 2002). Similarly, maple (sugar and red) and black ash stands, like paper birch, differ structurally from aspen stands. Aspen is the most abundant hardwood species in this region and much of northern Minnesota (Reich et al., 2001; Wolter & White, 2002; Pastor et al., 2005). Hence, it follows that any reduction in confusion between aspen and other forest types, especially paper birch, can result in substantial improvements to overall accuracy, especially when aspen occurs in mixtures that include conifer species (Wolter et al., 1995; Wolter & White, 2002).

Black ash and paper birch models also weighted L-band PALSAR variables PHM and PVM (neighborhood mean amplitude for HH and HV, respectively) equally strong but with opposite signs (Fig. 7). Return energy for L-band HH polarization has been linked to trunk-ground terms in backscatter models and to a much lesser degree canopy volume scattering (Wang et al., 1995). L-band HV cross polarization return energy is known to be affected principally by lower canopy elements and much less by other upper canopy elements (Pope et al., 1994; Ranson & Sun, 1994; Ahern et al., 1995; Wang et al., 1995; Dobson et al., 1995). In our study area, paper birch trees frequently occur in clumps (Cooper, 1913; Buell & Niering1957) that may produce a unique double bounce signature (e.g., Imhoff, 1995; Townsend, 2002). On the other hand, both PHM and PVM are positively loaded on

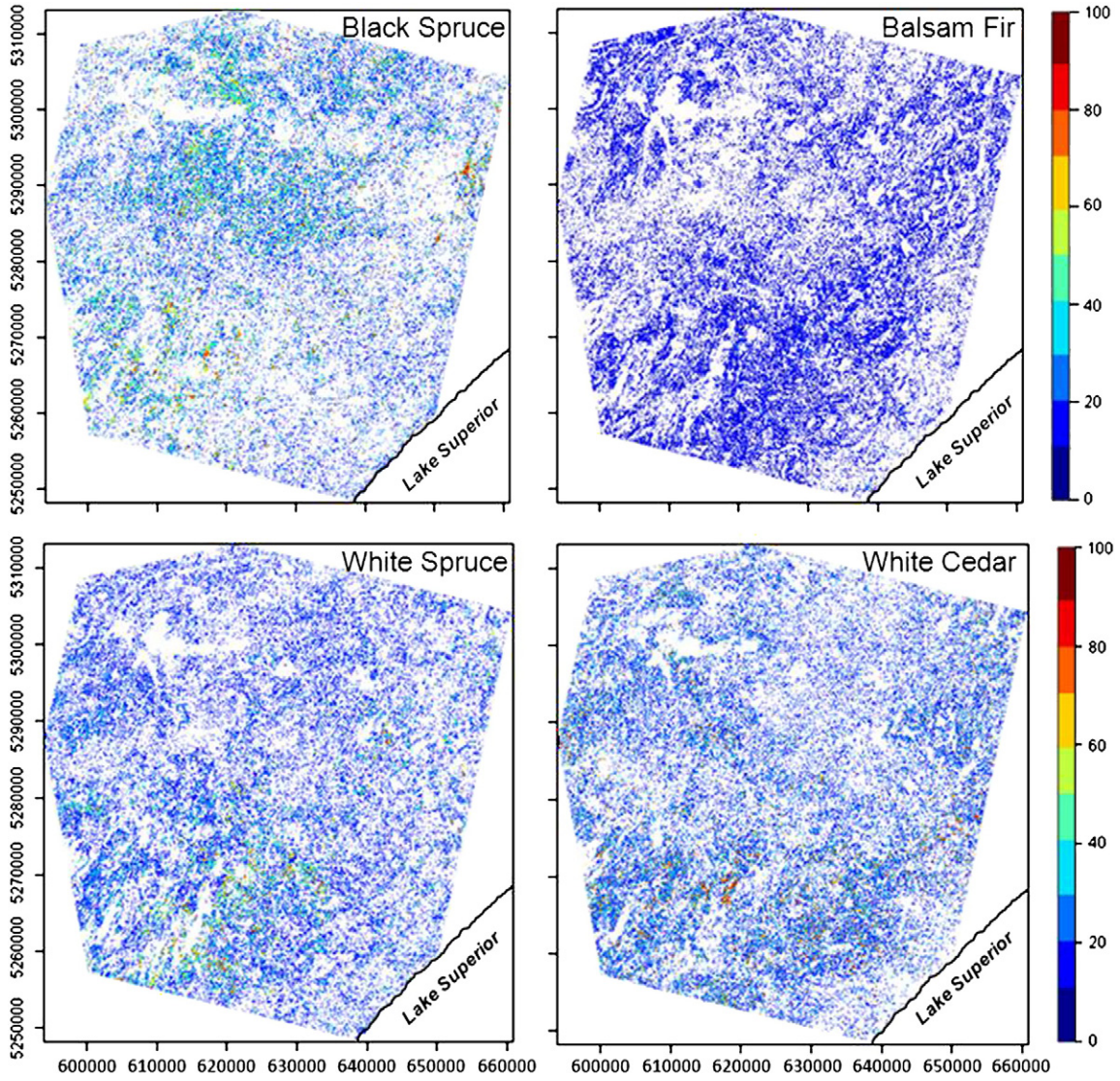

Fig. 10. Study-wide relative basal area (RBA) estimates for spruce, fir, and cedar. Black spruce distribution is greatest within the peatlands to the southwest and in the BWCA wilderness to the north. White spruce is more ubiquitous with highest RBA in the south. While balsam fir is ubiquitous, it rarely has RBA values above 20%. On the other hand, white cedar shows distinct patches of high RBA along the southern portion of the study area. Coordinates on the map axes are UTM zone 15 in meters.

ash and negatively loaded on paper birch. Black ash stands in this region of Minnesota are almost exclusively lowland species where copious areas of standing water are common. Though black ash stands were not totally flooded when PALSAR imaged this area (Table 2), it is conceivable that any standing water in these stands could enhance double bounce characteristics (e.g. Townsend, 2002).

4.1.2. Synthetic aperture radar—conifers

Overall, loading magnitudes for Radarsat-1variables were slightly stronger for conifers than hardwoods, while PALSAR variables were clearly stronger for conifers (Fig. 7), which was similar to previous studies (see Sader, 1987). For C-band, this is logical since conifers (except eastern larch) generally lack unique phenological cues that are typical of many hardwood species. Rather, conifers in this study have substantial differences in needle size, needle clumping, branching arrangement, and canopy architecture compared to their hardwood counterparts, which C-band SAR is particularly sensitive to (Dobson et al., 1992b; Ahern et al., 1993). As a result, the subtle optical differences between conifer species (e.g., white and red pine or white spruce and balsam fir) were complemented by SAR's sensitivity to structural differences as well as tree physiognomic differences (Rignot et al., 1994).

Among the three host tree species for spruce budworm (balsam fir and the two spruces), balsam and white spruce stands have both strong optical similarities and strong structural differences (Appendix A). Only one Radarsat-1 variable (+R6M, mean neighborhood amplitude 2/28/2007) showed relatively strong loading for balsam fir. Black and white spruce showed equally strong loadings on four (−R5S, +R6M, −R23, and +R26) and five (−R2S, +R4S, +R15, +R23, and −R26) of the Radarsat-1 variables, respectively (Fig. 7). Apparent sensitivity to structural and physiognomic differences, and not underlying soil moisture differences (black spruce stands generally occupy much wetter sites), may have also been facilitated, in part, by the presence of snow cover and frozen soil (Wegmuller, 1990; Pulliainen et al., 1999; Fransson et al., 2001). Radarsat-1 winter image variables and summer to winter ratio variables generally had stronger loadings for conifers (including fir and spruce) than summer C-band image variables (Fig. 7).

The Radarsat-1 variable R6M was the only neighborhood mean amplitude variable selected for use in the best combined sensor model (Table 4, Fig. 7). Moderate loadings on this variable for jack pine (−R6M), balsam fir (+R6M), black spruce (+R6M), and white cedar (+R6M) emphasizes the importance of including leaf-off SAR data when conifers occur in mixture with hardwoods. Balsam fir, the preferred host for spruce budworm in this region (Batzer, 1969), is a

prime example in that it frequently exists as an understory component in aspen-fir associations and rarely forms pure stands. Where fir is found in the overstory it usually shares the position with white spruce and aspen. In either case, the absence of hardwood foliage in the winter months, that would otherwise affect C-band backscatter (Rignot et al., 1994), affords a less obstructed view of this species in terms of optical satellite data (Wolter et al., 2008) and C-band SAR data.

