Remote Sensing of Environment 113 (2009) 532–545

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

A two-step nearest neighbors algorithm using satellite imagery for predicting forest structure within species composition classes

Ronald E. McRoberts

Northern Research Station, U.S. Forest Service, Saint Paul, Minnesota, USA

ARTICLE INFO

Article history:
Received 23 April 2008
Received in revised form 25 September 2008
Accepted 1 October 2008

Keywords:
Mulitnomial logistic regression
Discriminant analysis
Landsat
Model-based inference

ABSTRACT

Nearest neighbors techniques have been shown to be useful for predicting multiple forest attributes from forest inventory and Landsat satellite image data. However, in regions lacking good digital land cover information, nearest neighbors selected to predict continuous variables such as tree volume must be selected without regard to relevant categorical variables such as forest/non-forest. The result is that non-zero volume predictions may be obtained for pixels predicted to be non-forest, and volume predictions for pixels predicted to be forest may be erroneously small due to non-forest nearest neighbors. For users who wish to circumvent this discrepancy, a two-step algorithm is proposed in which the class of a relevant categorical variable such as land cover is predicted in the first step, and continuous variables such as volume are predicted in the second step subject to the constraint that all nearest neighbors must come from the predicted class of the categorical variable. Nearest neighbors, multinomial logistic regression, and discriminant analysis techniques were investigated for use in the first step. The results were generally similar for the three techniques, although the multinomial logistic regression technique was slightly superior. The k-Nearest Neighbors technique was used in the second step because many continuous forest inventory variables do not satisfy the distributional assumptions necessary for parametric multivariate techniques. The results for six 15-km×15-km areas of interest in northern Minnesota, USA, indicate that areal estimates of tree volume, basal area, and density obtained from pixel predictions are comparable to plot-based estimates and estimates by conifer and deciduous classes are also comparable to plot-based estimates. When a mixed conifer/deciduous class was included, predictions for the mixed and deciduous class were confused.

Published by Elsevier Inc.

1. Introduction

1.1. Background and motivation

Forest area and structure by tree species composition classes have received increased attention in recent years as indicators of forest sustainability and biodiversity. The Ministerial Conference on the Protection of Forests in Europe (MCPFE, 2008) includes area by forest type as an indicator for a criterion related to maintaining forest resources and wood production as an indicator for a criterion related to maintaining the productive function of forests. The Montréal Process (Montréal Process, 2005) includes the same indicators for criteria related to maintaining ecosystem biodiversity and maintaining forest productivity. Forest type has been widely used as an inventory variable in North American, Mediterranean, central European countries, and recently Nordic countries. Finally, Action E43 (Harmonization of the national forest inventories of Europe) (COST E43, 2007) of the European program Cooperation in the field of Scientific and Technical Research has selected forest type as an indicator for biodiversity assessments.

Moderate resolution (100-m×100-m or finer) information on tree species composition and forest structural variables such as tree volume and density are crucial for a variety of forest assessments. Relationships between forest species composition and structure, on the one hand, and productivity, economic returns, site occupation, and nutrient use, on the other hand, have been documented in numerous studies (Buongiorno et al., 1994; Sterba and Monserud, 1995; Mård, 1996; Önal, 1997; Edgar and Burk, 2001; Chen & Klinka, 2003; Chen et al., 2003). Commercial forest enterprises rely on regional and national assessments of forest structure by species composition classes to support decisions regarding establishment or expansion of facilities such as paper mills. Ellenwood and Krist (2007) argue persuasively that a moderate resolution data set that includes information on forest species composition and structure is necessary to support strategic insect and disease risk assessments. Numerous authors have reported that greater tree species and size diversity lead to greater overall diversity (Ambuel & Temple, 1983; Schuler & Smith, 1988; Buongiorno et al., 1994; Önal, 1997; Kerr, 1999). MacArthur and MacArthur (1961), Murdoch et al. (1972), and Cody (1975) all reported

E-mail address: rmcroberts@fs.fed.us.

0034-4257/$ – see front matter. Published by Elsevier Inc.
doi:10.1016/j.rse.2008.10.001

that greater tree species and size diversity lead to greater occupation of forest land by birds.

Avian species constitute the majority of terrestrial vertebrate species in most North American boreal forest communities: 71% in the boreal forest of northeastern Ontario, Canada (D'Eon and Watt, 1994); 81% in the western Canadian boreal forest (Smith, 1993); and 70% in the Superior National Forest of Minnesota, United States of America (USA) (Niemi et al., 1998). Habitat for these species depends heavily on forest species composition and structure. MacArthur and MacArthur (1961) attributed 80% of the diversity in bird species to vegetation diversity in five eastern states of the USA. Numerous studies have documented positive avian responses to mixed conifer-deciduous composition and size diversity (Willson, 1974; DesGranges, 1980; Collins et al., 1982; James & Warner, 1982; Ambuel & Temple, 1983; Clark et al., 1983; Rice et al., 1984; Sherry & Holmes, 1985). Hobson and Bayne (2000) found that increases in the conifer component of older western Canadian boreal aspen forests resulted in increases in the numbers and types of niches available for breeding birds. Rempel (2007) reported that hardwood–softwood diversity and structural diversity were two of the four distinct factors associated with species occurrence patterns. Girard et al. (2004) concluded that mixed conifer/deciduous forests are perceived by selected bird communities as a distinct habitat. Kirk et al. (1996) noted that mature and old mixedwood forests provide core breeding populations for species of neo-tropical migrants and that excess individuals in these populations contribute to populating other habitats. The results of these studies support the assertion that mixed conifer/deciduous forests make important contributions to avian habitat.

1.2. Mapping forest composition and structure

National forest inventories (NFI) conducted in North America, Europe and elsewhere are the most important sources of comprehensive information for assessing forest species composition and structure for large geographic areas. Because complete, enumerative inventories are prohibitively expensive, NFIs sample populations of interest and report plot-based estimates of forest resources. For valid sampling designs and corresponding estimators, these plot-based approaches produce asymptotically unbiased estimates of area by conifer, deciduous, and mixed classes, and estimates of structural variables such as tree volume and density by class. However, these approaches are unable to produce credible inferences for small areas due to insufficient sample sizes, and they are unable to depict spatial distributions of forest attributes and related landscape metrics. Therefore, model-based approaches to inference that produce areal estimates consistent with corresponding satellite image-based forest attribute maps are attracting greater interest.

Thessler et al. (2005), however, noted that accurate forest attribute maps are difficult to construct from satellite imagery unless pixels are clearly distinguishable with respect to both vegetation structure and spectral signature. Whereas patterns of species composition and forest structure in intensively managed homogeneous forest stands are relatively discrete with easily identifiable boundaries, patterns in naturally regenerated, uneven-aged, mixed species forests change gradually. As a result, distinguishing species composition and forest structure classes in naturally regenerated, uneven-aged, mixed species forests can be difficult and subjective (Salovaara et al., 2005). Thus, inclusion of a mixed species composition class may contribute to better class prediction.

Natural resource mapping applications typically entail constructing statistical models of relationships between land cover attributes and ancillary variables including satellite image spectral variables, and then predicting attribute values for image pixels. Multivariate statistical modelling approaches are necessary because separate univariate approaches can produce inconsistent predictions such as large tree volume for a pixel predicted to have non-forest land cover. However, parametric multivariate statistical methods generally assume that observations of response variables follow Gaussian distributions, an assumption that is violated for many forest inventory variables. A variety of non-parametric multivariate techniques have been investigated and reported for forestry applications including regression trees (Xu et al., 2005), bootstrap aggregation or bagging (Steele et al., 2003), random forests (Hudak et al., 2008), multivariate adaptive regression splines (Prasad et al., 2006), neural networks (Kimes et al., 2006), and nearest neighbors techniques. For forest inventory mapping and small area estimation applications using satellite imagery, nearest neighbors techniques have enjoyed increasing international popularity (e.g., Koukal et al., 2007, Ohmann et al., 2007, Chirici et al., 2008, Hudak et al., 2008, LeMay et al., 2008, Tomppo et al., 2008).

1.3. Prediction and mapping techniques

With nearest neighbors techniques, predictions for satellite image pixels without observations (i.e., target pixels) are calculated as linear combinations of observations for pixels that are nearest or most similar to the target pixels with respect to a selected distance metric in the space defined by the ancillary variables. For countries with mostly unchanging land uses and good spatial land cover information, classes of relevant categorical land cover variables can be assigned to pixels. For example, Finland has digital map data that can be used with confidence to assign pixels to land cover classes such as forest, agriculture, and urban and to classes of soil types (Katila and Tomppo 2002, Tomppo & Halme, 2004). Thus, in Finland nearest neighbors selected for predicting volume for a forest target pixel seldom include non-forest pixels. However, in North America where land uses change frequently and accurate spatial land cover information is not readily available, nearest neighbors techniques typically are implemented without regard to whether the nearest neighbor pixels are of the same class of a relevant categorical variable. Thus, even though the land cover class predicted for a target pixel may be non-forest, the nearest neighbors may include forest pixels. The consequences are that non-zero tree volume and density may be predicted for pixels predicted to be non-forest, and tree volume and density predictions for pixels predicted to be forest may be erroneously small because of non-forest nearest neighbors.

A potential solution that avoids these consequences is to use a two-step algorithm in which the class of a categorical variable such as species composition class is predicted first, and then the values of continuous variables such as tree volume and density are predicted subject to the constraint that all nearest neighbors must be selected from the predicted class of the categorical variable. Although any technique that predicts classes of categorical variables can be used in the first step of a two-step algorithm, nearest neighbors techniques are selected for the second step because multiple, continuous, non-Gaussian variables must be predicted simultaneously. Among the techniques that can be used in the first step, nearest neighbors techniques are an obvious possibility because they are used in the second step. Nearest neighbors techniques have been used frequently with Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) satellite image data for predicting species composition classes in temperate and boreal forests. Franco-Lopez et al. (2001) predicted hardwood, softwood, and mixed classes for northern Minnesota, USA; McRoberts et al. (2002) and Haapanen et al. (2004) predicted forest/non-forest for northern Minnesota, USA; Ohmann et al. (2007) predicted forest type classes for coastal Oregon, USA; Koukal et al. (2007) predicted deciduous, conifer, and mixed classes in Austria; and Tomppo et al. (2009-this issue) predicted classes of dominant tree species in Finland and conifer and deciduous classes in Italy.

A second possibility for the first step technique is multinomial logistic regression also characterized as logistic discriminant analysis (Seber, 1984). This technique is commonly used for predicting the classes of a categorical forest response variable from continuous

satellite image variables. Koutsias and Karteris (1998, 2000) used a binomial logistic regression model to predict the probability that TM pixels belonged to a burned area in Greece; Coops et al. (2006) used a binomial logistic model and TM data to predict classes of insect damage to forest stands in British Columbia, Canada; Pasher et al. (2007) used a binomial logistic model with TM data to map potential habitat for a bird species in Ontario, Canada; and McRoberts (2006) used a binomial logistic regression model with TM data to predict the probability of forest cover for Minnesota, USA. Tomppo (1992) used multinomial logistic regression with TM data to predict multiple forest soil fertility classes in Finland, and Calef et al. (2005) used it with TM data to predict four vegetation classes for Alaska, USA.

A third possibility for the first step technique is discriminant analysis, another multivariate technique that predicts classes of categorical response variables from continuous ancillary variables (Tomppo et al., 2001). Discriminant analysis may produce more accurate predictions than multinomial logistic regression if the ancillary variables follow multivariate Gaussian distributions. Among the forestry applications, Tomppo (1992) used discriminant analysis with TM data to predict forest soil fertility classes in Finland; Bentz and Endreson (2003) used it with TM data to predict lodgepole pine mortality caused by mountain pine beetle in Montana, USA; Sachs et al. (1997) used the technique to assign TM pixels to conifer, deciduous, and mixed classes in British Columbia, Canada; and Mallinis et al. (2003) used the technique to assign TM pixels to one forest and three non-forest classes in northern Greece.

