Object-based Analysis of Ikonos-2 Imagery for Extraction of Forest Inventory Parameters

Michael S. Chubey, Steven E. Franklin, and Michael A. Wulder

Abstract

A method is presented for deriving forest inventory information from Ikonos-2 imagery based on the analysis of image objects rather than more conventional pixel-based image analysis approaches. For a 77 km^2 study area in southwestern Alberta, Canada, image objects representing homogeneous landscape components were delineated from Ikonos-2 data using an image segmentation routine. Decision tree statistical analyses were used to identify correlations between metrics derived from spectral and spatial properties of the image objects and field-derived samples of individual forest inventory parameters. The strongest relationships were observed for classes of discrete land-cover types, species composition, and crown closure.

Introduction

Remote sensing has played an important role in forestry for several decades, particularly as a tool for acquiring information about the location, extent, composition, and structure of the forest resource as part of industrial forest inventories. Forest inventory information is required for all levels of forest management and planning, and is becoming increasingly important for other applications including conservation and environmental management (Leckie and Gillis, 1995). The conventional approach of deriving forest inventory information through manual interpretation of aerial photographs works well for traditional timber management applications, but is costly in terms of time and resources, delivers inconsistent results, and often does not fulfil information requirements for non-timber uses in terms of detail, accuracy, and currency (King, 2000; Wulder, 1998). It has long been anticipated that manual photo-based forest inventory procedures will be superseded by semi-automated, digital remote sensing approaches that promise greater efficiency and consistency (Bergen *et al.*, 2000; Caylor, 2000; Pitt *et al.*, 1997). However, the limitations thus far of conventional pixel-based remote sensing techniques to realize these expectations suggests that an entirely different approach to digital image analysis, such as one based on *image objects* rather than pixels, may be needed in order to make an imagery-derived forest inventory a reality at the operational level.

Satellite-based remote sensing offers several desirable characteristics from a forest inventory mapping perspective including stable platforms, large footprint sizes, and regular data collection cycles. Regional-scale estimates of some forest variables have been obtained from medium spatial resolution satellites (e.g., pixel sizes $\geq$ 10 meters on a side) such as SPOT (10 to 20 m) and Landsat (15 to 30 m); however, high spatial resolution sensors (e.g., $\leq$ 4 m pixels) are more appropriate for forest stand-level parameterization (Wulder *et al.*, 2004). Although high spatial detail imagery from the relatively new generation of high-resolution commercial satellites (e.g., Ikonos (1 to 4 m^2), QuickBird (0.6 to 2.8 m^2)) have potential to provide information suitable for modern forest inventory requirements (Bergen *et al.*, 2000; Green, 2000; Waring and Running, 1999), extraction of forest information from high spatial resolution digital data poses unique challenges that will require new interpretation procedures (Culvenor, 2003).

A fundamental problem for incorporating digital imagery into the forest inventory process is that for a given forest stand, spectral response is represented in digital imagery as a series of discrete pixels covering a wide range of spectral values; yet, for inventory purposes, the stand is interpreted as a single homogeneous polygon (Hall, 2003). One solution to this dilemma is to aggregate the individual pixels representing the forest stand into an *image object* represented spectrally as the combined response of all underlying pixels. The image objects, rather than the underlying pixels, become the carriers of image information and form the basic units of subsequent analysis.

Object-based analysis of multispectral imagery was introduced early on in the remote sensing literature (Ketting and Landgrebe, 1976); however, the object-based approach has largely been ignored in favour of pixel-based methods, which have been easier to implement (Lobo, 1997). Several object-based image analysis techniques have been used successfully for forest information extraction purposes (Hay *et al.*, 1996; Pekkarinen, 2002; St-Onge and Cavayas, 1997); but application of these techniques has been limited, likely due to the complex nature of some of the procedures, or due to specialized custom software requirements (Flanders *et al.*, 2003). The recent release of new commercial object-based image analysis software such as eCognition® (Definiens Imaging, 2002) and Feature Analyst® (Visual Learning Systems, Inc., 2002), coupled with the limitations of pixel-based methods to provide acceptable results for many applications, has led to new interest in object-based image analysis of remotely sensed data (Hay *et al.*, 2003; Thomas

Michael S. Chubey is with the Department of Earth and Atmospheric Science, University of Alberta, Edmonton, AB T6G 2E3 (mchubey@ualberta.ca).