Black spruce stands (least preferred spruce budworm host species in this region) have frequently been confused with jack pine stands in Landsat-based classifications (Shen et al., 1985; Hall et al., 1991; Wolter et al., 1995; Peddle et al., 2004). While black spruce typically forms pure stands on lowland sites, pure jack pine stands in this region and elsewhere (see Saatchi & Rignot, 1997) frequently occurs on upland sites where the ground is relatively smooth and dry compared black spruce sites. It is important to note that these species coexist on upland and lowland sites (more so farther north) which has limited the use of ancillary wetlands data for reconciling such confusion in this region (Wolter & White, 2002). Structurally, black spruce stands (with the strongest overall −PHV loading) were shorter than jack pines (mean HT = 12.8 m versus 15.1 m), had smaller diameter trunks and crowns (mean DBH = 16.1 cm versus 22.8 cm and mean CDIA = 2.3 m versus 3.9 m), and had lower crown closure than jack pine stands (mean CC = 52% versus 85%) (Appendix A), characteristics known to effect HV backscatter (Sader, 1987). Saatchi and Rignot (1997) also noted greater height of jack pine over black spruce stands, but that foliar biomass of jack pine was lower than black spruce. We did not quantify foliar biomass, but is interesting that the L-band HH polarization variable was more strongly loaded for jack pine (−PHM) than black spruce stands (+PHM) given the known strong influences of these collective stand attributes on HH double bounce returns (Le Toan et al., 1992; Wang et al., 1993; Moghaddam & Saatchi, 1995). For other SAR variables, loadings for jack pine were generally of similar (but opposite) magnitude compared to L- and C-band loadings that were important for black spruce (Fig. 7).

L-band HV backscatter, on the other hand, is typically stronger over forests and more sensitive to differences in above ground biomass than HH backscatter (Sader, 1987; Le Toan et al., 1992; Rignot et al., 1994; Ranson & Sun, 1994; Imhoff, 1995; Wang et al., 1995; Almeida-Filho et al., 2007). In this study, HV cross polarization data were more strongly loaded for black spruce (−PVM), white cedar (+PVM), and white spruce (+PVM), respectively, than for any other hardwood or conifer species (Fig. 7). In terms of Landsat data, white cedar is optically similar to white spruce, red pine, and white pine (Wolter et al., 1995). However, white cedar is structurally unique in that very old trees tend to be relatively short and stout (e.g., mean height 11.9 m and DBH 28.7 cm) with vertical length of live crown (LC) values (mean = 7.2 m, σ = 2.9 m) similar to much taller pine trees. Cedar stands also had high basal area (range 44–80 m^2 ha^{-1}) with numerous trees that leaned 20–40° from vertical. This leaning tendency and high basal area may explain why within stand abundance of cedars was underestimated (Fig. 6) and why +PVM was a strong predictor over all other SAR variables (Fig. 7). Relatively strong loadings on two C-band neighborhood standard deviation variables (−R4S and +R5S) may also be indicative of the chaotic canopy structure observed within white cedar stands.

In contrast, white pine stands in our study area contained the largest conifers with tree heights often in excess of 26 m (max. = 31.3 m). However, models of both white pine and eastern larch generally relied less on SAR variables compared to the other conifers, but showed moderate dependence on one L-band variable (−PVM). For these species, unique structure identified among the SPOT-based variables (i.e., LC for white pine) or unique phenological cues (i.e., leaf-off eastern larch) captured in Landsat imagery (Table 3, Fig. 7) were strong compliments to the structural differences that C- and L-band SAR data are known to be sensitive to (Rignot et al., 1994).

With specific reference to spruce budworm host species, L-band cross polarization variables (PVM and PVS) were unique (Fig. 7). For PVM, the combination of relatively strong magnitudes with opposite loadings for black and white spruce distinguished these species from each other and balsam fir (lower −PVM magnitude), as well as from other conifer species. Perhaps more important, balsam fir had the strongest −PVS loading of any other forest species analyzed (Fig. 7). Neighborhood variability in HV backscatter at L-band for stands containing balsam fir may be linked to this specie's frequent subcanopy position under a hardwood overstory, as discussed above. Such stands where larger aspen (or aspen and white spruce) boles occur in combination with smaller, more densely spaced balsam fir boles, is distinctive in this region, but how this may relate to HV neighborhood backscatter variability remains unclear. Whether L-band signatures observed for spruces and balsam fir in this study are regionally robust is not known. However, the fact that balsam fir is detectable with any degree of acceptable accuracy under these conditions is remarkable since it had been common practice until recently (e.g., Wolter et al., 2008) to combine balsam fir with either black spruce (Beaubien, 1979), white spruce (Bauer et al., 1994; Wolter & White, 2002; Bauer et al., 2009), or more general conifer categories (Reese et al., 2002) to avoid optical discrimination errors among satellite-based forest classification efforts. Hence, the use of PALSAR data in combination with C-band SAR and optical remote sensing data is recommended for distinguishing spruce budworm host tree species in northeastern Minnesota.

Extraction of SAR amplitude statistics (C- and L-band), according to pixel-wise, optical, Euclidean neighborhoods (see Wolter et al., 2009), provides a context within which otherwise noisy SAR data (Dell'Acqua et al., 2006) can be used in conjunction with optical data (via PLS regression) to distinguish different species and facilitate estimation of their relative abundance. While results (Fig. 7) seem to suggest leaf-off Radarsat-1 data (three variables selected) are more useful than leaf-on data in this regard (two variables selected), ratios of summer to winter (leaf-on/off) Radarsat-1 data were clearly important for reconciling relative BA for several optically similar conifer species, especially spruce budworm host tree species (Fig. 7).

4.1.3. Spot-based forest structure

Maps of forest structure derived from SPOT imagery (Wolter et al., 2009) were moderately strong predictors of relative BA for species lacking unique and/or temporally divergent phenological signatures such as conifers, but also aspen and paper birch (Fig. 7). A salient example of this is white pine and red pine, which are optically similar in Landsat imagery. As a result, white pine has commonly been pooled with red pine or less specific categories (Bauer et al., 1994; Wolter et al., 1995; Wolter & White, 2002; Reese et al., 2002; Bauer et al., 2009) to reduce error since white pine is generally less abundant than red pine in this region (Nichols, 1935; Moore & Bauer, 1990). Structurally, however, white pine trees are unique in that they attain greater height (mean HT = 25.8 m), have broader (mean CDIA = 7.1 m), deeper canopies (mean LC = 13.7 m), and larger boles (mean DBH = 46.9 cm) compared to the other conifer stands measured (Appendix A). White pine is known to rapidly attain dominant canopy position (Frothingham, 1914), often towering over associated tree species (Nichols, 1935) as super-dominants. Given this capacity, one would have expected the SPOT-based height (HT) variable to be a stronger predictor of white pine relative BA. Rather, the SPOT-based length of live crown (LC) variable eclipsed all other image variables in this capacity (Fig. 7).

On average, white pine live crowns were 5 to 7 m longer than that of other conifer species measured in this study (Appendix 1) regardless of canopy position. Whether these dimensions are unique to white pine is unclear. Published equations for determining leaf and live branch biomass allocations (Perala & Alban, 1994; Jenkins et al., 2003) are generally insufficient for determining species-specific LC trends. Moreover, while crown properties including LC and live crown ratio (HT/LC) are used to gauge growth of white pine saplings (O'Connell & Kelty, 1994) or calculate optimum stocking levels among mature stands (Bechtold, 2003; Jordan & Ducey, 2007), LC as uniquely large for white

pine is not mentioned. It is, therefore, uncertain whether SPOT structure data of this nature would produce similar results if applied across the range of white pine where it co-occurs with conifer species not found (or rare) in northern Minnesota, such as eastern hemlock (*Tsuga canadensis*). The ability to more accurately map white pine distribution and abundance offers substantial potential for ecological management to identify and preserve critical habitat for a variety of forest bird species (Green, 1992; Niemi et al., 1997; Schulte et al., 2005) and large mammals (Rogers & Lindquist, 1992).

White cedar stands also exhibited unique structural characteristics, having high basal area (mean 63.7 m^2 ha^{-1}, $\sigma = 17.2$ m^2 ha^{-1}), moderate trunk diameter (mean = 28.7 cm), but relatively short trees (mean = 11.9 m) (Appendix A). Thus, it is reasonable that SPOT-based measures of tree height (−HT) and basal area (−BA) were relatively strong predictors of abundance, followed by moderate loading on + CDIA and −LC. Likewise, black spruce stands (optically similar to jack pine, white spruce, balsam fir, and white cedar) are typically shorter and more open compared to other conifer species (Appendix A). SPOT-based + HT and + CC were moderate compliments to the higher ranked −PVM and Landsat variables for estimating black spruce relative abundance (Fig. 7). We should note that interpretation of SPOT-based structure estimate loadings for balsam fir is confounded by the fact that these estimates were based solely on summer (leaf-on) imagery (see Wolter et al., 2009). As previously discussed, balsam fir is frequently, but not always, an understory component of mixed aspen/fir/spruce stands in the study area. Therefore, SPOT-based structure estimates for such stands represent the dominant overstory species—typically combinations of aspen and white spruce.