1.4. Objectives

The overall objective of the study was to construct unbiased maps depicting tree volume (V), basal area (BA), and density (T) within conifer, deciduous, and mixed species composition classes that are suitable for small area estimation and sustainability analyses. The primary data sources were NFI and Landsat TM/ETM+ data. Intermediate and supporting objectives included evaluating nearest neighbors, multinomial logistic regression, and discriminant analysis techniques for predicting non-forest and species composition classes in the first step of two-step algorithms.

2. Data

2.1. Satellite Imagery

The study area was defined by the portion of the row 27, path 27, Landsat scene in northern Minnesota, USA (Fig. 1). Land use for the study area consists of forest land dominated by aspen–birch and spruce–fir associations, agriculture, wetlands, and water. Imagery was acquired for three dates corresponding to early, peak, and late seasonal vegetative stages (Yang et al., 2001), April 2000, July 2001, and November 1999. Two sets of spectral variables were considered: (1) SPEC18, consisting of data for TM/ETM+ bands 1–5 and 7 for each of the three image dates, and (2) SPEC12, consisting of the normalized difference vegetation index (NDVI) (Rouse et al., 1973) transformation and the three tassel cap (TC) transformations (brightness, greenness, and wetness) (Kauth and Thomas, 1976; Crist and Cicone, 1984) for each of the three image dates. To evaluate areal estimates, six 15-km × 15-km areas of interest (AOI) within the study area were selected using a systematic grid of locations.

2.2. Forest inventory data

The Forest Inventory and Analysis (FIA) program of the U.S. Forest Service conducts the NFI of the USA. The program has established field plot centers in permanent locations using a sampling design that produces an equal probability sample (Bechtold and Patterson, 2005; McRoberts et al., 2005). The sampling design is based on a tessellation

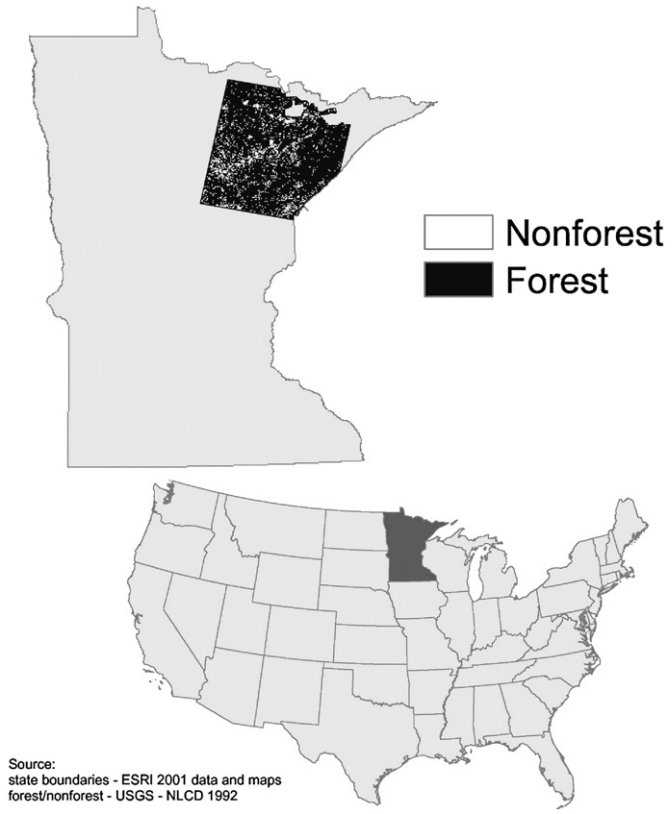

Fig. 1. Study area.

of the USA into approximate 2400-ha (6000-ac) hexagons and features a permanent plot at a randomly selected location within each hexagon. Some states, including Minnesota in which the study area is located, provide additional funding to double the sampling intensity to approximately one plot per 1200 ha. Each plot consists of four 7.32-m (24-ft) radius circular subplots for a total area of 672 m^2. The subplots are configured as a central subplot and three peripheral subplots with centers located at 36.58 m (120 ft) and azimuths of 0°, 120°, and 240° from the center of the central subplot. In general, locations of forested or previously forested plots are determined using global positioning system receivers, whereas locations of non-forested plots are verified using aerial imagery and digitization methods.

Field crews observe species and measure diameter at-breast-height (dbh) (1.37 m, 4.5 ft) and height for all trees with dbhs of 12.7 cm (5 in) or greater. Tree data are aggregated at the subplot level to obtain totals for V, BA, and T. Area (A) by land cover class is obtained for plots by collapsing ground land use conditions into non-forest and multiple forest classes subject to the FIA definition of forest land: area of at least 0.4 ha (1 ac), continuous, external crown-to-crown width of at least 36.58 m (120 ft), and stocking of at least 10 percent. All plots were measured between 1999 and 2003.

For this study, portions of subplots were assigned to three land cover classes: (1) *non-forest* (*NF*) for portions that field crews determined did not qualify as forest land, (2) *conifer* (*C50*) for forested portions with proportion of total BA in conifer trees of at least 0.50, and (3) *deciduous* (*D50*) for forested portions with proportion of total BA in deciduous trees greater than 0.50. Portions of subplots were also assigned to four land cover classes: (1) *non-forest* (*NF*) for portions that field crews determined did not qualify as forest land, (2) *conifer* (*C75*) for forested portions with proportion of total BA in conifer trees of at least 0.75, (3) *deciduous* (*D75*) for forested portions with proportion of total BA in deciduous trees of at least 0.75, and (4) *mixed* (*M75*) for forested portions with proportions of total conifer BA and total deciduous BA both less than 0.75. The proportion 0.75 has ample precedent in both North America and Europe

for use as a threshold for distinguishing among conifer, deciduous, and mixed classes (Hansen & Hahn, 1992; Bossard et al., 2000; Homer et al., 2000; Young et al., 2005). An additional class designated *For* was formed by aggregating all the forest classes.

Plots were assigned to land cover classes in two ways to accommodate two purposes. First, the non-forest and forested portions of plots were assigned to the same sets of three and four land cover classes in the same manner as portions of subplots were assigned. With this approach, forested portions of plots could be assigned to only one land cover class of each set of classes. This approach was used to facilitate combining plot and satellite image data (Section 2.3). Second, portions of plots were assigned to land cover classes by aggregating the assignments of their subplots to classes. With this approach, forested portions of plots could be assigned to multiple land cover classes of each set, depending on the assignments of their constituent subplots. This approach was used to calculate estimates of means and totals that used plot data only, that would be similar to estimates calculated by inventory programs (Section 3.4.1), and that could be used to assess the unbiasedness of pixel-based estimators of means and totals both over all land cover classes and within classes (Section 3.4.2).

2.3. Combining FIA data and satellite image data

The spatial configuration of the FIA subplots with centers separated by 36.58 m and the 30-m×30-m spatial resolution of the TM /ETM+ imagery permits individual subplots to be associated with individual image pixels. The subplot area of 672 m^2 is approximately 19 percent of the 900 m^2 pixel area and is assumed to be sufficient to characterize entire pixels containing subplot centers. The inventory subplot data were converted to a per ha basis and combined with the pixel level satellite image data to construct a subplot/pixel training data set to be used for classification and prediction. When constructing this data set, data were omitted for subplots that were not completely forested or non-forested and for subplots that had any portions classified by field crews as forest land but with no trees with dbh ≥ 12.7 cm because of uncertainty as to tree cover.

Subplot location errors and plot-to-image registration errors may cause subplots to be associated with incorrect and/or multiple pixels. Examples of the consequences of such errors include association of forest ground attribute data with spectral data for pixels with non-forest land cover and association of conifer ground data with spectral data for pixels with deciduous land cover. The probability of these consequences is non-negligible for the naturally regenerated, uneven-aged, mixed species forest stands characteristic of the study area. An alternative that could at least partially alleviate the consequences of these errors is to associate ground data aggregated over the four subplots with the means of spectral variables for 3x3 blocks of pixels centered on the pixels containing the plot centers. Therefore, a plot/pixel block training data set was constructed by using plot data assigned to classes, scaling to a per ha basis, and combining with mean spectral data for 3x3 pixel blocks. Data were omitted for plots that were not completely forested or non-forested and for plots that had any portions classified by field crews as forest land but with no trees with dbh ≥ 12.7 cm. The primary advantage of the plot/pixel block data set is that the detrimental effects of plot location and registration errors may be reduced. The primary disadvantages are that the number of observations is reduced by a factor of at least four; the total plot area is only approximately 8 percent of the 3x3 pixel block area; multiple land cover classes occur more frequently for pixel blocks than for individual pixels; and the relationship between ground observations and spectral data may be weaker. A summary of numbers of subplots and plots by land cover class is provided in Table 1.

Table 1
Distribution of subplots and plots

Class	Reference data		AOIs
	Subplots	Plots	Plots
Total	7533	1383	93.00
NF	2500	534	25.42
C50	2351	391	33.97
D50	2682	458	33.61
C75	1940	276	25.42
D75	2227	310	26.68
M75	866	263	15.49

3. Methods

3.1. Nearest neighbors techniques

Let **Y** denote a possibly multivariate vector of response variables with observations for a sample of size *n* from a finite population of size *N*, and let **X** denote a vector of ancillary variables with observations for all population units. For this study, the population is a set of 30-m×30-m Landsat pixels; $\mathbf{Y}_i$ consists of subplot or plot observations of *A*, *V*, BA, *T*, over all land cover classes and by individual land cover classes, conifer BA proportion, and deciduous BA proportion for the *i*th subplot or plot; and $\mathbf{X}_i$ is the vector of image spectral data for the pixel containing the center of the *i*th subplot or the mean vector of image spectral data for the 3×3 pixel block whose center pixel contains the center of the *i*th plot. In the terminology of nearest neighbors techniques, the set of pixels or pixel blocks for which observations of both response and ancillary variables are available is designated the *reference set*; the set of pixels or pixel blocks for which predictions of response variables are desired is designated the *target set*; and the space defined by the ancillary variables, **X**, is designated the *feature space*. Each target pixel or pixel block is assumed to have a complete set of observations for all feature space variables.

For continuous response variables such as *A*, *V*, BA, and *T*, the k-NN prediction for the *i*th target pixel is,

$$\tilde{y}_i = \frac{1}{W_i}\sum_{j=1}^{k} w_{ij} y_j^i, \tag{1}$$

where $\{y_j^i, j=1,2,\ldots, k\}$ is the set of observations for the *k* reference pixels that are nearest to the *i*th target pixel in feature space with respect to a distance metric, *d*; and w_{ij} is the weight assigned to the *j*th nearest neighbor with $W_i = \sum_{j=1}^{k} w_{ij}$. For categorical variables such as forest/non-forest or land cover class, the prediction for the *i*th target pixel is the most heavily weighted class among the k nearest neighbors. Frequent choices for the weights are $w_{ij}=d_{ij}^t$ where $t\in[-2, 0]$. Many distance measures may be expressed in matrix form as,

$$d_{ij} = (\mathbf{X_i}-\mathbf{X_j})'\mathbf{S}(\mathbf{X_i}-\mathbf{X_j}),$$

where *i* denotes a target pixel for which a prediction is sought, *j* denotes a reference pixel, $\mathbf{X}_i$ and $\mathbf{X}_j$ are the vectors of observations of feature space variables for the *i*th and *j*th pixels, respectively, and **S** is a square matrix. Popular choices for **S** include the identity matrix which results in squared Euclidean distance and a diagonal matrix which results in weighted squared Euclidean distance. Mahalanobis distance results when **S** is the inverse of the covariance matrix of the feature space variables (Kendall and Buckland, 1982). Other choices for **S** are based on canonical correlation analyses (Moeur and Stage, 1995; Temesgen et al., 2003; LeMay et al., 2008) or canonical correspondence analyses (Ohmann and Gregory, 2002; Ohmann et al., 2007). Chirici et al. (2008) and Hudak et al. (2008) evaluated additional distance metrics. For this study, all implementations of the k-NN technique featured the squared Euclidean distance metric and equal weighting of neighbors.