Steven E. Franklin is with the Office of the Vice-President Research, University of Saskatchewan, Saskatoon, SK S7N 4J8 (steven.franklin@usask.ca).

Michael A. Wulder is with the Canadian Forest Service, Pacific Forestry Centre, Victoria, BC V8Z 1M5 (mike.wulder@ncan.gc.ca).

Photogrammetric Engineering & Remote Sensing
Vol. 72, No. 4, April 2006, pp. 383–394.

0099-1112/06/7204–0383/$3.00/0

et al., 2003; Wulder and Seemann, 2003); however, protocols for implementing these new tools for individual applications, or select inventory attributes, are needed.

One of the advantages of object-based image analysis is the multitude of additional information that can be derived from image objects versus the amount of information available from individual pixels. A pixel typically contains a vector of information representing each band or layer in a data set. In the case of digital imagery, the spectral response information is related as digital numbers (DN). In contrast, image objects are composed of multi-pixel groups, enabling the calculation of aggregative statistics, such as mean and standard deviation, from an object's underlying DNs. In addition to spectral-based information, information based on object size, shape, and context, can be calculated, as can information pertaining to an object's sub- or super-objects if a multi-level image object hierarchy has been created.

In order to take full advantage of the potentially rich set of available image object-derived information, a classification approach capable of processing potentially exhaustive numbers of inputs is required. One such approach is based on the use of decision tree data analysis. Decision trees are conceptually simple yet powerful data analysis tools suitable for identifying relationships between variables in multivariate data sets (Breiman *et al.*, 1984; Crawley, 2002; Hastie *et al.*, 2001; Venables and Ripley, 2002). Tree models identify relationships between a single response (dependent) variable and multiple explanatory (independent) variables through a process known as *binary recursive partitioning*, where the data are split repeatedly into increasingly homogeneous groups, or nodes, using combinations of explanatory variables that best distinguish the variation of the response variable (Breiman *et al.*, 1984). The decision tree algorithm examines all possible splits of the data and selects a split threshold value of the explanatory variable that produces maximum dissimilarity, or deviance, between the resulting subsets. Decision trees have been applied to the analysis of remotely sensed data in several contexts including delineating forest species (Sugumaran *et al.*, 2003), estimating percent tree cover (Hansen *et al.*, 2002), estimating tree canopy density (Huang *et al.*, 2001), and generating membership rules for land-cover classification (Brown de Colstoun *et al.*, 2003).

In the present study, we develop and test methods for extracting forest inventory information from high spatial resolution satellite imagery using commercially-available image object analysis software (Definiens Imaging, 2002) and decision tree statistical analysis. In particular, we identify relationships between some typical forest inventory attributes (species type, percent species, crown closure, age, and height) and image object metrics derived from Ikonos-2 imagery.

Methods

Study Area

The study was undertaken in a mature forest ecosystem located in the foothills of the Rocky Mountains in southwestern Alberta, Canada (Figure 1). The study area measures approximately 11 km by 7 km and encompasses several land-cover types including forest, shrubland, grassland, lakes, rivers, and bare rock. The terrain is variable with elevations ranging from approximately 1,200 m ASL on the alpine meadows to more than 2,000 m ASL along the higher mountain ridges. Forests within the study area consist mainly of coniferous species including lodgepole pine (*Pinus contorta*) and white spruce (*Picea glauca*) occurring in pure and mixed stands. Deciduous forests consisting mainly of trembling aspen (*Populus tremuloides*) occurring in pure stands and mixed with conifers are present at lower elevations.