Lower overall use of the SPOT-base structure data among hardwood species compared to the conifers was expected given the distinct, interspecific hardwood phenology cues in the study area (Wolter et al., 1995). However, several of the loading results on these structure data validated ground observations among hardwood stands. For instance, maple and paper birch (each posting moderate + CDIA loading) had similarly sized canopy diameters that were larger on average than that of either aspen or black ash (Appendix A). Average live crown lengths for paper birch (moderate −LC loading) were approximately 2 m greater than that of aspen (weak + LC loading), but similar to black ash (moderate + LC loading). While many of the SPOT-based structure loadings seem reasonable when compared with field measurements, several do not, and hence, are more difficult to explain, especially when evaluated in context with the full complement of SAR and Landsat variables (63 total). Nevertheless, the fact that five of the six SPOT-based structure variables (DBH excluded) emerged from the automatic variable selection routine (for the best combined sensor model) is a testament to the importance of such data for reconciling species differences as well as variability in abundance. This is particularly important where key diagnostic properties for a species (e.g., white pine LC) are not adequately captured, due to suboptimal spatial or spectral resolution, when using conventional satellite sensor data (e.g., Landsat).

4.1.4. Landsat predictor variables

The fact that SWIR-related variables (SWIR5, SWIR7, SVR, SVR7, MSI, and Tasseled Cap wetness) and greenness-related variables (NIR, SR, SAVI, GEMI, and Tasseled Cap greenness) comprised over 57% of the 63 image variables retained by the best PLS model (Figs. 4, 7) was not surprising. SWIR-based Landsat variables have been shown to be sensitive to forest basal area and density in other studies (Horler & Ahern, 1986; Ahern et al., 1991; Brockhaus & Khorram, 1992; Ardö, 1992; Franklin et al., 2000; Wolter et al., 2008), while the importance of greenness indices among forest studies is well documented. In particular, the Landsat-based SWIR/visible ratio (SVR) is known to be more sensitive to conifer and hardwood relative basal area than other SWIR-based variables (Wolter et al., 2008), while NDVI and similar ratios are especially useful for distinguishing unique tree species phenology (Wolter et al., 1995).

With regard to phenology, 20 out of the 46 Landsat variables retained by the best PLS model were from the autumn images (Table 2, Fig. 7). Image variables from early leaf-flush (6/9/1997) were also important for quantifying relative BA for deciduous tree species in this region (Fig. 7), as the spectral properties among many forest species are more distinct during early-stages of leaf development than later in the growing season (Badhwar et al., 1986; Miller et al., 1991; Schriever & Congalton, 1995). Several of the Landsat variables (e.g., SVR, SR, SAVI, MSI, and wetness) from the autumn images (Table 2) were stronger for paper birch and aspen than for maple. Because the SWIR-base variables (SVR, MSI, and wetness) are not known to be sensitive to autumn leaf pigmentation differences and because leaf-off timing was not a factor (these species were all leaf-on in September and Leaf-off in October), we suspect a combination of forest type-specific differences such as crown closure (maple 95% vs. birch and aspen 75–76%, Appendix A) coupled with visibility of understory vegetation (Badhwar et al., 1986; Spanner et al., 1990; Wolter et al., 2008) and autumn shadow geometry (Wolter et al., 2009) were likely factors that contributed to the observed loadings.

In addition to observed sensitivities of autumn SAVI and SR to aspen and paper birch abundance, the 9/24/2001 autumn index (AI_S) also emerged as a moderately strong predictor of maple (+AI_S) and paper birch (−AI_S) relative BA, with weaker AI loading for aspen. AI was designed specifically for use with Landsat data (J. Vogelmann pers. com.) to enhance the collective increases in carotinoid (orange to yellow coloring), anthocyanin (scarlet to red coloring), and xanthophyll (yellow coloring) pigmentation among senescent maple stands (*Acer rubrum* and *A. saccharum*) with respect to the relative insensitivity of visible blue to these autumn pigmentation differences. The AI loadings for maple and paper birch are logical because maple (sugar and red) leaf pigmentation is typically orange to red in this region (i.e., stronger AI) and senescent paper birch leaves (reaching peak color a few days after maple, Ahlgren, 1957) are yellow (Sayn-Wittgenstein, 1961). Aspen (quaking and bigtooth) leaves also turn yellow but are typically still green while maple and paper birch are at their peak autumn color (Eder, 1989). Visual assessment of hardwood stands in the September images confirmed that aspen foliage was, indeed, green on both dates. It should also be noted that black ash stands had shed leaves prior to theses dates explaining the weaker −AI_S loading for this species.

Explanation of the strong AI loadings for red pine (+AI_S, top ranked variable), jack pine (+AI_S), and white cedar (−AI_S, top ranked variable) is not as straight forward. Differences in visible red reflectance between conifer stands have been linked to leaf area index (Spanner et al., 1990) and basal area (Franklin, 1986), while differences in visible blue reflectance have been attributed to conifer canopy condition (Nelson et al., 1984). However, there are differences in visible spectral properties of conifer foliage (Pinard & Bannari, 2003) that can be substantially altered in autumn as second and third year needles become senescent (Santos et al., 2010). For red pine, jack pine, and cedar, the combination of autumn-specific variance in visible spectral properties (especially red) and species-specific structural differences (Appendix A) may explain elevated loading on AI_S, as well as the other strong autumn variable loadings for other conifer species (Fig. 7).

In addition to phenology, strong loadings on older June (1997) and October (1985) Landsat variables may also indicate use of gross spectral differences related to forest age. Clear-cutting is the primary mode of wood removal in this region (Wolter & White, 2002), which allows forest age to be easily tracked using time-series Landsat imagery (James et al., in revision; Huang et al., 2009; Pastor et al., 2005). It was apparent from inspection of these older images (particularly 1985) that many relatively mature forest stands observed in this study were clearly distinguishable as even-aged regeneration ca. 20 years earlier. Thus, abundance estimation for species almost exclusively under even-aged management (e.g., jack pine, red pine, and aspen) were readily distinguished from slower growing, spectrally similar species that are rarely managed this way (e.g., black spruce, white cedar, and maple).

As anticipated, 3 March (R) 2008 (leaf-off with snow cover) and 4 May (M) 2007 (leaf-off) Landsat images were strong predictors for estimating total BA as well as relative BA for aspen, paper birch, black ash, and most conifers (Fig. 7). As noted for SAR, the combination of the leaf-off image data (one with snow cover and one without) and leaf-on imagery is especially important where conifers (especially balsam fir) and hardwoods grow in association with each other (Sayn-Wittgenstein, 1961; Wolter et al., 2008). Leaf-off optical imagery allows a clearer view of both understory and overstory conifers in mixed stands while snow cover provides a uniformly bright background that enhances tree canopies and their shadows (Seely, 1949) as well as spectral factors related to forest density and age (Horler & Ahern, 1986). Moreover, the snow cover effectively obscures spectrally variable forest floor targets that can confound species-specific structural signatures (Brown et al., 2000; Chen & Cihlar, 1996; White et al., 1995).

4.2. Forest composition estimation and mapping

4.2.1. Species distributions

The spatial distribution and abundance of forest species within the study area highlight some obvious and interesting patterns (Figs. 8–10). While aspen and birch are both ubiquitous on this landscape, there are some clear differences. At higher relative BA values the two species generally occupy the same locations on the landscape, running east and west through the middle latitudes and much of the northeast. Paper birch is generally less abundant through most of this area, but tends to dominate the region between the maximum dominance of maple on the ridges of the north shore highlands and the Lake Superior shore to the southeast (Fig. 8). On the other hand, black ash is generally of low abundance across the entire study area, but is slightly more prevalent from northwest to southeast, where several large, pure stands are located.

Eastern larch is also of low abundance throughout the study area with locally vast stands occurring in the southwest (Fig. 9), which is part of the Tamarack Lowlands Ecological Subsection (Minnesota Department of Natural Resources, 1999). Among the pines, white pine is much more prevalent in the northern half of the study area where much of the landscape is within protected wilderness. However, there is a small region of high relative BA white pine near Isabella, MN at the center of the study area. This area also contains a high dominance of red pine in addition to a few high relative BA red pine patches to the south and southeast (Fig. 9). Jack pine occurs over large areas (mostly in south central and the northeast) where it represents 20 to 40% of the total basal area. Interestingly, balsam fir is ubiquitous on the landscape, but never at high relative BA (usually <10–20%). This is consistent with field observations showing that balsam fir rarely dominates the overstory in stands where is occurs (see Wolter et al., 2008).

In contrast to balsam fir, black spruce, white spruce, and white cedar each occupies very distinct positions on the landscape (Fig. 10). While black spruce attains greatest relative BA in the southwest peatlands region and in the north as an associate with jack pine, white spruce is most dominant in the south central and western portions of the study area. On the other hand, white cedar, while fairly ubiquitous at low relative BA, frequently occurs in large pure patches in the southern portion of the study area.