3.2. Multinomial logistic regression

The relationship between a binomial dependent variable, Y, with values $y=0$ or $y=1$ (e.g., non-forest or forest) and continuous independent variables, $\mathbf{X}$, such as the spectral values of satellite imagery, is often expressed in the form,

$$p_i = f(\mathbf{X}_i;\boldsymbol{\beta}), \tag{2}$$

where i indexes pixels, p_i is the probability that $y_i=1$, and $\boldsymbol{\beta}$ is a vector of parameters to be estimated (Agresti, 2007). The function, $f(\mathbf{X}_i;\boldsymbol{\beta})$, expresses the statistical expectation of Y in terms of $\mathbf{X}$ and $\boldsymbol{\beta}$ and is often formulated using the logistic function as,

$$f(\mathbf{X}_i;\boldsymbol{\beta}) = \frac{\exp\left(\sum_{j=1}^{J}\beta_j x_{ij}\right)}{1+\exp\left(\sum_{j=1}^{J}\beta_j x_{ij}\right)}, \tag{3}$$

where $j=1, 2,\ldots, J$ indexes the independent variables, and exp(.) is the exponential function. The parameter vector, $\boldsymbol{\beta}$, is estimated by maximizing the likelihood,

$$L = \prod_{i=1}^{n} f(\mathbf{X}_i;\boldsymbol{\beta})^{y_i}[1-f(\mathbf{X}_i;\boldsymbol{\beta})]^{1-y_i},$$

where n is the number of observations. Thus, the predicted probability that $y=1$ for the ith pixel is,

$$\hat{p}_i^{\,1} = \frac{\exp\left(\sum_{j=1}^{J}\hat{\beta}_j x_{ij}\right)}{1+\exp\left(\sum_{j=1}^{J}\hat{\beta}_j x_{ij}\right)},$$

where the superscript denotes the class, and the predicted probability that y=0 is,

$$\hat{p}_i^{\,0} = 1-\hat{p}_i^{\,1} = \frac{1}{1+\exp\left(\sum_{j=1}^{J}\hat{\beta}_j x_{ij}\right)}.$$

Simple algebra yields,

$$\hat{p}_i^{\,1} = \hat{p}_i^{\,0}\exp\left(\sum_{j=1}^{J}\hat{\beta}_j x_{ij}\right).$$

The logistic regression model approach can be extended from binomial to multinomial response variables (Agresti, 2007); e.g., $y=1$ for NF, $y=2$ for C50, and $y=3$ for D50 . For a response variable with M classes, one class is selected arbitrarily and designated the Mth class. The estimate of the probability that the ith pixel belongs to the Mth class is,

$$\hat{p}_i^{\,M} = \frac{1}{1+\sum_{m=1}^{M-1}\sum_{j=1}^{J}\exp\left(\hat{\beta}_{mj}x_{ij}\right)} \tag{4a}$$

where m indexes the classes of the response variable. The estimate of the probability that the ith pixel belongs to the mth class ($m<M$) is,

$$\hat{p}_i^{\,m} = \hat{p}_i^{\,M}\exp\left(\sum_{j=1}^{J}\hat{\beta}_{mj}x_{ij}\right). \tag{4b}$$

The class prediction for the ith pixel is the class for which $\hat{p}_i^{\,m}$ ($m=1, 2,\ldots, M$) is the greatest. All parameters of the multinomial logistic regression model can be estimated simultaneously using a variety of software packages such as SAS with the CATMOD procedure or *R* with the VGAM library.

3.3. Discriminant analysis

Discriminant analysis can also be used to predict classes of a categorical variable using continuous ancillary variables (McLachlan, 2004). For this study, quadratic discriminant analysis was used where the squared distance in the space of the ancillary satellite image variables from the ith pixel to be classified and the mth class is calculated as,

$$d_m^2(\boldsymbol{X}_i) = \left(\boldsymbol{X}_i-\hat{\boldsymbol{\mu}}_m\right)' \boldsymbol{S}_m \left(\boldsymbol{X}_i-\hat{\boldsymbol{\mu}}_m\right), \tag{5}$$

where $\mathbf{X}_i$ is the vector of ancillary variables for the ith target pixel, and $\hat{\boldsymbol{\mu}}_m$ is the mean vector and $\mathbf{S}_m$ is the inverse of the covariance matrix of the ancillary variables for the mth class as observed for the reference set. The estimate, $\hat{p}_i^{\,m}$ of the probability that the ith target pixel belongs to the mth class is,

$$\hat{p}_i^{\,m} = \frac{\exp\left\{-0.5\left[d_m^2(\boldsymbol{X}_i)-\ln|\boldsymbol{S}_m|-2\cdot\ln(q_m)\right]\right\}}{\sum_{u=1}^{M}\exp\left\{-0.5\left[d_u^2(\boldsymbol{X}_i)-\ln|\boldsymbol{S}_u|-2\cdot\ln(q_u)\right]\right\}}, \tag{6}$$

where exp(.) is the exponential function, ln(.) is the natural logarithm function, $|\mathbf{S}_m|$ is the determinant of $\mathbf{S}_m$, and q_m is the proportion of reference pixels observed to belong to the mth class. The class prediction for the ith pixel is the class for which $\hat{p}_i^{\,m}$ ($m=1, 2,\ldots, M$) is the greatest.

3.4. Areal inference

3.4.1. Probability-based inference

Natural resource applications require estimates of areal parameters such as population means or totals and means or totals by land cover classes for AOIs rather than just observations for individual plots or predictions for individual pixels. The traditional forest inventory approach to estimating these parameters is based on a sample of the population obtained using a sampling design for which each population unit has a known probability of selection. The FIA program assumes an infinite population framework and uses a sampling design in which each population unit has an equal probability of selection (McRoberts et al., 2005; Bechtold & Patterson, 2005). The program assigns means or totals of data for all four subplots to the center point of the central subplot. Inferences regarding population parameters based on probability samples are characterized as *probability-based inference* because of the dependence of the observations on the probability of selection of population units into the sample (Hansen et al., 1983). Under the simple random sampling assumption, a probability-based estimator, $\overline{Y}^{\text{prob}}$, of the population mean, μ, is the sample mean,

$$\overline{Y}^{\text{prob}} = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{a_{\text{plot}}}, \tag{7a}$$

where n is the size of the sample, and a_{plot} is the constant known total plot area. The population total, Y_{tot}, is estimated as

$$\hat{Y}_{\text{tot}} = A_{\text{tot}}\overline{Y}^{\text{prob}}, \tag{7b}$$

where A_{tot} is the constant known population area. The probability-based estimator for the variance of $\overline{Y}^{\text{prob}}$ is,

$$\hat{\text{var}}\left(\overline{Y}^{\text{prob}}\right) = \frac{\sum_{i=1}^{n}\left(y_i-\overline{Y}^{\text{prob}}\right)^2}{n(n-1)}, \tag{8a}$$

and the estimator of the variance of $\hat{Y}_{\text{tot}}$ is

$$\hat{\text{var}}\left(\hat{Y}_{\text{tot}}\right) = A^2\hat{\text{var}}\left(\overline{Y}^{\text{prob}}\right). \tag{8b}$$

These probability-based estimators were used to calculate estimates of population means and totals for *V*, BA, and *T* over all land cover classes and totals by individual classes. Observations for all four subplots of all plots with centers in AOIs were used with the probability-based estimators, regardless of whether all subplots were in the AOI.

For estimation of means of *V*, BA, and *T* within land cover classes, the ratio-of-means (rom) estimator was used because portions of plots can be assigned to different land cover classes, and the areas of those portions vary from plot-to-plot. The rom estimator of the mean for the *m*th land cover class is,

$$\overline{Y}_m^{\text{rom}} = \frac{\overline{Y}_m}{\overline{A}_m}, \tag{9}$$

where

$$\overline{Y}_m = \frac{1}{n}\sum_{i=1}^{n} y_{im},$$

$$\overline{A}_m = \frac{1}{n}\sum_{i=1}^{n} a_{im}$$

and y_{im} and a_{im} are the observations of the forest attribute and area, respectively, for the mth land cover class for the *i*th plot (Cochran, 1977, Section 2.11; Bechtold & Patterson, 2005; Tomppo, 2006). A Taylor series expansion was used to obtain an approximate variance estimator,

$$\hat{\text{var}}\left(\overline{Y}_m^{\text{rom}}\right) = \frac{\hat{\text{var}}\left(\overline{Y}_m\right) - 2\overline{Y}_{\text{rom}}\hat{\text{cov}}\left(\overline{Y}_m, \overline{A}_m\right) + \overline{Y}_{\text{rom}}^2\hat{\text{var}}\left(\overline{A}_m\right)}{\overline{A}_m^2}, \tag{10}$$

where,

$$\hat{\text{cov}}\left(\overline{Y}_m, \overline{A}_m\right) = \frac{\sum_{i=1}^{n}\left(y_{im}-\overline{Y}_m\right)\left(a_{im}-\overline{A}_m\right)}{\sqrt{\sum_{i=1}^{n}\left(y_{im}-\overline{Y}_m\right)^2}\sqrt{\sum_{i=1}^{n}\left(a_{im}-\overline{A}_m\right)^2}}$$

(Cochran, 1977).

Estimators are characterized as unbiased if for all sample sizes, the statistical expectation of the estimator, $\hat{\mu}$, over all possible samples equals the parameter, i.e., if $E(\hat{\mu})=\mu$ (Kendall and Buckland, 1982). Estimators are further characterized as asymptotically unbiased if the estimator produces estimates that approach the parameter value as the sample size increases, i.e., if $\lim_{n\to\infty}\hat{\mu}=\mu$. For the simple random sampling design, Eqs. (7a) and (7b) are both unbiased and asymptotically unbiased as estimators of the population mean and total.

3.4.2. Model-based inference

Although the FIA probability-based estimators have the desirable properties of being both unbiased and asymptotically unbiased, variance estimates can be large for small areas due to small sample sizes and no spatial products result. A class of estimators that circumvents these difficulties is based on aggregating model predictions for multiple pixel AOIs. An estimator, $\overline{Y}^{\text{mod}}$, of a population mean, μ, for a finite population defined by *N* satellite image pixels, is,

$$\overline{Y}^{\text{mod}} = \frac{1}{N}\sum_{i=1}^{N}\hat{y}_i, \tag{11}$$

where $\hat{y}_i$ is a pixel prediction obtained using a model. In this context, a prediction obtained using a nearest neighbors technique is considered a model prediction. Inferences based on aggregating model predictions for multiple pixel AOIs are characterized as *model-based inference* because of the role of the models. McRoberts (2006) and McRoberts et al. (2007) provide additional background on the distinctions between the probability-based and model-based modes of inference for remote sensing applications. An important distinction between these approaches is that whereas probability-based estimators used by forest inventory programs are usually both unbiased and asymptotically unbiased, such cannot be asserted for model-based estimators. Thus, the compromise with model-based estimators is that although they can produce more precise small area estimates and spatial products, their unbiasedness is not assured.

One approach to assessing the unbiasedness of model-based estimators is to exploit the unbiasedness of probability-based estimators. For future reference, estimates obtained using probability-based estimators are designated *plot-based*, and estimates obtained using model-based estimators with satellite imagery are designated *pixel-based*. If the sample size for an AOI is sufficiently large (usually a subjective decision), and if absolute values of ratios of deviations between plot-based and pixel-based estimates and standard errors of the deviations are less than approximately 2.0, then a measure of confidence can be asserted for a claim of unbiasedness for the model-based estimator. However, an approach for combining variance estimates based on both probability-based and model-based estimators to produce the required standard error of the deviations is not apparent because of the very different assumptions underlying the two modes of inference. In addition, the few reports of variance estimators for nearest neighbors techniques (Baffeta et al., 2009-this issue; Magnussen et al., 2009-this issue; McRoberts et al., 2007) indicate that variance estimation is complex and computationally intensive. Therefore, for this study, plot-based standard errors of estimates were used as approximations to the standard errors of deviations between estimates although they are underestimates because no accommodation is made for the uncertainty in the pixel-based estimates. To compensate for the underestimation, the standard for the absolute value of the ratio of deviations between pixel-based and a plot-based estimates and plot-based standard errors was relaxed from 2.0 to 2.25 for this study.