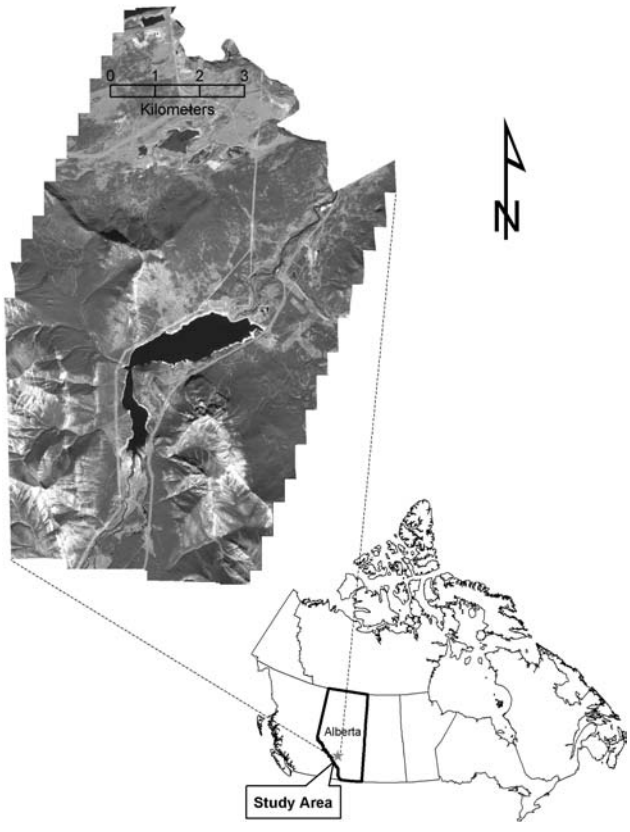

Figure 1. Location and extent of study area in southwestern Alberta, Canada. Study area extent is depicted by Ikonos-2 panchromatic imagery.

Data Acquisition

Digital image data were acquired over the study area by the Ikonos-2 satellite on 15 August 2001. The Ikonos-2 data set consisted of single-band panchromatic imagery (450 to 900 nanometers (nm)) with a spatial resolution of 1×1 m, and 4-band multispectral imagery with a spatial resolution of 4×4 m divided into the following spectral bands: blue (445 to 516 nm), green (506 to 595 nm), red (632 to 698 nm), and near-infrared (757 to 853 nm). Ancillary data for this study consisted of forest inventory information from the Alberta Vegetation Inventory (AVI), and topographic information in the form of a digital elevation model (DEM) with an original spatial resolution of 10×10 m. The AVI is a digital inventory developed to identify the type, extent, and conditions of vegetation in forested regions of Alberta (Alberta Sustainable Resource Development, 1991). Information including species composition, crown closure, stand height, and stand age is derived primarily though manual aerial photo-interpretation and stored in a geographic information system (GIS) database. Limited field surveys are used to verify the reliability of the AVI photo interpretation to an accuracy standard of 80 percent.

Preliminary processing of the image data included orthorectification, as well as subsetting of the imagery to the geographic extent shown in Figure 1. Orthorectification was based on ground control points obtained in the field using a differential global positioning system (GPS) receiver resulting in average rectification errors of 1.07 pixels for the panchromatic imagery and 0.27 pixels for the multispectral imagery.

Field plot locations were selected according to a proportional stratified sampling scheme in an attempt to characterize the diversity within the study area in terms of forest

species composition and structural attributes. Initially, AVI polygons representing pure stands of each of the three major forest species in the study area (lodgepole pine, white spruce, and trembling aspen) were identified through database queries. Pure stands were defined as polygons where the leading species comprised ≥ 90 percent of the land-cover according to the AVI label. AVI labels were used to further stratify the selected forest polygons on the basis of crown closure, age, and height. Sample polygons (i.e., polygons in which sample plots would be established) were selected on a proportional basis; first, according to relative abundance of each of the three major forest species in the study area (according to the AVI data), and then according to crown closure, age, and height. In all, 51 sample polygons were selected representing 24 pine-, 15 spruce-, and 12 aspen-dominated stands and covering the full range of crown closure, age, and height for each species. Three field plots were located at random within each sampling polygon for a total of 153 plots. Variable radius plot surveys, using a basal area factor (BAF) of 4, were conducted at each selected plot location. Per-plot estimates of standard forest inventory parameters including percent species composition, percent crown closure, tree height, and tree age were derived from measurements recorded at each field plot.