Tree species distribution and abundance estimates such as these are invaluable for modeling disturbance dynamics at the large landscape scale (James et al. in revision). Detailed spatial information about the abundance of spruces and balsam fir is a benefit to the study of spruce budworm dynamics in this region, whereas being able to distinguish pines, especially jack pine, from spruce and fir benefits the study of both spruce budworm and jack pine budworm (*Choristoneura pinus*) by providing increased spatial precision among species distribution estimates that were previously lacking in studies of disturbance dynamics. The discrimination of aspen from birch is crucial for the study of forest tent caterpillar dynamics, as remote determination of the spatial distributions of these two ubiquitous hardwood types has generally been unsatisfactory in the past (Moore & Bauer, 1990; Wolter et al., 1995; Franco-Lopez et al., 2001; Wolter & White, 2002). While we demonstrate the utility of including pixel-wise forest structure information for estimating species-wise relative BA, we suspect satisfactory results may be attained using SAR and Landsat data exclusively (Fig. 4).

4.2.2. Dominant forest type mapping and accuracy

While our species-specific estimates of forest relative BA are very useful as separate information layers, it is commonplace to attempt to produce a single layer that encompasses information on dominant cover types (Moore & Bauer, 1990; Schriever & Congalton, 1995; Wolter et al., 1995; Reese et al., 2002; Bauer et al., 2009). Many possible strategies may be used to accomplish this, but we chose to compile simple dominant forest species using a method that ignores species with slightly lower abundance values. Approximately one third (67) of the 205 ground data plots used for assessing the accuracy of the dominant species map were designed specifically for integration with 30 m satellite data (see Fig. 2A). The remaining plot data were not designed for such use, and consisted of three rather than five subplots per plot separate by 50 m rather than 30 m (Fig. 2B). Such differences in subplot arrangement and number could have resulted in disparities when determining a single dominant species. Also, there were concerns over the positional accuracy of these supplemental ground data points as their GPS locations were not differentially corrected (standard positioning service horizontal error ≤ 15 m). Thus, the overall accuracy of 78% (Khat = 0.75) could be substantially understated. In any event, based on forest classification work the author performed in previous studies (Wolter & White, 2002), the resulting dominance map produced here (Fig. 11) appears to be more accurate than past efforts using Landsat data alone.

It should also be noted that while ROC curve analysis performed on individual species was an excellent way to reliably determine presence or absence of a species for a given pixel, it does provide information on species co-occurrence or dominance. Thus, alternative approaches to accuracy assessment such as using fuzzy sets (Gopal & Woodcock, 1994; Foody, 1995; Townsend, 2000) may be warranted (Table 8).

5. Conclusions

In addition to being the first demonstration of the use of PALSAR data with other imagery to map forest species distribution and abundance, five major conclusions emerged from this study. First, PLS regression is a powerful tool for integrating or fusing satellite data from different

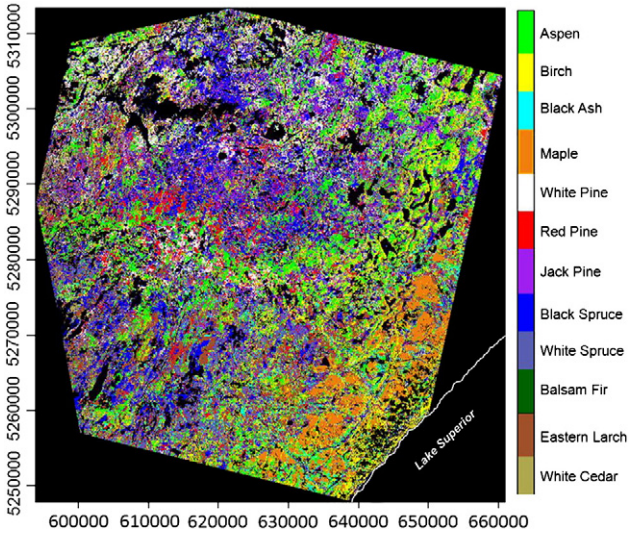

Fig. 11. Forest cover type map derived from the species-wise relative basal area (RBA) estimates for the study area. Though this map depicts pixel-wise species dominance, the RBA estimates for individual forest species can be combined in many different ways depending on specific needs. Coordinates on the map axes are UTM zone 15 in meters.

Table 8
Accuracy assessment for the dominant species cover type map of the study area that was derived from species-wise relative basal area estimates. See Table 4 for descriptions of the abbreviated forest species row and column headings.

		Ground Reference Data													
		ASP	BIR	ASH	MPL	WP	RP	JP	BS	WS	FIR	WC	EL	Row Total	User Acc.
Classified Data	ASP	20	3	2										25	0.80
	BIR	1	9		1									11	0.82
	ASH		1	11										12	0.92
	MPL		1	1	13									15	0.87
	WP	3				9	2	1			2			17	0.53
	RP					4	10		1	3				18	0.56
	JP							12		1				13	0.92
	BS							3	35	4		1	1	44	0.80
	WS								2	21				23	0.91
	FIR								1		4			5	0.80
	WC									2		8		10	0.80
	EL								3	2			7	12	0.58
Col. Total		24	14	14	14	13	12	16	42	33	6	9	8	205	
Prod. Acc.		0.83	0.64	0.79	0.93	0.69	0.83	0.75	0.83	0.64	0.67	0.89	0.88	Diag. Total	159
														Overall Accuracy	0.78
														Khat	0.75

sensors for the purpose of estimating forest composition. Second, the use of SPOT-based spectral neighborhoods for extracting C- and L-band amplitude statistic from inherently noisy radar data were invaluable for distinguishing several forest species. SAR's unique sensitivity to species-specific structural differences (branching, bole shape, or leaf arrangement) among otherwise optically similar forest species was the key to discriminating many forest species and estimating respective abundance, especially for spruce budworm and forest tent caterpillar host species. Third, pixel-wise forest structure information derived from 5 and 10 m SPOT-5 satellite data was especially important for estimating conifer forest composition (particularly white pine and white cedar). We suspect phenological differences among the hardwood species eclipsed subtle structural differences, whereas conifer species have stronger structural differences and only slight phenology differences (except eastern larch). Fourth, receiver operating characteristic (ROC) curve analysis was very useful for identifying the optimal lower limits among relative BA estimates for determining species presence or absence. However, the modest overall accuracy (78% correct, Khat 0.75) reported for the dominant species map is likely understated due to positional uncertainty and suboptimal design of a portion (ca. 66%) of the validation data set used, while our simple methods for determining dominant species among both ground and satellite-based relative BA estimates ignored secondary species with only slightly lower abundance. In the future, more elaborate iterative approaches to accuracy assessment of these dynamic data such as using fuzzy sets are probably more appropriate. Finally, partial least squares regression and a multi-sensor approach, when combined with both composition and structure data, provide a unique, powerful approach to mapping forest type and species abundance for regions that are both spatially and compositionally heterogeneous.

Acknowledgments

This research was supported by funding from the U.S. Forest Service via the National Fire Plan (Project 01.NCS.C.2) and a grant from the USDA Cooperative State Research, Education and Extension Service (CSREES) Managed Ecosystems program (2005-35101-16342), as well as fellowship support from the Department of Forest and Wildlife Ecology at the University of Wisconsin. Finally, thanks to Peter Crump and Aditya Singh for programming support, and to the Natural Resources Research Institute of the University of Minnesota—Duluth for providing supplemental forest plot data and for transportation assistance.

Appendix A

Mean and standard deviation for stand height (HT), bole diameter at breast height (DBH), canopy diameter (CDIA), vertical length of live crown (LC), canopy closure (CC), and total basal area (BA) for the 12 forest species studied. See Table 4 for descriptions of the abbreviated forest species column headings.

		ASP	BIR	MPL	ASH	WP	RP	JP	WC	FIR	WS	BS	EL
HT (m)	Mean	16.86	16.19	18.46	13.63	25.81	19.84	15.11	11.94	12.34	14.82	12.75	13.60
	Std. dev	3.57	3.16	4.36	2.15	3.80	4.14	3.96	2.80	1.98	3.23	2.93	1.77
DBH (cm)	Mean	20.29	21.57	28.18	20.67	46.93	28.61	22.77	28.73	12.86	22.22	16.10	15.88
	Std. dev	6.88	6.64	8.68	4.43	13.68	9.24	6.14	9.26	1.87	4.89	4.23	3.52
CDIA (m)	Mean	4.40	5.17	5.31	3.91	7.08	3.86	3.85	3.19	3.42	4.10	2.32	2.69
	Std. dev	0.98	0.81	1.62	0.50	1.71	1.32	0.61	0.72	0.43	0.76	0.47	0.68
LC (m)	Mean	5.42	7.69	8.60	7.02	13.73	7.29	6.32	7.18	8.36	8.54	7.53	7.88
	Std. dev	1.57	2.22	2.45	1.94	3.37	2.14	1.16	2.90	1.78	2.50	2.41	1.91
CC (%)	Mean	0.76	0.75	0.95	0.76	0.76	0.78	0.85	0.78	0.26	0.82	0.52	0.63
	Std. dev	0.08	0.13	0.05	0.08	0.08	0.10	0.01	0.08	0.18	0.11	0.20	0.13
BA (m^2 ha^{-1})	Mean	29.31	28.20	35.20	32.12	34.80	42.70	33.20	63.65	5.36	37.60	30.20	26.07
	Std. dev	7.21	4.34	4.17	4.63	6.26	11.84	3.17	12.74	2.09	5.46	14.73	4.67

References

Aber, J. D., Ollinger, S. V., Federer, C. A., Reich, P. B., Goulden, M. L., Kicklighter, D. W., et al. (1995). Predicting the effects of climate change on water yield and forest production in the northeastern United States. *Climate Research*, *5*, 207–222.