3.5. Analyses

The k-NN, multinomial logistic regression, and discriminant analysis techniques were compared with respect to the accuracies of their land cover class predictions. Three measures of accuracy were used: *overall accuracy* (OA), which is the proportion of observations correctly classified; *user's accuracy*, which is the ratio of the number of correct predictions and the total number of predictions for a class; and *producer's accuracy*, which is the ratio of the number of correct predictions and the total number of observations for a class. In addition, the *Kappa coefficient*, a measure of association describing the agreement between two classifications (Kraemer, 1983) was calculated. When one classification is based on observations and the other is based on predictions, the Kappa coefficient is considered a measure of accuracy and is estimated as,

$$\hat{K} = \frac{n\sum_{m=1}^{M} c_{ii} - \sum_{m=1}^{M}\left(c_{i+}\cdot c_{+i}\right)}{n^2 - \sum_{m=1}^{M}\left(c_{i+}\cdot c_{+i}\right)}, \tag{12}$$

where *n* is the number of observations; *M* is the number of land cover classes; c_{ii} is the number of plots both observed and predicted to be in the *i*th land cover class; c_{i+} and c_{+i} are the numbers of observations observed in the *i*th land cover class and predicted to be in the *i*th land cover class, respectively; and the dot symbol (·) denotes multiplication (Congalton, 1991). Perfect agreement between the observations and predictions is indicated when $K=1$, whereas $K=0$ indicates no agreement beyond that expected by chance. Both the accuracy and Kappa analyses included comparisons of results obtained using the SPEC18 and SPEC12 spectral data and the subplot/pixel and plot/pixel

block data sets. For each of the four data set combinations, all the measures were calculated for both the set of three land cover classes and the set of four classes.

Four approaches for predicting land cover class and *V*, BA, and *T* were investigated. A single-step k-NN algorithm was used for which all response variables were predicted simultaneously for each pixel without regard to the land cover classes of the nearest neighbors. In addition, three two-step algorithms were used with either the k-NN, multinomial logistic regression model, or discriminant analysis technique in the first step to predict the land cover class of the target pixel and the k-NN technique in the second step to predict *V*, BA, and *T* subject to the constraint that all nearest neighbors were from the land cover class predicted in the first step. For all four approaches, k-NN reference sets were constructed using the four combinations of the SPEC18 and SPEC12 spectral data and the subplot/pixel and the plot/pixel block data sets. For each of the four data set combinations, estimates were calculated for both the set of three land cover classes and the set of four classes.

For all four approaches, estimates of A by land cover class were calculated as,

$$\hat{A}_m = A_{\text{tot}}\hat{p}_m \tag{13}$$

and

$$\hat{\text{var}}\left(\hat{A}_m\right) = A_{\text{tot}}^2 \frac{\hat{p}_m\left(1-\hat{p}_m\right)}{N}, \tag{14}$$

where m designates land cover class, A_{tot} is total area, *N* is the total number of pixels, and p_m is the proportion of the *N* pixels predicted to be in the *m*th class.

Accuracy assessments for the prediction approaches were conducted at the pixel- and pixel block-levels using the reference set data with the leave-one-out cross-validation technique (Lachenbruch & Mickey, 1986) and at the areal-level by comparing pixel-based and plot-based estimates for the six AOIs. For the reference set assessments, the observations and predictions of *V*, BA, and *T* at the pixel level were compared using,

$$\bar{R}^2 = \frac{1}{3}\left(R_V^2 + R_{\text{BA}}^2 + R_T^2\right) \tag{15}$$

where

$$R^2 = \frac{SS_{\text{mean}} - S_{\text{k-NN}}}{SS_{\text{mean}}},$$

SS_{mean} is the sum of squared differences between the reference set observations and their mean, and $SS_{\text{k-NN}}$ is the sum of squared differences between reference set observations and their predictions. Larger values of R^2 and $\bar{R}^2$ indicate that predictions are closer to observations as is desired, whereas negative values indicate that the mean is better than the predictions, a particularly undesirable condition. The R^2 criterion is independent of the ranges of values of the forest attribute variables and can be easily modified to accommodate unequal weighting of the variables.

Pixel-based areal estimates obtained using the two-step algorithms and plot-based estimates were compared for the six AOIs collectively and separately. First, the data for all six AOIs were aggregated, and then pixel-based means of V, BA, and T over all land cover classes were calculated and compared to plot-based estimates. Second, estimates of mean V, BA, and T over all land cover classes were obtained for each AOI, and deviations from corresponding plot-based estimates were compared using root mean square deviation (RMSD), mean absolute deviation, and mean deviation. For all areal analyses using the plot/pixel block reference set data, means of spectral variables over 3×3 blocks of pixels were calculated for the AOIs so that the reference set and target set data would be compatible.

Estimates within land cover classes were also calculated. However, the numbers of subplots by land cover class within individual AOIs were small, often less than 5, resulting in plot-based standard errors for individual AOIs that were too large to permit meaningful comparisons of plot-based and pixel-based estimates. Therefore, estimates within land cover classes for individual AOIs were not evaluated. Instead, estimates of mean V, BA, and T within land cover classes over all six AOIs were calculated and compared to plot-based estimates calculated using the rom estimator (Eq. (9)). In addition, for purposes of integrating estimates of both within class means and class areas, totals within land cover classes over all six AOIs were also calculated.

4. Results and discussion

4.1. Extrapolations

Predictions for pixels whose values of feature space variables are outside the ranges represented in the reference set are characterized as extrapolations. If the number of extrapolations, as a proportion of the total number of predictions is large, bias may be a serious problem. Based on the convex hull technique used by Thessler et al. (2005) and illustrated by McRoberts (2009-this issue), the proportions of pixels for the six AOIs that required extrapolations were slightly more than 0.01 for each of the SPEC18 and SPEC12 spectral data sets. For each target pixel requiring an extrapolation, the number of forested pixels among the 30 nearest neighbor reference pixels was never greater than one and was zero for more than half the cases. Therefore, because the proportions of target pixels requiring extrapolations were small, and because their nearest neighbors were overwhelmingly NF, no special consideration was given to these pixels.

4.2. Reference set analyses

4.2.1. Single-step algorithm

The primary issue underlying this study is that without good ancillary information that can be used to assign classes of relevant categorical variables to reference pixels, users of nearest neighbors techniques have two choices: (1) accept that nearest neighbors for predicting continuous variables may include some pixels of a class different than the class predicted for the target pixel and any errors that result, or (2) use a two-step algorithm in which the class of the categorical variable is predicted in the first step, and then continuous variables are predicted in the second step subject to the constraint that all nearest neighbor pixels must come from the predicted class of the categorical variable. If the proportions of nearest neighbor pixels

Table 2
Proportions of seven nearest neighbors by three land cover classes for the reference set using the single-step k-NN algorithm with the SPEC12 spectral data[a]

Land cover class of neighbor pixel	Land cover class of reference pixel						
	Three land cover classes			Four land cover classes			
	NF	C50	D50	NF	C75	D75	M75
Subplot/pixel data set							
NF	0.89	0.04	0.07	0.88	0.04	0.07	0.02
C50/C75	0.07	0.71	0.22	0.07	0.63	0.12	0.18
D50/D75	0.16	0.20	0.64	0.16	0.14	0.57	0.13
M75				0.06	0.27	0.26	0.41
Plot/pixel block data set							
NF	0.93	0.02	0.05	0.92	0.02	0.05	0.02
C50/C75	0.05	0.79	0.16	0.06	0.74	0.03	0.17
D50/D75	0.14	0.13	0.73	0.16	0.05	0.68	0.11
M75				0.06	0.20	0.19	0.55

[a] NF=non-forest, C50=conifer BA proportion ≥0.50, D50=deciduous BA proportion >0.50, C75=conifer BA proportion ≥0.75, D75=deciduous BA proportion ≥0.75, M75=conifer and deciduous BA proportions both <0.75.

for predicting the continuous variables that are not from the predicted classes of the categorical variable are large, then two-step algorithms merit consideration. For this study, when the predicted class was NF, relatively small proportions of nearest neighbor reference pixels were from forest classes (Table 2). However, when one of the forest classes (D50, C50, D75, C75, M75) was predicted, the proportion of nearest neighbors reference pixels from different classes was much greater. In particular, confusion between the NF and deciduous classes was approximately twice the confusion between the NF and conifer classes. The consequences of these erroneous nearest neighbor selections were that pixel-based estimates of mean V, BA, and T for the NF class were non-zero, and estimates for at least some forest classes were substantially less than the plot-based estimates (Table 3). Deviations between the pixel-based and plot-based estimates were greatest for the deciduous and mixed classes as expected based on the greater proportions of NF reference pixels erroneously selected for these classes than for the conifer classes (Table 2). For the deciduous and mixed classes, ratios of absolute values of deviations between pixel-based and plot-based means and plot-based standard errors ranged from approximately 5 to greater than 13, too large to assert unbiasedness for the model-based estimator. Similar results were found when using the SPEC18 spectral data and the plot/pixel block data set. The problem of non-zero estimates for the NF class can be remedied easily by setting to zero predictions for all forest attribute variables for pixels with NF predictions, but the result would be that estimates of mean V, BA, and T over all land cover classes would then be biased downward. Thus, consideration of two-step algorithms merits consideration.

4.2.2. Predicting land cover classes

Differences in accuracies of land cover class predictions obtained using the k-NN, multinomial logistic regression, and discriminant analysis techniques were not great (Table 4). OA and K values were similar for both the SPEC18 and SPEC12 spectral data, regardless of the technique. Better results were obtained for the plot/pixel block data set than for subplot/pixel data set, although the two data sets are not directly comparable because of spatial resolution and size differences. Producer's and user's accuracies were generally similar for the k-NN, multinomial logistic regression, and discriminant analysis techniques.

User's and producer's accuracies were generally comparable to those reported for similar studies. For predicting conifer and deciduous classes, Franco-Lopez et al. (2001) reported user's accuracies of 0.63–0.90 and producer's accuracies of approximately 0.80 for the same geographic area as this study. Tomppo et al. (2009-this issue) reported slightly greater accuracies for conifer and deciduous classes for an Italian study. For predicting conifer, deciduous, and mixed classes, Franco-Lopez et al. (2001) reported similar accuracies, although those for the M75 class were slightly greater. Wickham et al. (2004) reported similar accuracies for the deciduous and mixed classes but smaller accuracies for the conifer class for an accuracy assessment of the Great Lakes portion of the 1992 National Land Cover Data Set (Vogelmann et al., 2001). Koukal et al. (2007) reported similar accuracies for the same three classes for a study in Austria, and Martin et al. (1998) reported greater accuracies for a study in Massachusetts, USA.