Image Object Information

The Ikonos-2 images were partitioned into image objects using the eCognition® software package (Definiens Imaging, 2002). Generation of image objects was achieved through an image segmentation procedure in eCognition® termed *multiresolution segmentation*. Multiresolution segmentation partitions an image into homogeneous multi-pixel regions based on several user-defined parameters. The user can influence the output of the segmentation process through specification and weighting of input data and definition of parameters affecting the size, spectral homogeneity, spatial homogeneity, and shape of the resulting image objects. The underlying proprietary segmentation algorithm is described as a region-merging technique in which individual pixels are merged into small objects, followed by successive iterations in which small objects are incrementally merged into larger ones in such a way that heterogeneity of the resulting image objects is minimized. The merging process continues until a threshold derived from the user-defined parameters is reached (Baatz and Schäpe, 2000).

Decisions regarding selection and weighting of inputs to the segmentation process were made based on spectral and spatial characteristics of individual Ikonos-2 bands, correlation analysis of the Ikonos-2 multispectral data, and experimentation. In terms of general spectral characteristics, near-infrared reflectance is most often associated with the ability to discriminate changes in vegetation, although all of the spectral bands recorded by the Ikonos-2 sensor are potentially useful for this purpose to some degree (Jensen, 2000). In terms of spatial characteristics, the Ikonos-2 panchromatic band, with a spatial resolution of 1 m captures more spatial detail than the 4 m multispectral data. Although data layers of differing spatial resolutions can be used simultaneously within eCognition®, input layers are automatically re-sampled to the spatial resolution of the highest resolution input layer. Preliminary correlation analysis of the image data revealed a high degree of correlation among the three bands representing the visible portion of the spectrum (bands 1 through 3), but much lower correlation between the near-infrared band (band 4) and the visible bands. In response to the above considerations and observations, the weightings of the multispectral bands were arranged such that the three visible bands were assigned equal weights, and the sum of the weights assigned to the three visible bands combined equalled the weighting assigned to the near-infrared band. Additionally, the panchromatic band also was assigned equal weighting to that of the near-infrared band. The relative weights of the segmentation input layers are summarized as follows:

$$(w_{bnd1} + w_{bnd2} + w_{bnd3}) = w_{bnd4} = w_{pan} \quad (1)$$

where w = weighting factor; bnd = multispectral band; and pan = panchromatic band.

The segmentation input/weighting scenario described above was tested against several other input/weighting combinations, including assigning equal weights to all bands, and omitting individual bands. Each input scenario was evaluated on its ability to delineate meaningful landscape components (e.g., homogeneous forest stands and clearings) based on visual inspection of the segmentation output; however, the input scenario described in Equation 1 produced the most satisfactory segmentation results. Settings for the remaining eCognition® segmentation parameters (scale, color, shape, smoothness, and compactness) were determined through a similar experimental process.

The objectives of the segmentation phase were: (a) to delineate homogeneous forest stand components; (b) to isolate field plot locations within these areas; and (c) to characterize the resulting image objects in terms of image texture. These tasks were accomplished using a multi-step segmentation strategy. The imagery were segmented initially at a resolution that was coarse enough to aggregate groups of pixels representing homogeneous areas into recognizable forest stand objects while preserving appropriate within-stand variability (similar to what might be expected in an aerial photo interpretation stratification) (Plate 1a). This initial image object level served as a means of calibrating the input parameters in such a way as to avoid arbitrary delineations as much as possible (we used the following settings for scale (166), color (0.8), shape (0.2), smoothness (0.7), and compactness (0.3)). A second image object level was created by sub-segmenting the original image objects using the same input parameter scenario, but with a scale parameter setting 50 percent finer than the original. This second segmentation level was designed to isolate homogeneous areas immediately surrounding the field plot locations in order to increase the likelihood that field plot data were representative of the image objects in which the plots were located (Plate 1b).

A third image object level was created for the purpose of object-based texture analysis. The premise behind the texture analysis was that forest stand texture may be described according to the size, shape, and spatial arrangement of real forest features as captured in high spatial resolution digital imagery (Hay *et al.*, 1996). The texture objects were created by sub-segmenting the original image object level using a very fine scale parameter setting (one-eighth of the original) with the goal of delineating decisive texture structures comprised of actual forest features associated with stand-to-stand variations in species composition, stem/crown density, crown architecture, and shadowing (Plate 1c). The texture-level segmentation was based exclusively on the 1×1 m panchromatic data in order to maximize the spatial accuracy of the resulting objects.