Ahern, F. J., Landry, R., Paterson, J. S., Boucher, D., & McKirdy, I. (1995). Forest landcover information content of multi-frequency multi-polarized SAR data of a boreal forest. *Proceedings of Seventeenth Canadian Symposium on Remote Sensing* (pp. 537–549). Saskatoon, Saskatchewan.

Ahern, F. J., Leckie, D. J., & Drieman, D. A. (1993). Seasonal changes in relative C-band backscatter of northern forest types. *IEEE Transactions on Geoscience and Remote Sensing*, *31*(3), 668–680.

Ahern, F. J., Skelly, W. C., McDonald, K., & Pearce, C. M. (1991). Studies of microwave sensitivity to black spruce forest biomass with the MIMICS model. *Proceedings of Fourteenth Canadian Symposium on Remote Sensing* (pp. 234–239). Calgary, Alberta.

Ahlgren, C. (1957). Phenological observations of nineteen native tree species in northern Minnesota. *Ecology*, *38*, 622–628.

Almeida-Filho, R., Rosenqvist, A., Shimabukuro, Y. E., & Silva-Gomez, R. (2007). Detecting deforestation with multitemporal L-band SAR imagery: A case study in western Brazilian Amazonia. *International Journal of Remote Sensing*, *28*(6), 1383–1390.

Amini, J., & Sumantyo, J. T. S. (2009). Employing a method on SAR and optical images for forest biomass estimation. *IEEE Transactions on Geoscience and Remote Sensing*, *47* (12), 4020–4026.

Ardö, J. (1992). Volume quantification of coniferous forest compartments using spectral radiance recorded by Landsat Thematic Mapper. *International Journal of Remote Sensing*, *13*(9), 1779–1786.

Badhwar, G. D., MacDonald, R. B., Hall, F. G., & Carnes, J. G. (1986). Spectral characterization of biophysical characteristics in a boreal forest: Relationship between thematic mapper band reflectance and leaf area index for aspen. *IEEE Transactions on Geoscience and Remote Sensing*, *GE-24*(3), 322–326.

Baker, W. L. (1989). Landscape ecology and nature reserve design in the Boundary Waters Canoe Area, Minnesota. *Ecology*, *70*(1), 23–35.

Batzer, H. O. (1969). Forest character and vulnerability of balsam fir to spruce budworm in Minnesota. *Forest Science*, *15*(1), 17–25.

Batzer, H. O., Martin, M. P., Mattson, W. J., & Miller, W. E. (1995). The forest tent caterpillar in aspen stands: Distribution and density estimation of four life stages in four vegetation strata. *Forest Science*, *41*(1), 99–121.

Bauer, M. E., Burk, T. E., Ek, A. R., Coppin, P. R., Lime, S. D., Walsh, T. A., et al. (1994). Satellite inventory of Minnesota forest resources. *Photogrammetric Engineering and Remote Sensing*, *60*(3), 287–298.

Bauer, M., Loeffelholz, B., & Shinneman, D. (2009). Border Lakes land-cover classification—Metadata, description of land-cover classification methods, fold-out map, and digital versions of classification. *U.S.D.A Forest Service, Northern Research Station, RMAP NRS-1* 14 pp.

Beaubien, J. (1979). Forest type mapping from Landsat digital data. *Photogrammetric Engineering and Remote Sensing*, *45*(8), 1135–1144.

Bechtold, W. A. (2003). Crown-diameter prediction models for 87 species of stand-grown trees in the eastern United States. *South Journal of Applied Forestry*, *27*(4), 269–278.

Betts, R. A., Cox, P. M., Lee, S. E., & Woodward, F. I. (1997). Contrasting physiological and structural vegetation feedbacks in climate change simulations. *Nature*, *387*, 796–799.

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.

Bolstad, P. V., & Lillesand, T. M. (1992). Improved classification of forest vegetation in northern Wisconsin through a rule-based combination of soils, terrain, and Landsat Thematic Mapper data. *Forest Science*, *38*(1), 5–20.

Bonan, G. B. (2008). Forests and climate change: Forcings, feedbacks, and the climate benefits of forests. *Science*, *320*, 1444–1449.

Bouchard, M., Kneeshaw, D., & Bergeron, Y. (2006). Forest dynamics after successive spruce budworm outbreaks in mixedwood forests. *Ecology*, *87*(9), 2319–2329.

Brockhaus, J. A., & Khorram, S. (1992). A comparison of SPOT and Landsat-TM data for use in conducting inventories of forest resources. *International Journal of Remote Sensing*, *13*(6), 3035–3043.

Brown, L., Chen, J. M., Leblanc, S. G., & Cihlar, J. (2000). A shortwave infrared modification to the simple ratio for LAI retrieval in boreal forests: An image and model analysis. *Remote Sensing of Environment*, *71*, 16–25.

Buell, M. F., & Niering, W. A. (1957). Fir-spruce-birch forest in northern Minnesota. *Ecology*, *38*(4), 602–610.

Campbell, N. A. (1993). Towards more quantitative extraction of information from remotely sensed data. *Proceedings Advanced Remote Sensing*, *2*. (pp. 29–40) Sydney, Australia.

Chen, J. M., Blanken, P. D., Black, T. A., Guilbeault, M., & Chen, S. (1997). Radiation regime and canopy architecture in a boreal aspen forest. *Agricultural and Forest Meteorology*, *86*, 107–125.

Chen, J. M., & Cihlar, J. (1996). Retrieving leaf area index of boreal conifer forests using Landsat TM images. *Remote Sensing of Environment*, *55*, 153–162.

Chopping, M., Moisen, G. G., Su, L., Laliberte, A., Rango, A., Martonchik, J. V., et al. (2008). Large area mapping of southwestern forest crown cover, canopy height, and biomass using the NASA Multiangle Imaging Spectro-Radiometer. *Remote Sensing of Environment*, *112*, 2051–2063.

Cohen, W. B., Maiersperger, T. K., Gower, S. T., & Turner, D. P. (2003). An improved strategy for regression of biophysical variables and Landsat ETM+ data. *Remote Sensing of Environment*, *84*, 561–571.

Congalton, R. G., & Green, K. (1999). *Assessing the accuracy of remotely sensed data: principles and practices*.Boca Raton: Lewis Publishers 133 pp.

Cooper, W. S. (1913). The climax forest of Isle Royale, Lake Superior, and its development. I. *Botanical Gazette*, *55*(1), 1–44.

Couturier, S., Gastellu-Etchegorry, J. -P., Patin˜o, P., & Martin, E. (2009). A model-based performance test for forest classifiers on remote-sensing imagery. *Forest Ecology and Management*, *257*, 23–37.

Crist, E. P., & Kauth, R. J. (1986). The tasseled cap de-mystified. *Photogrammetric Engineering and Remote Sensing*, *52*(1), 81–86.

Curran, P., & Hay, A. (1986). The importance of measurement error for certain procedures in remote sensing at optical wavelengths. *Photogrammetric Engineering and Remote Sensing*, *52*(2), 229–241.

Dell'Acqua, F., Gamba, P., & Trianni, G. (2006). Semi-automatic choice of scale-dependent features for satellite SAR image classification. *Pattern Recognition Letters*, *27*, 244–251.

Dobson, M. C., Pierce, L. E., Sarabandi, K., Ulaby, F. T., & Sharik, T. (1992). Preliminary analysis of ERS-1 SAR for forest ecosystem studies. *IEEE Transactions on Geoscience and Remote Sensing*, *30*, 203–211.

Dobson, M. C., Ulaby, F. T., LeToan, T., Beaudoin, A., Kasischke, E. S., & Christensen, N. (1992). Dependence of radar backscatter on conifer forest biomass. *IEEE Transactions on Geoscience and Remote Sensing*, *30*, 412–415.

Dobson, M. C., Ulaby, F. T., & Pierce, L. E. (1995). Land cover classification and estimation of terrain attributes using synthetic-aperture radar. *Remote Sensing of Environment*, *51*, 199–214.

Durden, S. L., Van Zyl, J. J., & Zebker, H. A. (1989). Modeling and observation of the radar polarization signature of forested areas. *IEEE Transactions on Geoscience and Remote Sensing*, *27*(3), 290–301.

Eder, J. J. (1989). Don't shoot unless it's autumn. *Journal of Forestry*, *87*(6), 50–51.

Ellenwood, J., & Krist, F. (2007). Building a nationwide 30-meter forest parameter dataset for forest health risk assessment. In M. (Ed.), *Forests and remote sensing: methods and operational tools. Proceedings of ForestSaT*, *200*, 5-7 November, Montpellier, France.