4.2.3. Predicting V, BA, and T

The two-step algorithms with the k-NN, multinomial logistic regression, and discriminant analysis techniques in the first step were used to predict V, BA, and T for reference pixels using leave-one-out cross-validation. Generally, the three algorithms produced similar results, regardless of whether the SPEC18 or the SPEC12 spectral data were used (Table 5), but the plot/pixel block data set produced better results than the subplot/pixel data set. Generally, the single-step algorithm produced slightly better results than the two-step

Table 4
Accuracy measures for land cover class predictions for the reference set using SPEC12 spectral data and leave-one-out cross-validation

Accuracy measure[a]	Land cover class[b]	Subplot/pixel data set			Plot/pixel block data set		
		k-NN	Log.[c]	Discr.[c]	k-NN	Log.[c]	Discr.[c]
Three land cover classes (NF, C50, D50)							
$\hat{K}$	All	0.57	0.58	0.56	0.67	0.72	0.71
OA	All	0.72	0.72	0.71	0.78	0.81	0.81
UA	NF	0.91	0.86	0.92	0.94	0.92	0.95
	C50	0.70	0.69	0.67	0.77	0.80	0.78
	D50	0.62	0.64	0.62	0.67	0.72	0.71
PA	NF	0.68	0.73	0.63	0.76	0.83	0.76
	C50	0.68	0.70	0.72	0.75	0.79	0.84
	D50	0.79	0.72	0.76	0.83	0.81	0.83
Four land cover classes (NF, C75, D75, M75)							
$\hat{K}$	All	0.51	0.52	0.50	0.58	0.64	0.61
OA	All	0.65	0.66	0.64	0.69	0.74	0.71
UA	NF	0.89	0.85	0.92	0.95	0.91	0.96
	C75	0.62	0.59	0.60	0.68	0.70	0.69
	D75	0.56	0.59	0.56	0.58	0.66	0.59
	M75	0.28	0.33	0.26	0.49	0.57	0.50
PA	NF	0.70	0.74	0.63	0.75	0.85	0.75
	C75	0.69	0.74	0.74	0.67	0.73	0.78
	D75	0.79	0.76	0.78	0.75	0.74	0.80
	M75	0.10	0.01	0.10	0.54	0.55	0.45

[a] OA = overall accuracy, PA = producer's accuracy, UA = user's accuracy.
[b] See Table 2 footnote for land cover class definitions.
[c] Log = multinomial logistic regression model, Discr = discriminant analysis.

Table 3
Estimates of means within land cover classes for the reference set using the single step algorithm, the SPEC12 spectral data, the pixel/subplot data set, the best choice of k, and leave-one-out cross-validation[a]

Variable		Three land cover classes[b]			Four land cover classes[b]			
		NF	C50	D50	NF	C75	D75	M75
Plot-based estimates								
V (m^3/ha)	Mean	0.00	1192.46	1373.54	0.00	1115.08	1348.68	1523.08
	SE	0.00	31.42	49.46	0.00	35.68	58.72	38.20
BA (m^2/ha)	Mean	0.00	66.01	73.04	0.00	62.02	71.15	83.53
	SE	0.00	1.40	1.86	0.00	1.58	2.18	1.84
T (count/ha)	Mean	0.00	180.88	174.06	0.00	174.57	168.30	206.26
	SE	0.00	3.41	2.27	0.00	3.90	2.49	4.11
Pixel-based estimates								
V (m^3/ha)	Mean	**401.98**	1143.34	**1130.87**	**401.98**	1132.86	**1118.22**	**1192.80**
BA (m^2/ha)	Mean	**21.46**	63.91	**59.84**	**21.46**	63.51	**58.84**	**65.25**
T (count/ha)	Mean	**53.17**	175.03	**143.43**	**53.17**	176.26	**139.41**	**166.01**

[a] Bold entries indicate absolute values of ratios of deviations between plot-based and pixel-based estimates and plot-based standard errors exceed 2.25.
[b] See Table 2 footnote for land cover class definitions.

Table 5
$\bar{R}^2$ for reference set predictions using leave-one-out cross-validation[a]

Data set	Spectral data	Single-step algorithm	Two-step algorithm[b]		
			1st step: k-NN	1st step: Log.	1st step: Discr.
Three land cover classes (NF, C50, D50)[c]					
Subplot/ pixel	SPEC18	0.11	0.08	0.08	0.09
	SPEC12	0.11	0.08	0.09	0.08
Plot/ pixel block	SPEC18	0.23	0.19	0.22	0.22
	SPEC12	0.24	0.20	0.23	0.22
Four land cover classes (NF, C75, D75, M75)[c]					
Subplot/ pixel	SPEC18	0.11	0.08	0.08	0.08
	SPEC12	0.11	0.08	0.09	0.08
Plot/ pixel block	SPEC18	0.23	0.19	0.22	0.23
	SPEC12	0.24	0.20	0.24	0.22

[a] $\bar{R}^2 = \frac{1}{3}\left(R_V^2 + R_{TD}^2 + R_{BA}^2\right)$ where $R^2 = \frac{SS_{mean} - SS_{k\text{-}NN}}{SS_{mean}}$ Eq. (15).
[b] Log = multinomial logistic regression, Discr = discriminant analysis.
[c] See Table 2 footnote for land cover class definitions.

algorithms, suggesting that the loss of potential nearest neighbors with the two-step algorithms had a minor detrimental effect on estimates of mean *V*, BA, and *T*. The values of the $\bar{R}^2$ criterion were small, suggesting a weak relationship between the spectral data and the *V*, BA, and *T* reference set observations. This result is supported by the large k-values associated with the maximum values of the $\bar{R}^2$ criterion (Fig. 2). Although pixel- and pixel block-level analyses are useful for comparing techniques and data sets, most inventory applications require areal estimates for multiple pixel AOIs.

4.3. Areal analyses

4.3.1. Operational standarization

When applying the two-step algorithms to data for the six AOIs, three operational features of the k-NN components of these algorithms were standardized to reduce computational intensity. First, for predicting land cover classes in the first step the *k*-value was set to $k=7$. For the reference set analyses, *k*-values that produced maximum OAs ranged from $k=7$ to $k=30$, depending on whether the SPEC18 or the SPEC12 spectral data were used and whether subplot/pixel or plot/pixel block data sets were used. However, for each combination, a wide range of *k*-values produced OAs that were approximately the same as the maximum values. Selection of $k=7$ did not reduce any OA by more than 0.01. Second, for predicting *V*, BA, and *T* in the second step of the algorithm, the *k*-value was set to $k=30$. For the reference set analyses, *k*-values that produced maximum values of the $\bar{R}^2$ criterion were as large as $k=55$. However, selection of $k=30$ did not reduce any value of the $\bar{R}^2$ criterion by more than 0.01 (Fig. 2). Third, because results obtained using the SPEC18 and the SPEC12 spectral data were nearly indistinguishable, the SPEC12 spectral data were used exclusively for the AOI analyses.

4.3.2. Estimates over all land cover classes

Data for all six AOIs were aggregated to produce a single population encompassing 135,000 ha, and pixel-based estimates of mean V, BA, and T over all land cover classes were calculated and compared to plot-based estimates. When using the subplot/pixel data set, absolute values of ratios of deviations between pixel-based and plot-based estimates of means and plot-based standard errors were all less than 2.0 for all techniques for both three and four land cover classes (Table 6). When using the plot/pixel block data set, ratios were generally less than 2.0, although ratios for *V* were between 2.5 and 3.0 for the algorithms with the k-NN and discriminant analysis techniques

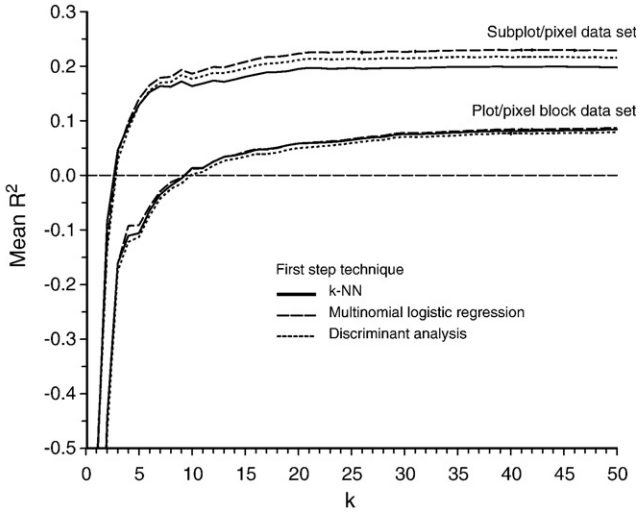

Fig. 2. Mean R^2 ($\bar{R}^2$) criterion (Eq. (15)) versus k for two-step algorithms with SPEC12 spectral data for three land cover classes (NF, C50, and D50).

Table 6
Estimates of areal means over all land cover classes for data aggregated over six AOIs using SPEC12 spectral data[a]

Type of estimates		V (m^3/ha)	BA (m^2/ha)	T (count/ha)
Plot-based estimates				
Mean		62.38	11.47	325.50
SE		5.46	0.97	27.65
Pixel-based estimates: three land cover classes				
Subplot/ pixel	single-step	64.90	11.57	320.76
	two-step, 1st step: k-NN	70.89	12.61	348.64
	two-step, 1st step: Log.[b]	69.67	12.41	342.51
	two-step, 1st step: Discr.[c]	73.10	12.99	357.82
Plot/pixel block	single-step	70.41	12.44	340.82
	two-step, 1st step: k-NN	**76.61**	13.48	366.92
	two-step, 1st step: Log.[b]	72.45	12.76	347.52
	two-step, 1st step: Discr.[c]	**78.37**	**13.78**	373.30
Pixel-based estimates: four land cover classes				
Subplot/pixel	single-step	64.90	11.57	320.76
	two-step, 1st step: k-NN	70.14	12.47	345.64
	two-step, 1st step: Log.[b]	67.17	11.94	331.43
	two-step, 1st step: Discr.[c]	72.12	12.80	353.44
Plot/pixel block	single-step	70.41	12.44	340.82
	two-step, 1st step: k-NN	**76.25**	13.39	363.24
	two-step, 1st step: Log.[b]	70.62	12.44	339.29
	two-step, 1st step: Discr.[c]	**78.42**	**13.78**	373.70

[a] Bold entries indicate absolute values of ratios of deviations between plot-based and pixel-based estimates and plot-based standard errors exceed 2.25.
[b] Log = multinomial logistic regression.
[c] Discr = discriminant analysis.

in the first step for both three and four land cover classes. For both the subplot/pixel and plot/pixel block data sets for both three and four land cover classes, the algorithm with the multinomial logistic regression model in the first step was superior among two-step algorithms with no ratio greater than 1.35 when using the subplot/pixel data set. However, the single-step algorithm was slightly superior to all the two-step algorithms for predicting mean *V*, BA, and *T* over all land cover classes when using both the subplot/pixel and plot/pixel block data, a result that is consistent with the result obtained for the reference set analyses (Table 5). In summary, when using the subplot/pixel data set, all three two-step algorithms produced estimates of means over all land cover classes that were very similar to the plot-based estimates, and the two-step algorithm with the multinomial logistic regression model in the first step produced estimates similar to the plot-based estimates using both the subplot/pixel and plot/pixel block data sets.

To evaluate the algorithms for smaller AOIs, estimates of mean V, BA, and T were calculated for the 22,500-ha individual AOIs. Estimates for individual AOIs are not reported, because the AOIs are considered random selections from within the study area. Instead, features of the distributions of the deviations between plot-based and pixel-based estimates of means were assessed using RMSD, mean absolute deviation, and mean deviation. Results are reported only for the two-step algorithm with the multinomial logistic regression model in the first step because of its superiority to the single-step algorithm and the other two-step algorithms. Absolute values for all three measures were close to or less than the smallest standard error among the six AOIs, except for values of RMSD and mean absolute deviation for A_{For} which were still less than 2.0 standard errors (Table 7). Because the *V*, BA, and *T* variables are highly correlated, no particular meaning should be attributed to the observation that mean deviations for all three variables were of the same sign. In summary, for the two-step algorithm with the multinomial logistic regression model in the first step using both the subplot/pixel and plot/pixel block data sets, absolute values of all measures of the deviations were smaller than twice the smallest standard error among the six AOIs.