Image object information in the form of metrics derived from the image and DEM data were calculated by eCognition® for each image object. In total, 87 metrics were selected for analysis and exported from the image objects from the field plot level segmentation (i.e., the second segmentation level). The selected image object metrics included mean, standard deviation, and ratio measures for each Ikonos-2 band and the DEM layer, metrics derived from differences among neighbouring objects and between segmentation levels, and sub-object-based texture measures.

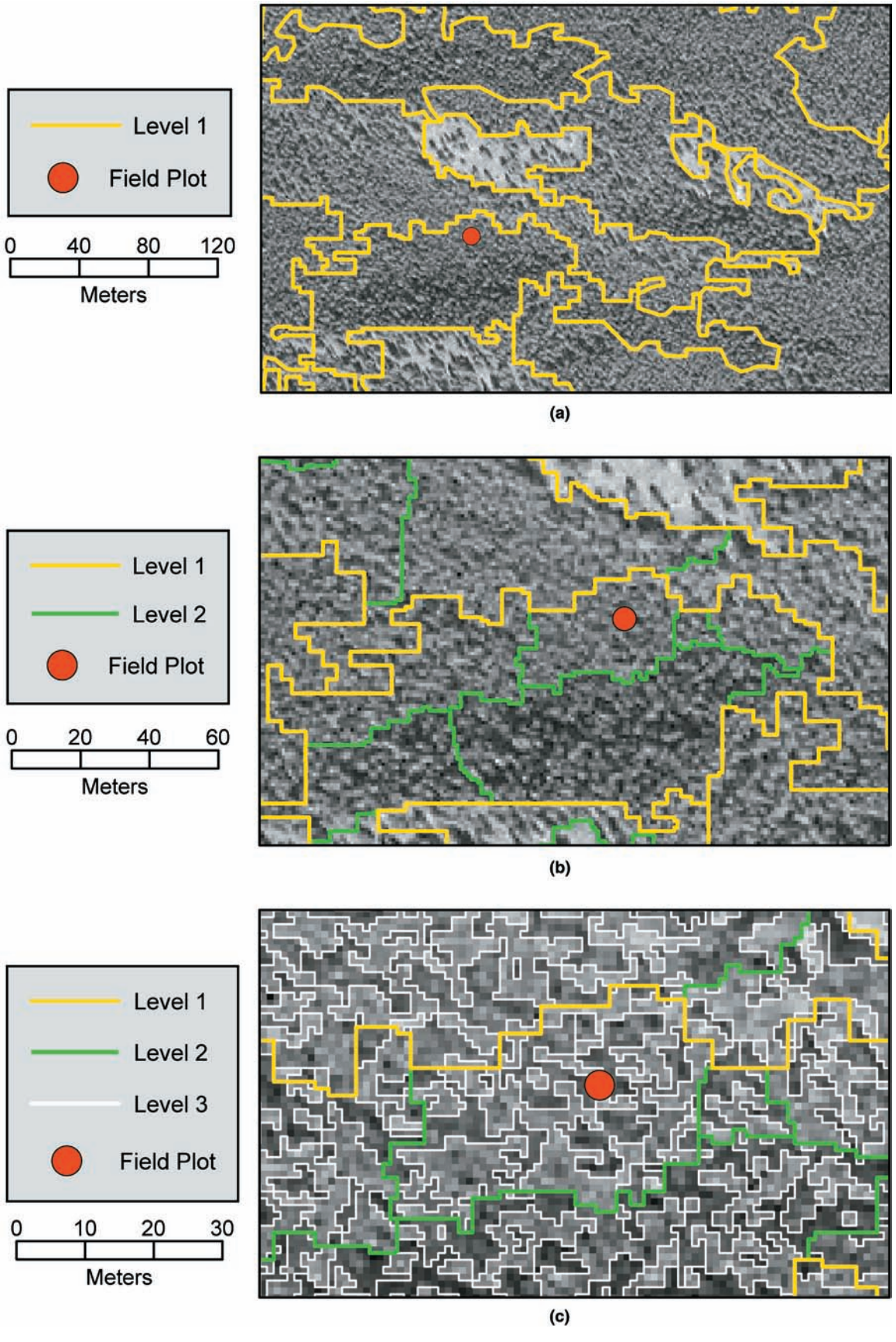

Plate 1. Image segmentation hierarchy. Initial segmentation level delineates image objects representing homogeneous forest stand components (a). A second segmentation level is created by sub-segmenting the original objects in order to better isolate field plot locations (b). The third segmentation level, created for object-based texture analysis, delineates image objects representing real features such as individual trees, tree assemblages, canopy gaps, and shadow (c).

Decision Tree Analysis

Data analysis was performed using a series of decision trees to identify relationships between the image object metrics extracted from the input data and individual forest inventory parameters. The decision tree approach was chosen for this particular study for several reasons: (a) capability for handling high-dimensional data sets (i.e., the 87 selected image object metrics); (b) capability for handling both non-continuous and continuous variables (i.e., forest species and land-cover classes, as well as continuous measurements of species composition, crown closure, height, and age); (c) nonparametric nature of the approach (i.e., no assumptions required concerning the statistical frequency distribution of the input data); (d) model transparency (straight-forward interpretation of results); and (e) ease of implementation (important from an operational perspective).

Where sufficiently abundant, the input data were stratified into three groups prior to analysis, ensuring that the class distribution of the entire data set was maintained in each subset; 60 percent of the data were selected for growing the initial trees, 20 percent were set aside for pruning, and the final 20 percent reserved for accuracy assessment (DeFries *et al.*, 1998; Prince and Steininger, 1999). Pruning was performed based on the results of a 10-fold, cross-validation routine (Venables and Ripley, 2002). Model results were tested against the 20 percent evaluation data set from which 100 randomly-chosen evaluation point locations per class were selected proportional to object size and compared with classification maps generated from each decision tree model. Accuracy statistics including producer's accuracy (omission error), user's accuracy (commission error), overall accuracy, and KIA (kappa index of agreement: a measure of overall accuracy that accounts for chance agreement), were calculated based on the construction of confusion matrices (Congalton, 1991).