Engdahl, M. E., Pulliainen, J. T., & Hallikainen, M. T. (2004). Boreal forest coherence-based measures of interferometric pair suitability for operational stem volume retrieval. *IEEE Transactions on Geoscience and Remote Sensing*, *1*(3), 228–231.

Foody, G. M. (1995). Cross-entropy for the evaluation of the accuracy of a fuzzy land cover classification with fuzzy ground data. *ISPRS Journal of Remote Sensing*, *17*, 2–12.

Franco-Lopez, H., Ek, A. R., & Bauer, M. E. (2001). Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbors method. *Remote Sensing of Environment*, *77*, 251–274.

Franklin, J. (1986). Thematic Mapper analysis of coniferous forest structure and composition. *International Journal of Remote Sensing*, *7*(10), 1287–1301.

Franklin, S. E., Moskal, L. M., Lavinge, M. B., & Pugh, K. (2000). Interpretation and classification of partially harvested forest stands in the Fundy Model Forest using multitemporal Landsat TM data. *Canadian Journal of Remote Sensing*, *26*(4), 318–333.

Fransson, J. E. S., Smith, G., Askne, J., & Olsson, H. (2001). Stem volume estimation in boreal forests using ERS-1/2 coherence and SPOT XS optical data. *International Journal of Remote Sensing*, *22*(14), 2777–2791.

Frelich, L. E., & Reich, P. B. (1995). Spatial patterns and succession in a Minnesota southern-boreal forest. *Ecological Monographs*, *65*, 325–346.

Friedman, S. K., & Reich, P. B. (2005). Regional legacies of logging: Departure from presettlement forest conditions in northern Minnesota. *Ecological Applications*, *15* (2), 726–744.

Frothingham, E. H. (1914). White pine under forest management. *U.S. Department of Agriculture Bulletin*, *13*, 70 pp.

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.

Gopal, S., & Woodcock, C. E. (1994). Theory and methods for accuracy assessment of Thematic Maps using fuzzy sets. *Photogrammetric Engineering and Remote Sensing*, *60*, 181–188.

Green, J. C. (1992). Ecological features of white pine stands for wildlife. In R. A., & M. J. (Eds.), *Proceedings of White Pine Symposium: History, Ecology, Policy and Management* (pp. 44–53). Duluth, MN: Department of Forestry Resources, University of Minnesota, St. Paul 16-18 September,.

Grosenbaugh, L. R. (1952). Plotless timber estimates—New, fast, easy. *Journal of Forestry*, *50*, 322–337.

Hall, F. G., Botkin, D. B., Strebel, D. E., Woods, K. D., & Goetz, S. J. (1991). Large-scale patterns of forest succession as determined by remote sensing. *Ecology*, *72*(2), 628–640.

Hallikainen, M. T., Tares, T., Hyyppa, J., Somersalo, E., Ahola, P., Toikka, M., et al. (1990). Helicopter-borne measurement of radar backscatter from forests. *International Journal of Remote Sensing*, *11*(7), 1179–1191.

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, *143*, 29–36.

Hanowski, J., Danz, N., Lind, J., & Niemi, G. (2005). Breeding bird response to varying amounts of basal area retention in riparian forests. *Journal of Wildlife Management*, *62*(2), 689–698.

He, H. S., Mladenoff, D. J., & Gustafson, E. J. (2002). Study of landscape change under forest harvesting and climate warming-induced fire disturbance. *Forest Ecology and Management*, *155*, 257–270.

Heinselman, M. L. (1973). Fire in the virgin forests of the Boundary Waters Canoe Area, Minnesota. *Quaternary Research*, *3*, 329–382.

Holmgren, P., & Thuresson, T. (1998). Satellite remote sensing for forestry planning—A review. *Scandinavian Journal of Forest Research*, *13*, 90–110.

Horler, D. N. H., & Ahern, F. J. (1986). Forestry information content of thematic mapper data. *International Journal of Remote Sensing*, *7*, 405–428.

Huang, C., Goward, S. N., Schleeweis, K., Thomas, N., Masek, J. G., & Zhu, Z. (2009). Dynamics of national forests assessed using the Landsat record: Case studies in eastern United States. *Remote Sensing of Environment*, *113*, 1430−1442.

Huang, C., Yang, L., Wylie, B., & Homer, C. A. (2001). Strategy for estimating tree canopy density using Landsat7 ETM+and high resolution images over large areas. *Proceedings of Third International Conference on Geospatial Information in Agriculture and Forestry* Denver, Colorado, CD-ROM.

Huete, A. R. (1988). A soil adjusted vegetation index (SAVI). *Remote Sensing of Environment*, *25*, 295−309.

Hyde, P., Dubayah, R., Walker, W., Blair, J. B., Hofton, M., & Hunsaker, C. (2006). Mapping forest structure for wildlife habitat analysis using multi-sensor (Lidar, SAR/InSAR, ETM+, Quickbird) synergy. *Remote Sensing of Environment*, *102*, 63−73.

Imhoff, M. L. (1995). Radar backscatter and biomass saturation: ramifications for global biomass inventory. *IEEE Transactions on Geoscience and Remote Sensing*, *33*, 511−518.

James, P.M.A., Sturtevant, B.R., Townsend, P.A., Wolter, P.T., &. Fortin, M-J. (in revision). Spatial legacies of forest management in spruce budworm host tree species in the Border Lakes Landscape. Ecological Applications.

Jenkins, J. C., Chojnacky, D. C., Heath, L. S., & Birdsey, R. A. (2003). National-scale biomass estimators for United States tree species. *Forest Science*, *49*(1), 12−35.

Jensen, J. R. (2005). *Introductory digital image processing. A remote sensing perspective.* New Jersey, USA: Prentice Hall 526 pp.

Jordan, C. F. (1969). Derivation of leaf area index from quality of light on the forest floor. *Ecology*, *50*, 663−666.

Jordan, G. J., & Ducey, M. J. (2007). Predicting crown radius in eastern white pine (*Pinus strobus* L.) stands in New Hampshire. *Northern Journal of Applied Forestry*, *24*(1), 61−64.

Kellndorfer, J. M., Dobson, M. C., & Ulaby, F. T. (1997). Multi-ecoregion vegetation mapping using combined ERS/JERS SAR imagery. *Proceedings of Third ERS Symposium on Space at the service of our Environment* (pp. 1735−1740). 14-21 March, Florence, Italy.

Kerekes, J. (2008). Receiver operating characteristic curve confidence intervals and regions. *IEEE Geoscience and Remote Sensing Letters*, *5*(2), 251−255.

Koskinen, J. T., Pulliainen, J. T., Hyyppa, J. M., Engdahl, M. E., & Hallikainen, M. T. (2001). The seasonal behavior of interferometric coherence in boreal forest. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(4), 820−829.

Kwok, R., Rignot, E., Way, J. B., Freeman, A., & Holt, J. (1994). Polarization signatures of thawed and frozen forests of varying biomass. *IEEE Transactions on Geoscience and Remote Sensing*, *32*, 371−381.

Le Toan, T., Beaudoin, A., Riom, J., & Guyon, D. (1992). Relating forest biomass to SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, *30*, 403−411.

Leckie, D. G. (1990). Synergism of synthetic aperture radar and visible/infrared data for forest type discrimination. *Photogrammetric Engineering and Remote Sensing*, *56*(9), 1237−1246.

Leckie, D. G., Tinis, S., Nelson, T., Burnett, C., Gougeon, F. A., Cloney, E., et al. (2005). Issues in species classification of trees in old growth conifer stands. *Canadian Journal of Remote Sensing*, *31*(2), 175−190.

Lee, J. (1981). Speckle analysis and smoothing of synthetic aperture radar images. *Computer Graphics and Image Processing*, *17*(1), 24−32.

Lim, H. H., Swartz, A. A., Yueh, H. A., Kong, J. A., Shin, R. T., & van Zyl, J. (1989). Classification of earth terrain using polarimetric synthetic aperture radar images. *Journal of Geophysical Research*, *94*(B6), 7049−7057.

Lin, R. (1967). Review of the electrical properties of wood and cellulose. *Forestry Products Journal*, *17*, 54−61.

Lopes, A., Nezry, E., Touzi, R., & Laur, H. (1993). Structure detection and statistical adaptive speckle filtering in SAR images. *International Journal of Remote Sensing*, *14*(9), 1735−1758.

MacArthur, R. H., & MacArthur, J. W. (1961). On bird species diversity. *Ecology*, *42*(3), 594−598.

McRoberts, R. E. (2008). Using satellite imagery and the k-nearest neighbors technique as a bridge between strategic and management forest inventories. *Remote Sensing of Environment*, *112*, 2212−2221.

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*, 532−545.

McRoberts, R. E., Tomppo, E. O., Finley, A. O., & Heikkinen, J. (2007). Estimating areal means and variances of forest attributes using the k-Nearest Neighbors technique and satellite imagery. *Remote Sensing of Environment*, *111*, 466−480.

Meng, Q., Cieszewski, C. J., Madden, M., & Borders, B. E. (2009). K Nearest Neighbor method for forest inventory using remote sensing data. *GIScience & Remote Sensing*, *44*(2), 149−165.