Table 7
Deviations between plot-based estimates and pixel-based estimates of AOI means using SPEC12 spectral data

Measure	Three land cover classes (NF, C50, D50)[a]				Four land cover classes (NF, C75, D75, M75)[a]			
	A_{For}	V	BA	T	A_{For}	V	BA	T
Plot-based estimates								
Minimum SE among 6 AOIs	0.06	11.52	1.97	53.40	0.06	11.52	1.97	53.40
Subplot/pixel data set; single-step algorithm								
RMSD	0.11	19.11	2.92	76.61	0.11	19.11	2.92	76.61
Mean absolute deviation	0.08	14.56	2.24	61.32	0.08	14.56	2.24	61.32
Mean deviation	0.01	−14.44	−2.02	−39.93	0.01	−14.44	−2.02	−39.93
Subplot/pixel data set; two-step algorithm with multinomial logistic regression model								
RMSD	0.09	12.03	1.89	56.17	0.09	10.88	1.74	52.46
Mean absolute deviation	0.07	10.73	1.70	46.99	0.07	10.00	1.55	45.12
Mean deviation	−0.01	−7.88	−1.03	−19.19	−0.01	−5.36	−0.57	−8.12
Plot/pixel data set: single-step algorithm								
RMSD	0.11	13.81	2.14	62.70	0.11	13.81	2.14	62.70
Mean absolute deviation	0.09	11.19	1.71	53.29	0.09	11.19	1.71	53.29
Mean deviation	0.04	−8.61	−1.07	−17.50	0.04	−8.61	−1.07	−17.50
Plot/pixel data set: two-step algorithm with multinomial logistic regression model								
RMSD	0.11	15.40	2.33	63.54	0.12	14.38	2.19	60.47
Mean absolute deviation	0.09	12.32	1.88	51.69	0.10	11.57	1.74	49.65
Mean deviation	0.04	−10.65	−1.39	−24.20	0.06	−8.81	−1.06	−15.98

[a] See Table 2 footnote for land cover class definitions.

4.3.3. Estimates within land cover classes

When predicting means and totals for *V*, BA, and *T* within land cover classes, the algorithm with the multinomial logistic regression model in the first step was superior to the other algorithms; therefore, only results for this algorithm are reported. When using three land cover classes, ratios of deviations between pixel-based and plot-based estimates of means and plot-based standard errors were less in absolute value than 1.3 for all classes using both the subplot/pixel and plot/pixel block data sets, except for V for the D50 class using the plot/pixel block data set where the ratio was slightly greater than 2.0 (Table 8a). Pixel-based estimates of totals within land cover classes were also calculated for purposes of integrating estimates of both means and areas within classes. For the three land cover classes using the subplot/pixel data set, ratios for totals were less in absolute value than 2.0 except for *V* and *A* for the D50 class where the ratios were approximately 2.2 and 2.3, respectively. For the plot/pixel block data set, ratios were less than approximately 2.0, except for *V* where the ratio was approximately 3.0. In summary, when used with the subplot/pixel data set, the two-step algorithm with the multinomial logistic regression model in the first step produced estimates of within class means and totals that were very similar to the plot-based means and totals. When used with the plot/pixel-block data set, only the estimate for total *V* for the D50 class was substantially larger than the plot-based estimate.

When using four land cover classes, ratios of deviations between pixel-based and plot-based estimates of means and plot-based standard errors were less in absolute value than 1.0 for all classes using the subplot/pixel data set (Table 8b). When using the plot/pixel block data set, ratios for means within classes were less than 1.7, except for BA and *V* for the D75 class where the ratios were approximately 2.3 and 3.0 respectively. For totals, many ratios were greater than 4.0 when using the subplot/pixel data set. These poor results are attributed to confusion between the D75 and M75 classes rather than to pixel-based estimates of mean *V*, BA, and *T* within classes which are close to plot-based estimates. The cause of the confusion is partially attributed to the arbitrary threshold distinguishing the D75 and M75 classes and to the small relative proportion of reference observations for the M75 class for the subplot/pixel data set (Table 1). For the plot/pixel block data set, absolutes values of ratios were less than 1.5 except for estimates of BA and *V* for the D75 class which were approximately 2.2 and 3.0, respectively. In summary, none of the two-step algorithms produced acceptable within class estimates for all variables for four land cover classes.

Because of its overall superiority with respect to estimates of means and totals within the NF, C50, and D50 classes, the two-step algorithm with the multinomial logistic regression model in the first step was used to construct maps of per ha means of *V*, BA, and *T* within these classes for each AOI (Fig. 3).

Table 8a
Estimates for three land cover classes for aggregation of data for the six AOIs using SPEC12 spectral data and the two-step algorithm with the multinomial logistic regression model in the first step[a]

Variable	Means within classes[b]			Totals within classes[b]		
	Estimate	C50	D50	Estimate	C50	D50
Plot-based estimates						
A				Total (ha × 10^3)	45.41	44.92
				SE (ha × 10^3)	5.17	5.15
V	Mean (m^3/ha)	91.64	99.37	Total ($m^3 \times 10^6$)	4.00	4.42
	SE (m^3/ha)	8.89	8.00	SE ($m^3 \times 10^6$)	0.61	0.58
BA	Mean (m^2/ha)	17.26	17.80	Total ($m^2 \times 10^6$)	0.75	0.90
	SE (m^2/ha)	1.63	1.22	SE ($m^2 \times 10^6$)	0.11	0.10
T	Mean (count/ha)	546.55	456.42	Total (count × 10^6)	23.75	20.19
	SE (count/ha)	48.24	27.16	SE (count × 10^6)	3.63	2.49
Pixel-based estimates; subplot/pixel data set						
A				Total (ha × 10^3)	45.41	**56.79**
V	Mean (m^3/ha)	82.38	99.78	Total ($m^3 \times 10^6$)	3.74	5.67
BA	Mean (m^2/ha)	15.47	17.12	Total ($m^2 \times 10^6$)	0.70	0.97
T	Mean (count/ha)	484.98	426.37	Total (count × 10^6)	22.02	24.92
Pixel-based estimates; plot/pixel block data set						
A				Total (ha × 10^3)	41.47	53.49
V	Mean (m^3/ha)	86.76	115.59	Total ($m^3 \times 10^6$)	3.60	**6.18**
BA	Mean (m^2/ha)	16.49	19.43	Total ($m^2 \times 10^6$)	0.68	1.04
T	Mean (count/ha)	522.02	472.36	Total (count × 10^6)	21.65	25.27

[a] Bold entries indicate absolute values of ratios of deviations between plot-based and pixel-based estimates and plot-based standard errors exceed 2.25.
[b] See Table 2 footnote for land cover class definitions.

Table 8b
Estimates for four land cover classes for aggregation of data for the six AOIs using SPEC12 spectral data and the two-step algorithm with the multinomial logistic regression model in the first step[a]

Variable	Means within classes[b]				Totals within classes[b]			
	Estimate	C75	D75	M75	Estimate	C75	D75	M75
Plot-based estimates								
A					Total (ha × 10^3)	33.97	35.66	20.70
					SE (ha × 10^3)	5.08	4.86	3.21
V	Mean (m^3/ha)	83.98	94.83	113.78	Total ($m^3 \times 10^6$)	2.72	3.34	2.36
	SE (m^3/ha)	9.81	8.69	12.02	SE ($m^3 \times 10^6$)	0.50	0.51	0.44
BA	Mean (m^2/ha)	15.74	16.96	21.11	Total ($m^2 \times 10^6$)	0.51	0.60	0.44
	SE (m^2/ha)	1.78	1.30	2.22	SE ($m^2 \times 10^6$)	0.09	0.09	0.08
T	Mean (count/ha)	512.60	454.01	547.36	Total (count × 10^6)	16.80	15.88	11.26
	SE (count/ha)	59.28	33.36	51.27	SE (count × 10^6)	3.14	2.36	2.10
Pixel-based estimates; subplot/pixel data set								
A					Total (ha × 10^3)	**46.43**	**54.63**	**0.39**
V	Mean (m^3/ha)	76.61	100.07	116.54	Total ($m^3 \times 10^6$)	3.56	**5.47**	**0.05**
BA	Mean (m^2/ha)	14.46	17.08	21.41	Total ($m^2 \times 10^6$)	0.67	**0.93**	**0.01**
T	Mean (count/ha)	460.07	424.04	571.37	Total (count × 10^6)	21.36	**23.16**	**0.22**
Pixel-based estimates; plot/pixel block data set								
A					Total (ha × 10^3)	30.84	40.20	21.53
V	Mean (m^3/ha)	82.46	**120.63**	99.42	Total ($m^3 \times 10^6$)	2.54	**4.85**	2.14
BA	Mean (m^2/ha)	15.97	**19.94**	17.86	Total ($m^2 \times 10^6$)	0.49	0.80	0.38
T	Mean (count/ha)	539.72	478.42	460.93	Total (count × 10^6)	16.64	19.23	9.92

[a] Bold entries indicate absolute values of ratios of deviations between plot-based and pixel-based estimates and plot-based standard errors exceed 2.25.
[b] See Table 2 footnote for land cover class definitions.

4.4. Predicting for M75 mixed conifer/deciduous class

Although precedent exists for using the proportion 0.75 as a threshold for distinguishing among the conifer, deciduous, and mixed classes, accuracies associated with this threshold were small (Fig. 4). Although accuracies for the three forest classes, C75, D75, and M75, considered collectively, decreased as the threshold increased, accuracies for the M75 class increased as the threshold increased, possibly as a result of the increase in the number of reference pixels assigned to the M75 class. Nevertheless, accuracies for the M75 class were very small for thresholds less than approximately 0.85. Although the species composition classes are considered discrete classes of a categorical response variable, in fact they are defined with respect to an arbitrary threshold of a continuous variable, BA proportion. Therefore, difficulty predicting the class of pixels whose BA proportions are near the threshold should be expected. In addition, the difficulty is exacerbated when more classes are used and when the proportions of reference pixels assigned to individual classes are small. When 0.75 is used as BA proportion threshold, the proportion of reference pixels assigned to the mixed class was only approximately 0.15.

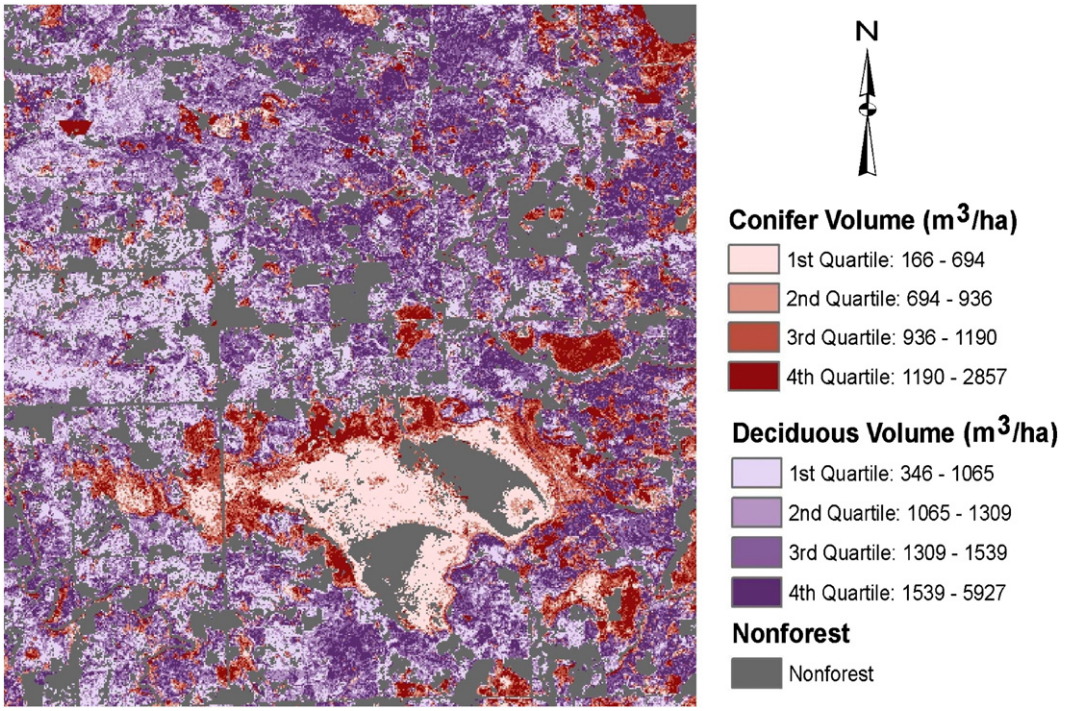

Fig. 3. Predictions of volume (m^3/ha) within conifer (C50) and deciduous (D50) species composition classes for a 15-km × 15-km AOI using the SPEC12 spectral data, the subplot/pixel data set, and the two-step algorithm with the multinomial logistic regression model in the first step.

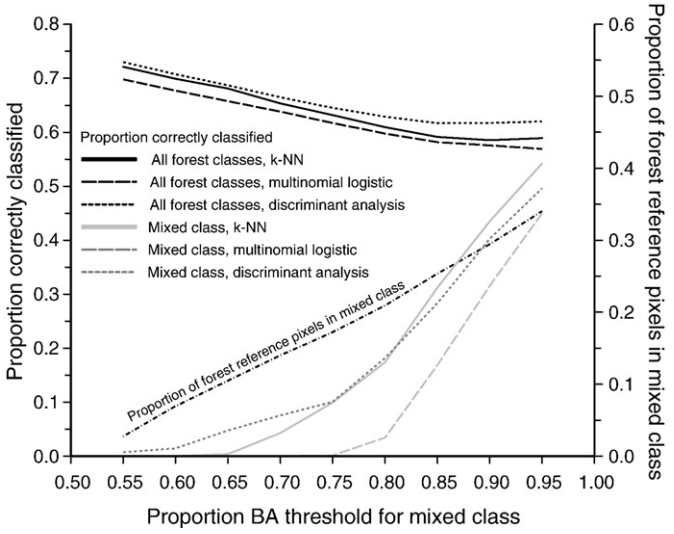

Fig. 4. Proportions of forest reference set pixels in mixed class and proportions of correctly classified forest and mixed reference set pixels using SPEC12 spectral data and subplot/pixel data set.

4.5. Sustainability and habitat applications

The two-step k-NN algorithm with the multinomial logistic regression model in the first step produced species composition and forest structure maps that are useful for both wildlife and sustainability applications. The similarity between the pixel-based and plot-based estimates of conifer and deciduous areas and means of *V*, BA, and *T* suggests that the maps are unbiased at the 15-km × 15-km scale; further, the maps and estimates obtained from them are compatible. For sustainability analyses, the maps produce areal estimates of conifer- and deciduous-dominated forest, and *V* and *T* estimates by these species composition classes. For avian habitat analyses, the maps depict the relative positions and interspersion of conifer- and deciduous-dominated forest areas (Fig. 3). In addition, the combination of *V* and *T* predictions may be used to infer locations of forest areas with smaller or larger trees. Finally, the maps are also useful for characterizing small and irregularly shaped areas such as riparian zones for which insufficient numbers of plots are available for reliable plot-based estimates.

5. Conclusions

Four specific conclusions can be drawn from this study. First, the single-step k-NN algorithm produced non-zero estimates of mean *V*, BA, and *T* for areas predicted to be in the NF class and erroneously small estimates of mean *V*, BA, and *T* for some forest classes. Therefore, without current and accurate land cover information, the single-step k-NN algorithm may produce biased estimates for at least some individual land cover classes.

Second, the problems associated with the single-step algorithm can be avoided by using a two-step algorithm featuring a multinomial logistic regression model in the first step to predict land cover class and a k-NN technique in the second step to predict *V*, BA, and *T* subject to the constraint that all nearest neighbors in the second step must be of same land cover class as that predicted in the first step. Pixel-based estimates of means and totals within the NF, C50, and D50 classes obtained using this algorithm with the subplot/pixel data set were within the standard of 2.25 standard errors of the corresponding plot-based estimates with one very minor exception; the pixel-based estimate of A for D50 class deviated from the plot-based estimate by 2.3 standard errors. When a mixed class was included for construction of the NF, C75, D75, and M75 classes, considerable confusion between the deciduous (D75) and mixed (M75) classes resulted, and totals for all variables were poorly estimated.

Third, although the single-step algorithm produced slightly better estimates of mean V, BA, and T over all land cover classes for data aggregated for the six AOIs, the two-step algorithm with the multinomial logistic regression model in the first step produced acceptable estimates for the six AOIs collectively and the best estimates for the 15-km × 15-km AOIs individually.

Fourth, comparisons of results for the SPEC18 and SPEC12 spectral data revealed only minor differences, whereas comparisons of the subplot/pixel and plot/pixel block data sets suggested the former was superior for areal estimation.

The general conclusions are threefold: (1) the single-step algorithm may be satisfactory for users who are not interested in estimates by land cover class, who are willing to accept discrepancies such as non-zero forest estimates for pixels predicted to be forested and attribute them to classification error, who are concerned about loss of information resulting from classification in the first step of the two-step algorithm, or who prefer continuous class probability estimates rather than categorical class predictions; (2) the two-step algorithm with the multinomial logistic regression model in the first step using the subplot/pixel data set circumvented the discrepancies documented with the single-step algorithm, produced pixel-based estimates of means over all land cover classes that were within 2.0 plot-based standard errors of plot-based estimates, and produced pixel-based estimates of mean and total V, BA, and T within the NF, C50, and D50 land cover classes that were within or very close to the standard of 2.25 plot-based standard errors of plot-based estimates; and (3) additional research and perhaps additional ancillary data are necessary to produce acceptable estimates for land cover classes that include a mixed conifer/deciduous class.

Acknowledgements

The author acknowledges the assistance of Lisa G. Mahal in constructing maps and the assistance of Daniel J. Kaisershot in providing aerial photography for visual map checking. Also, the suggestions of two anonymous reviewers are acknowledged.

References

Agresti, A. (2007). *An introduction to categorical data analysis.* Hoboken, NJ: Wiley-Interscience.

Ambuel, B., & Temple, S. A. (1983). Area dependent changes in the bird communities and vegetation of southern Wisconsin forest. *Ecology*, *64*, 1057–1068.

Baffeta, F., Fattorini, L., Franceschi, S., & Corona, P. (2009). Design-based approach to k-NN technique for coupling field and remotely sensed data in forest surveys. *Remote Sensing of Environment*, *113*, 463–475 (this issue).

Bechtold, W. A., & Patterson, P. L. (Eds.). (2005). *The Enhanced Forest Inventory and Analysis program — national sampling design and estimation procedures.* General Technical Report SRS-80. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 85 p.

Bentz, B. J., & Endreson, D. (2003). Evaluating satellite imagery for estimating mountain pine beetle-cause lodgepole pine mortality: current status. In T. L. Shore, J. E. Brooks, & J. E. Stone (Eds.), *Mountain pine beetle symposium: challenges and solutions.* Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre, Information Report BC-C-399, Victoria, BC. 298 p.

Bossard, M., Feranec, J., & Otahel, J. (2000). CORINE land cover technical guide — addendum 2000. *Technical Report No. 40. European Space Agency. Copenhagen.*

Buongiorno, J., Dahir, S., Lu, H. C., & Lin, C. R. (1994). Tree size diversity and economic returns in uneven-aged forest stands. *Forest Science*, *40*, 83–103.

Calef, M. P., McGuire, A. D., Epstein, H. E., Rupp, T. S., & Shugart, H. H. (2005). Analysis of vegetation distribution in Interior Alaska and sensitivity to climate change using a logistic regression model. *Journal of Biogeography*, *32*, 863–878.

Chen, J., & Klinka, K. (2003). Aboveground production of mixed-species western hemlock and redcedar stands in British Columbia. *Forest Ecology and Management*, *184*, 46–55.

Chen, J., Klinka, K., Mathey, A. H., Wang, X., Varga, P., & Chourmouzis, C. (2003). Are mixed-species stands more productive than single-species stands: an empirical test of three forest types in British Columbia and Alberta. *Canadian Journal of Forest Research*, *33*, 1227–1237.

Chirici, G., Barbati, A., Corona, P., Marchetti, M., Travaglini, D., Maselli, F., & Bertini, R. (2008). Non-parametric and parametric methods using satellite imagery for