Mickelson, J. G., Civco, D. L., & Silander, J. A. (1998). Delineating forest canopy species in the northeastern United States using multi-temporal TM imagery. *Photogrammetric Engineering and Remote Sensing*, *64*(9), 891−904.

Miller, J. R., Boyer, M. G., Belanger, M., Hare, E. W., & Wu, J. (1991). Seasonal patterns in leaf reflectance red-edge characteristics. *International Journal of Remote Sensing*, *12*, 1509−1523.

Minnesota Department of Natural Resources (1999). *Ecological Subsections of Minnesota (Map).* St. Paul, MN, USA: Minnesota Department of Natural Resources.

Moen, R., Burdett, C. L., & Niemi, G. J. (2008). Movement and habitat use of Canada Lynx during denning in Minnesota. *Journal of Wildlife Management*, *72*(7), 1507−1513.

Moghaddam, M., & Saatchi, S. S. (1995). Analysis of scattering mechanism in SAR imagery over boreal forest: results from BOREAS'93. *IEEE Transactions on Geoscience and Remote Sensing*, *33*, 1290−1296.

Moisen, G. G., Freeman, E. A., Blackard, J. A., Frescino, T. S., Zimmermann, N. E., & Edwards, T. C. (2006). Predicting tree species presence and basal area in Utah: A comparison of stochastic gradient boosting, generalized additive models, and tree-based methods. *Ecological Modelling*, *199*, 176−187.

Moore, M. M., & Bauer, M. E. (1990). Classification of forest vegetation in north-central Minnesota using Landsat Multispectral Scanner and Thematic Mapper data. *Forest Science*, *36*(2), 330−342.

Nelson, R. F., Latty, R. S., & Mott, G. (1984). Classifying northern forests using Thematic Mapper simulator data. *Photogrammetric Engineering and Remote Sensing*, *50*(5), 607−617.

Nichols, G. E. (1935). The hemlock-white pine-northern hardwood region of eastern North America. *Ecology*, *16*(3), 403−422.

Nielsen, A. A. (2002). Multiset canonical correlations analysis and multispectral, truly multitemporal remote sensing data. *IEEE Transactions on Image Processing*, *11*, 293−305.

Niemi, G. J., Hanowski, J. M., Lima, A. R., Nicholls, T., & Weiland, N. (1997). A critical analysis on the use of indicator species in management. *Journal of Wildlife Management*, *61*, 1240−1252.

O'Connell, B. M., & Kelty, M. (1994). Crown architecture of understory and open-grown white pine (*Pinus strobus* L.) saplings. *Tree Physiology*, *14*, 89−102.

Pastor, J., Cohen, Y., & Moen, R. (1999). Generation of spatial pattern in boreal forest landscapes. *Ecosystems*, *2*, 439−450.

Pastor, J., & Post, W. M. (1986). Influences of climate, soil moisture, and succession on forest carbon and nitrogen cycles. *Biogeochemistry*, *2*, 3−27.

Pastor, J., Sharp, A., & Wolter, P. (2005). An application of Markov models to the dynamics of Minnesota's forests. *Canadian Journal of Forest Research*, *35*, 3011−3019.

Peddle, D. R., Johnson, R. L., Cihlar, J., & Latifovic, R. (2004). Large area forest classification and biophysical parameter estimation using the 5-Scale canopy reflectance model in Multiple-Forward-Mode. *Remote Sensing of Environment*, *89*, 252−263.

Peet, R. K., Wentworth, T. R., & White, P. S. (1998). A flexible, multipurpose method for recording vegetation composition and structure. *Castanea*, *63*(3), 262−274.

Perala, D. A., & Alban, D. H. (1994). Allometric biomass estimators for aspen-dominated ecosystems in the Upper Great Lakes. *U.S. Forest Service Research Paper NC-134* (pp. 38).

Percy, K. E., Awmack, C. S., Lindroth, R. L., Kubiske, M. E., Kopper, B. J., Isebrands, J. G., et al. (2002). Altered performance of forest pests under atmospheres enriched by CO_2 and O_3. *Nature*, *420*, 403−407.

Peters, R. L., & Darling, J. D. S. (1985). The greenhouse effect and natural reserves. *Bioscience*, *35*(11), 707−711.

Pinard, V., & Bannari, A. (2003). Spectroradiometric analysis in a hyperspectral use perspective to discriminate between forest species. *Proceedings of IEEE Transactions on Geoscience and Remote Sensing Symposium*, *7*. (pp. 4301−4303) 21–25 July, Toulouse, France.

Pinty, B., & Verstraete, M. M. (1992). GEMI: A non-linear index to monitor global vegetation from satellites. *Vegetatio*, *101*, 15−20.

Pohl, C., & Van Genderen, J. L. (1998). Multisensor image fusion in remote sensing: concepts, methods and applications. *International Journal of Remote Sensing*, *19*(5), 823−854.

Pope, K. O., Rey-Benayaas, J. M., & Paris, J. F. (1994). Radar remote sensing of forest and wetland ecosystems in the Central American tropics. *Remote Sensing of Environment*, *48*, 205−219.

Pulliainen, J., Engdahl, M., & Hallikainen, M. (2003). Feasibility of multi-temporal interferometric SAR data for stand-level estimation of boreal forest stem volume. *Remote Sensing of Environment*, *85*, 397−409.

Pulliainen, J. T., Kurvonen, L., & Hallikainen, M. T. (1999). Multitemporal behavior of L- and C-band SAR observations of boreal forests. *IEEE Transactions on Geoscience and Remote Sensing*, *37*(2), 927−937.

Radeloff, V. C., Mladenoff, D. J., & Boyce, M. S. (2000). The changing relation of landscape patterns and jack pine budworm populations during an outbreak. *Oikos*, *90*, 417−430.

Ranson, K. J., & Saatchi, S. S. (1992). C-band microwave scattering from small balsam fir. *IEEE Transactions on Geoscience and Remote Sensing*, *30*, 924−932.

Ranson, K. J., & Sun, G. (1994). Northern forest classification using temporal multifrequency and multipolarimetric SAR images. *Remote Sensing of Environment*, *47*(2), 142−153.

Ranson, K. J., & Sun, G. (2000). Effects of environmental conditions on boreal forest classification and biomass estimates with SAR. *IEEE Transactions on Geoscience and Remote Sensing*, *38*(3), 1242−1252.

Ranson, K. J., Sun, G., Kharuk, V. I., & Kovacs, K. (2001). Characterization of forests in western Sayani Mountains, Siberia from SIR-C SAR data. *Remote Sensing of Environment*, *75*, 188−200.

Ranson, K. J., Sun, G., Lang, R. H., Chauhan, N. S., Cacciola, R. J., & Kilic, O. (1997). Mapping of boreal forest biomass from spaceborne synthetic aperture radar. *Journal of Geophysical Research*, *102*(D24), 29, 599−29, 610.

Rauste, Y. (2005). Multi-temporal JERS SAR data in boreal forest biomass mapping. *Remote Sensing of Environment*, *97*, 263−275.

Reese, H. M., Lillesand, T., Nagel, D. E., Stewart, J. S., Goldmann, R. A., Simmons, T. E., et al. (2002). Statewide land cover derived from multiseasonal Landstat TM data—A retrospective of the WISCLAND project. *Remote Sensing of Environment*, *82*, 224−237.

Reich, P. B., Bakken, P., Carlson, D., Frelich, L. E., Friedman, S. K., & Grigal, D. F. (2001). Influence of logging, fire, and forest type on biodiversity and productivity in southern boreal forests. *Ecology*, *82*(10), 2731−2748.

Riaño, D., Chuvieco, E., Salas, J., & Aguado, I. (2003). Assessment of different topographic corrections in Landsat-TM data for mapping vegetation types. *IEEE Transactions on Geoscience and Remote Sensing*, *41*(5), 1056−1061.

Richards, J. A., Sun, G., & Simonett, D. S. (1987). L-band radar backscatter modeling of forest stands. *IEEE Transactions on Geoscience and Remote Sensing*, *GE-25*(4), 487−498.

Rignot, E. J. M., Williams, C. L., Way, J., & Viereck, L. A. (1994). Mapping of forest types in Alaskan boreal forests using SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *32*(5), 1051−1059.

Rock, B. N., Vogelmann, J. E., Williams, D. L., Vogelmann, A. F., & Hoshizaki, T. (1986). Remote detection of forest damage. *Bioscience*, *36*(7), 439−445.

Rogers, L. L., & Lindquist, E. L. (1992). Supercanopy white pine and wildlife. In R. A., & M. J. (Eds.), *Proceedings of the White Pine Symposium: History, Ecology, Policy and Management* (pp. 44−53). Duluth, MN: Department of Forestry Resources, University of Minnesota, St. Paul 16-18 September.

Rosenqvist, A., Shimada, M., Watanabe, M., & Yamauchi, K. (2006). The ALOS PALSAR Observation Strategy—A global mission concept. EUSAR 2006. *6th European Conference on Synthetic Aperture Radar* 16-18 May, Dresden, Germany, 4 pp.