estimating growing stock volume in alpine and Mediterranean forest ecosystems. *Remote Sensing of Environment*, *112*, 2686–2700.
Clark, K., Euler, D., & Armstrong, E. (1983). Habitat association of breeding birds in cottage and natural areas of central Ontario. *Wilson Bulletin*, *95*, 77–96.
Cochran, W. G. (1977). *Sampling techniques.* New York: Wiley.
Cody, M. L. (1975). Towards a theory of continental species diversities. In N. L. Cody, & J. M. Diamond (Eds.), *Ecology and evolution of communities* (pp. 214–257). Cambridge, MA: Belknap.
Collins, S. L., James, F. C., & Rixxer, P. G. (1982). Habitat relationships of wood warblers (Parulidae) in north central Minnesota. *Oikos*, *39*, 50–58.
Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, *37*, 35–46.
Coops, N. C., Wulder, M. A., & White, J. C. (2006). Integrating remotely sensed and ancillary data sources to characterize a mountain pine beetle infestation. *Remote Sensing of Environment*, *105*, 83–97.
COST E43 (2007). *Harmonization of the national forest inventories of Europe.* http://www.metla.fi/eu/cost/e43/. (Accessed March 2008).
Crist, E. P., & Cicone, R. C. (1984). Application of the tasseled cap concept to simulated Thematic Mapper data. *Photogrammetric Engineering and Remote Sensing*, *50*, 343–352.
D'Eon, R. G., & Watt, W. R. (1994). A forest habitat suitability matrix for northeastern Ontario. *Ontario Ministry of Natural Resources, Northeast Science and Technology, Technical Manual TM-004. Timmins, Ontario, Canada*.
DesGranges, J. L. (1980). Avian community structure of six forests stands in La Maurice National Park, Quebec. *Canadian Wildlife Service Occasional Paper, Number 41.*
Edgar, C. B., & Burk, T. E. (2001). Productivity of aspen forests in northeastern Minnesota, U.S.A., as related to stand composition and canopy structure. *Canadian Journal of Forest Research*, *31*, 1019–1029.
Ellenwood, J., & Krist, F. (2007). Building a nationwide 30-meter forest parameter dataset for forest health risk assessments. In M. DeShayes (Ed.), *Forests and remote sensing: methods and operational tools. Proceedings of ForestSat 2007. 5–7 November 2007, Montpellier, France. Cemegref, France*.
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.
Girard, C., Darveau, M., Savard, J. -P. L., & Huot, J. (2004). Are temperate mixedwood forests perceived by birds as a distinct forest type? *Canadian Journal of Forest Research*, *34*, 1895–1907.
Haapanen, R., Ek, A. R., Bauer, M. E., & Finley, A. O. (2004). Delineation of forest/non-forest land use classes using nearest neighbor methods. *Remote Sensing of Environment*, *89*, 265–271.
Hansen, M. H., & Hahn, J. T. (1992). Determining sotcking, forest type, and stand size class from forest inventory data. *Northern Journal of Applied Forestry*, *9*(3), 82–89.
Hansen, M. H., Madow, W. G., & Tepping, B. J. (1983). An evaluation of model-dependent and probability-sampling inferences in sample surveys. *Journal of the American Statistical Association*, *78*(384), 776–793.
Hobson, K. A., & Bayne, E. (2000). The effects of stand age on avian communities in aspen-dominated forests of central Saskatchewan, Canada. *Forest Ecology and Management*, *136*, 121–134.
Homer, C., Huang, C., Yang, L., Wylie, B., & Coan, M. (2000). Development of a 2001 national land-cover database for the United States. *Photogrammetric Engineering & Remote Sensing*, *70*, 829–840.
Hudak, A. T., Crookston, N. L., Evans, J. S., Hall, O. E., & Falkowski, M. J. (2008). Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. *Remote Sensing of Environment*, *112*, 2232–2245.
James, F. C., & Warner, N. O. (1982). Relationships between temperate forest bird communities and vegetation structure. *Ecology*, *63*, 159–171.
Katila, M., & Tomppo, E. (2002). Stratification by ancillary data in multisource forest inventories employing k-nearest neighbour estimation. *Canadian Journal of Forest Research*, *32*, 1548–1561.
Kauth, R. J., & Thomas, G. S. (1976). The Tasseled Cap — a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. *Proceedings of the symposium on machine processing of remotely sensed data* (pp. 41–51). West Lafayette, IN: Purdue University.
Kendall, M. G., & Buckland, W. R. (1982). *A dictionary of statistical terms*, 4th *ed*. New York: Longman Group, Ltd.
Kerr, C. (1999). Comparative productivity of monocultures and mixed-species stands. In M. J. Kelty (Ed.), *The ecology and silviculture of mixed species stands* (pp. 125–141). Netherlands: Kluwer Academic Publishers.
Kimes, D. S., Ranson, K. J., Sun, G., & Blair, J. B. (2006). Predicting LIDAR measured forest vertical structure from multi-angle data. *Remote Sensing of Environment*, *100*, 503–511.
Kirk, D. A., Diamond, A. W., Hobson, K. A., & Smith, A. R. (1996). Breeding bird communities of the western and northern Canadian boreal forest: relationship to forest type. *Canadian Journal of Zoology*, *74*, 1749–1770.
Koukal, T., Suppan, F., & Schneider, W. (2007). The impact of radiometric calibration on the accuracy of kNN-predictions of forest attributes. *Remote Sensing of Environment*, *110*, 431–437.
Koutsias, N., & Karteris, M. (1998). Logistic regression modelling of multitemporal Thematic Mapper data for burned area mapping. *International Journal of Remote Sensing*, *19*, 3499–3514.
Koutsias, N., & Karteris, M. (2000). Burned area mapping using logistic regression modeling of a single post-fire Landsat-5 Thematic Mapper image. *International Journal of Remote Sensing*, *21*, 673–687.
Kraemer, H. C. (1983). Kappa coefficient. In S. Kotz, & N. L. Johnson (Eds.), *Encyclopedia of statistical sciences*, *vol. 4*. New York: Wiley.
Lachenbruch, P. A., & Mickey, M. R. (1986). Estimation of error rates in discriminant analysis. *Technometrics*, *10*, 1–11.
LeMay, V. M., Maedel, J., & Coops, N. C. (2008). Estimating stand structural details using nearest neighbor analyses to link ground data, forest cover maps, and Landsat imagery. *Remote Sensing of Environment*, *112*, 2578–2591.
MacArthur, R. H., & MacArthur, J. W. (1961). On bird species diversity. *Ecology*, *42*, 594–598.
Magnussen, S., McRoberts, R. E., & Tomppo, E. (2009). Model based mean square error estimators for k-nearest neighbour predictions and applications using remotely sensed data for forest inventories. *Remote Sensing of Environment*, *113*, 476–488 (this issue).
Mallinis, G., Koutsias, N., Makras, A., & Karteris, M. (2003). Forest parameters estimation in a European Mediterranean landscape using remotely sensed data. *Forest Science*, *50*, 450–460.
Mård, H. (1996). The influence of birch shelter (*Betula* spp.) on the growth of young stands of *Picea abies*. *Scandinavian Journal of Forest Research*, *11*, 343–350.
Martin, M. E., Newman, S. D., Aber, J. D., & Congalton, R. G. (1998). Determining forest species composition using high spectral resolution remote sensing data. *Remote Sensing of Environment*, *65*, 249–254.
McLachlan, G. J. (2004). *Discriminant analysis and statistical pattern recognition.* New York: Wiley.
McRoberts, R. E. (2006). A model-based approach to estimating forest area. *Remote Sensing of Environment*, *103*, 56–66.
McRoberts, R. E. (2009). Diagnostic tools for nearest neighbors techniques when used with satellite imagery. *Remote Sensing of Environment*, *113*, 489–499 (this issue).
McRoberts, R. E., Nelson, M. D., & Wendt, D. G. (2002). Stratified estimation of forest area using satellite imagery, inventory data, and the *k*-Nearest Neighbors technique. *Remote Sensing of Environment*, *82*, 457–468.
McRoberts, R. E., Tomppo, E. O., Finley, A. O., & Heikkinen, J. (2007). Estimating areal means and variances of forest attributes using the k-Nearest Neighbor technique and satellite imagery. *Remote Sensing of Environment*, *111*, 466–480.
McRoberts, R. E., Bechtold, W. A., Patterson, P. L., Scott, C. T., & Reams, G. A. (2005). The enhanced Forest Inventory and Analysis program of the USDA Forest Service: Historical perspective and announcement of statistical documentation. *Journal of Forestry*, *103*, 304–308.
MCPFE (2008). *Ministerial Conference on the Protection of Forests in Europe.* http://www.mcpfe.org/. (Accessed March 2008).
Moeur, M., & Stage, A. R. (1995). Most similar neighbor — an improved sampling inference procedure for natural resource planning. *Forest Science*, *41*(2), 337–359.
Montréal Process (2005). *The Montréal Process.* http://www.rinya.maff.go.jp/mpci/. (Accessed March 2008).
Murdoch, W. W., Evans, F. C., & Peterson, C. H. (1972). Diversity and pattern in plants and insects. *Ecology*, *53*, 819–829.
Niemi, G., Hanowski, P., Helle, P., Howe, R., Mönkkönen, Vernier, L., & Welsh, D. (1998). Ecological sustainability of birds in boreal forests.*Conservation Ecology*, *2*(2) available at http://consecol.org/vol2/iss2/art17/
Ohmann, J. L., & Gregory, M. J. (2002). Predictive mapping of forest composition and structure with direct gradient analysis and nearest neighbor imputation in coastal Oregon, U.S.A.. *Canadian Journal of Forest Research*, *32*, 725–741.
Ohmann, J. L., Gregory, J. J., & Spies, T. A. (2007). Influence of environment, disturbance, and ownership on forest vegetation of coastal Oregon. *Ecological Applications*, *17*, 18–33.
Önal, H. (1997). Trade-off between structural diversity and economic objectives in forest management. *American Journal of Agricultural Economics*, *79*, 1001–1012.
Pasher, J., King, D., & Lindsay, K. (2007). Modelling and mapping potential hooded warber (*Wilsonia citrina*) habitat using remotely sensed imagery. *Remote Sensing of Environment*, *107*, 471–483.
Prasad, A. M., Iverson, L. R., & Liaw, A. (2006). Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems*, *9*, 181–199.
Rempel, R. S. (2007). Selecting focal songbird species for biodiversity conservation assessment: response to forest cover amount and configuration.*Avian Conservation and Ecology*, *2*(1) available at http://www.ace-eco.org/vol2/iss1/art6
Rice, J., Anderson, B. W., & Ohmart, R. D. (1984). Comparison of the importance of different habitat attributes to avian community organization. *Journal of Wildlife Management*, *48*, 895–911.
Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1973). Monitoring vegetation systems in the great plains with ERTS. *Proceedings of the Third ERTS Symposium, NASA SP-351, Vol. 1.* (pp. 309–317)Washington, DC.: NASA.
Sachs, D. L., Sollins, P., & Cohen, W. B. (1997). Detecting landscape changes in the interior of British Columbia from 1975 to 1992 using satellite imagery. *Canadian Journal of Forest Research*, *28*, 23–36.
Salovaara, K. J., Thessler, S., Malik, R. N., & Tuomisto, H. (2005). Classification of Amazonian primary rain forest vegetation using Landsat ETM+ satellite imagery. *Remote Sensing of Environment*, *97*, 39–51.
Schuler, T. M., & Smith, F. W. (1988). Effects of species mix on size/density and leaf-area relations in Southwest pinyon/juniper woodlands. *Forest Ecology and Management*, *25*, 211–220.
Seber, G. A. F. (1984). *Multivariate observations.* New York: Wiley.
Sherry, T. W., & Holmes, R. T. (1985). Disperson patterns and habitat responses of birds in northern hardwoods forests. In M. L. Cody (Ed.), *Habitat selection in birds* (pp. 283–309). New York: Academic Press, Inc.
Smith, A. (1993). Ecological profiles of birds in the boreal forest of western Canada. In D. H. Kuhnke (Ed.), *Birds in the boreal forest.* Proceedings of a workshop held 10–12 March 1992 in Prince Albert, Saskatchewan, Northern Forestry Centre, Forestry Canada, Edmonton, Alberta, Canada.

Steele, B. M., Patterson, D. A., & Redmond, R. L. (2003). Toward estimation of map accuracy without a probability test sample. *Environmental and ecological statistics*, *10*, 333–356.

Sterba, H., & Monserud, R. A. (1995). Potential volume yield for mixed-species Douglas fir stands in the northern Rocky Mountains. *Forest Science*, *41*, 531–545.

Temesgen, H., LeMay, V. M., Froese, K. L., & Marshall, P. L. (2003). Imputing tree-lists from aerial attributes for complex stands of south-eastern British Columbia. *Forest Ecology and Management*, *177*, 277–285.

Thessler, S., Ruokolainen, K., Tuomisto, H., & Tomppo, E. (2005). Mapping gradual landscape-scale floristic changes in Amazonian primary rain forests by combining ordination and remote sensing. *Global Ecology and Biogeography*, *14*, 315–325.

Tomppo, E. (1992). Satellite image aided forest site fertility estimation for forest income taxation.*Acta Forestalia Fennica*, *229*, 70 p.

Tomppo, E. (2006). The Finnish National Forest inventory. In A. Kangas, & M. Maltamo (Eds.), *Forest inventory: methodology and applications* Dordrecht, The Netherlands: Springer.

Tomppo, E., & Halme, M. (2004). Using coarse scale forest variables as ancillary information and weighting of variables in k-NN estimation: a genetic algorithm approach. *Remote Sensing of Environment*, *92*, 1–20.

Tomppo, E., Korhonen, K. T., Heikkinen, J., & Yli-Kojola, H. (2001). Multi-source inventory of the forests of the Hebei Forestry Bureay, Heilongjian, China. *Silva Fennica*, *35*(3), 309–328.

Tomppo, E. O., Gagliano, C., De Natale, F., Katila, M., & McRoberts, R. E. (2009). Predicting categorical variables using an improved k-Nearest Neighbors estimator. *Remote Sensing of Environment*, *113*, 500–517 (this issue).

Tomppo, E., Olsson, H., Ståhl, Nilsson, M., Hagner, O., & Katila, M. (2008). Combining national forest inventory field plots and remote sensing data for forest databases. *Remote Sensing of Environment*, *112*, 1982–1999.

Vogelmann, J. E., Howard, S. M., Yang, L., Larson, C. R., Wylie, B. K., & Van Driel, N. (2001). Completion of the 1990s National land cover data set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. *Photogrammetric Engineering and Remonte Sensing*, *67*, 581–587.

Wickham, J. D., Stehman, S. V., Smith, J. H., & Yang, L. (2004). Thematic accuracy of the 1992 national land-cover data for the western United States. *Remote Sensing of Environment*, *91*, 452–468.

Willson, M. F. (1974). Avian community organization and habitat structure. *Ecology*, *55*, 1017–1029.

Xu, M., Watanachaturaporn, P., Varshney, P. K., & Arora, J. K. (2005). Decision tree regression for soft classification of remote sensing data. *Remote Sensing of Environment*, *97*, 322–336.

Yang, L., Homer, C. G., Hegge, K., Huang, C., Wylie, B., & Reed, B. (2001). A Landsat 7 scene selection strategy for a National Land Cover Database. *Proceedings of the IEEE 2001 International Geoscience and Remote Sensing Symposium, Sydney, Australia, CD-ROM, 1 disc*.

Young, L., Betts, M. G., & Diamond, A. W. (2005). Do Blackburnian Warblers select mixed forest? The importance of spatial resolution in defining habitat. *Forest Ecology and Management*, *214*, 358–372.