Saatchi, S. S., & Rignot, E. (1997). Classification of boreal cover types using SAR images. *Remote Sensing of Environment*, *60*, 270−281.

Sader, S. A. (1987). Forest biomass, canopy structure, and species composition relationships with multipolarization L-band synthetic aperture radar data. *Photogrammetric Engineering and Remote Sensing*, *53*(2), 193−202.

Santoro, M., Askne, J., Smith, G., & Fransson, J. E. S. (2002). Stem volume retrieval in boreal forests from ERS-1/2 interferometry. *Remote Sensing of Environment*, *81*, 19−35.

Santoro, M., Fransson, J. E. S., Eriksson, L. E. B., Magnusson, M., Ulander, L. M. H., & Olsson, H. (2009). Signatures of ALOS PALSAR L-Band backscatter in Swedish forest. *IEEE Transactions on Geoscience and Remote Sensing*, *47*(12), 4001−4019.

Santos, M. J., Greenberg, J. A., & Ustin, S. L. (2010). Using hyperspectral remote sensing to detect and quantify southeastern pine senescence effects in red-cockaded woodpecker (*Picoides borealis*) habitat. *Remote Sensing of Environment*, *114*, 1242−1250.

Sayn-Wittgenstein, L. (1961). *Phenological aids to tree species identification on air photographs*.Technical Note No. 104, Forest Research Branch, Canada Department of Forestry, Ottawa 26 pp.

Scheller, R. M., Mladenoff, D. J., Crow, T. R., & Sickley, T. A. (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*, 396−411.

Schriever, J. R., & Congalton, R. G. (1995). Evaluating seasonal variability as an aid to cover-type mapping from Landsat Thematic Mapper in the Northeast. *Photogrammetric Engineering and Remote Sensing*, *61*(3), 321−327.

Schulte, L. A., Pidgeon, A. M., & Mladenoff, D. J. (2005). One hundred fifty years of change in forest bird breeding habitat: Estimates of species distributions. *Conservation Biology*, *19*(6), 1944−1956.

Seely, H. E. (1949). Air photography and its application to forestry. *Photogrammetric Engineering*, *15*(4), 548−554.

Sharma, R., Leckie, D., Hill, D., Crooks, B., Bhogal, A. S., Arbour, P., et al. (2005). Hypertemporal Radarsat SAR data of a forested terrain. *IEEE 2005 International Workshop, Analysis of Multi-Temporal Remote Sensing Images* (pp. 71−75). 16-18 May, Biloxi, Mississippi USA.

Shen, S. S., Badhwar, G. D., & Carnes, J. G. (1985). Separability of boreal forest species in the Lake Jennette area, Minnesota. *Photogrammetric Engineering*, *51*(11), 1775−1783.

Shimada, M. (2007). Coherence dependency of the PALSAR POLinSAR on forest in Japan and Amazon. *Proceedings IEEE International Symposium on Geoscience and Remote Sensing (IGARSS)* (pp. 2636−2639). 23-27 July, Barcelona, Spain.

Singhroy, V., & Saint-Jean, R. (1999). Effects of relief on the selection of RADARSAT-1 incidence angle for geological applications. *Canadian Journal of Remote Sensing*, *25*(3), 211−217.

Spanner, M. A., Pierce, L. L., Peterson, D. L., & Running, S. W. (1990). Remote sensing of temperate coniferous forest leaf area index. The influence of canopy closure, understory vegetation and background reflectance. *International Journal of Remote Sensing*, *11*(1), 95−111.

Stabach, J. A., Dabek, L., Jensen, R., & Wang, Y. Q. (2009). Discrimination of dominant forest types for Matschie's tree kangaroo conservation in Papua New Guinea using high-resolution remote sensing data. *International Journal of Remote Sensing*, *30*(2), 405−422.

Sturtevant, B. R., & Gustafson, E. J. (2004). Modeling disturbance and succession in forest landscapes using LANDIS: introduction. *Ecological Modelling*, *180*(1), 1−5.

Thome, K. J., Helder, D. L., Aaron, D., & Dewald, J. D. (2004). Landsat-5 TM and Landsat-7 ETM+ absolute radiometric calibration using the reflectance-based method. *IEEE Transactions on Geoscience and Remote Sensing*, *42*(12), 2777−2785.

Touzi, R., Landry, R., & Charbonneau, F. J. (2004). Forest type discrimination using calibrated C-band polarimetric SAR data. *Canadian Journal of Remote Sensing*, *30*(3), 543−551.

Townsend, P. A. (2000). A quantitative fuzzy approach to assess mapped vegetation classifications for ecological applications. *Remote Sensing of Environment*, *72*, 253−267.

Townsend, P. A. (2002). Estimating forest structure in wetlands using multi-temporal SAR. *Remote Sensing of Environment*, *79*, 288−304.

Treuhaft, R. N., Law, B. E., & Asner, G. P. (2004). Forest attributes from radar interferometric structure and its fusion with optical remote sensing. *Bioscience*, *54*(6), 561−571.

Turkar, V., & Rao, Y. S. (2008). Classification of polarimetric synthetic aperture radar images from SIR-C and ALOS PALSAR. *Proceedings: International Conference on Recent Advances in Microwave Theory and Application* (pp. 438−440). 21-24 November, Jaipur, India.

Verbyla, D. L. (2001). A test of detecting spring leaf flush within the Alaskan boreal forest using ERS-2 and Radarsat SAR data. *International Journal of Remote Sensing*, *22*(6), 1159−1165.

Vogelmann, J. E., Sohl, T. L., Campbell, P. E., & Shaw, D. M. (1998). Regional land cover characterization using Landsat thematic mapper data and ancillary data sources. *Environmental Monitoring and Assessment*, *51*, 415−428.

Wang, Y., Davis, F. W., & Melack, J. M. (1993). Simulated and observed backscatter at P-, L-, and C-bands from ponderosa pine stands. *IEEE Transactions on Geoscience and Remote Sensing*, *31*(4), 871−879.

Wang, Y., Hess, L. L., Filoso, S., & Melack, J. M. (1995). Understanding the radar backscattering from flooded and nonflooded Amazonian forests: Results from canopy backscatter modeling. *Remote Sensing of Environment*, *54*, 324−332.

Watanabe, M., Sato, M., Shimada, M., Ouchi, K., & Wang, H. P. (2007). Forest monitoring by using L-band coherence. *Proceedings Asian Pacific Microwave Conference* (pp. 2528−2531). 11-14 December, Bangkok, Thailand.

Watanabe, M., Shimada, M., Rosenqvist, A., Tadono, T., Matsuoka, M., Romshoo, S. A., et al. (2006). Forest structure dependency of the relation between L-band σ^o and biophysical parameters. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(11), 3154−3165.

Wegmuller, U. (1990). The effect of freezing and thawing on the microwave signatures of bare soil. *Remote Sensing of Environment*, *33*, 123−135.

White, H. P., Miller, J. P., Chen, J. M., & Peddle, D. R. (1995). Seasonal change in mean understory reflectance for boreal sites: preliminary results. *Proceedings of the 17th Canadian Symposium on Remote Sensing* (pp. 189−194). 13–15 June, Saskatoon, Saskatchewan, Canada.

Wold, H. (1966). Estimation of principal components and related models by iterative least squares. In P. R. Krishnaiah (Ed.), *Multivariate Analysis* (pp. 391−420). New York: Academic Press.

Wold, H. (1975). Soft modeling by latent variables; The non-linear iterative partial least squares approach. In J. (Ed.), *Perspectives in Probability and Statistics: Papers in Honour of M.S. Bartlett* (pp. 117−142). London: Academic Press.

Wold, S., Høy, M., Martens, H., Trygg, J., Westad, F., MacGregor, J., et al. (2009). The PLS model space revisited. *Journal of Chemometrics*, *23*(1–2), 67−68.

Wolter, P. T., Mladenoff, D. J., Host, G. E., & Crow, T. R. (1995). Improved forest classification in the northern Lake State using multi-temporal Landsat imagery. *Photogrammetric Engineering and Remote Sensing*, *61*(9), 1129−1143.

Wolter, P. T., Townsend, P. A., Kingdon, C. C., & Sturtevant, B. R. (2008). Remote sensing of the distribution and abundance of host species for spruce budworm in northern Minnesota and Ontario. *Remote Sensing of Environment*, *112*, 3971−3982.

Wolter, P. T., Townsend, P. A., & Sturtevant, B. R. (2009). Estimation of forest structural parameters using 5 and 10 meter SPOT-5 satellite data. *Remote Sensing of Environment*, *113*, 2019−2036.

Wolter, P. T., & White, M. A. (2002). Recent forest cover type transitions and landscape structural changes in northeast Minnesota. *Landscape Ecology*, *17*, 133−155.

Wu, S., & Sader, S. A. (1987). Multipolarization SAR data for surface feature delineation and forest vegetation characterization. *IEEE Transactions on Geoscience and Remote Sensing*, *GE-25*(1), 67−76.

Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer*, *3*, 32−35.