A total of six separate analyses were carried out using decision trees in conjunction with image object metrics and several different subsets of the field data. In cases where the response variable is categorical, the model is called a *classification tree*. When the response variable is continuous, the model is called a *regression tree*. Classification trees were used to assess forest species separability using image objects representing pure stands (i.e., where the leading species comprised ≥90 percent of the land-cover), and to generate decision rules for classifying the image data into discrete land-cover classes using pure- and mixed-stand data. Regression trees were used to distinguish classes of four individual forest inventory parameters (percent composition, percent crown closure, stand height, and stand age) using continuous measurement field data from pine dominant stands. All decision tree analyses were performed using the tree algorithm in the S-PLUS® statistical software package (Insightful Corporation, 2002).

Results and Discussion

Classification Tree Results

Species Separability

Due to the relatively scarce amount of pure-stand field data available for the species separability analysis (81 samples in total representing 33 pine, 19 spruce, and 29 aspen stands), the pure-stand data were not subdivided for pruning and quantitative analysis as described above in the Methods Section. Rather, all available data were used as input data to the classification tree process. The resulting model indicated that it was possible to correctly separate the 81 sample objects into discrete forest species classes using just four of the 87 available image object metrics (Figure 2). A key to the image object metric codes used in Figure 2 and in all subsequent tree diagrams is presented in Table 1. The small sample size for this particular analysis has undoubtedly contributed to an overly optimistic result in terms of accuracy (i.e., 100 percent). Nonetheless, the un-pruned tree shown in Figure 2 demonstrates the utility of the combination of classification tree analysis and image object metrics for separating spectrally homogeneous forest species classes.

As illustrated in Figure 2, all 29 aspen stands were separated from the conifer stands using *ratio of NIR*, calculated as the mean near-infrared value of an image object divided by the sum of all spectral layer mean values. This finding confirms the observation made during exploratory image data analysis that separation between conifer and deciduous samples should be possible based on near-infrared reflectance. In this case, the *ratio* metric provided

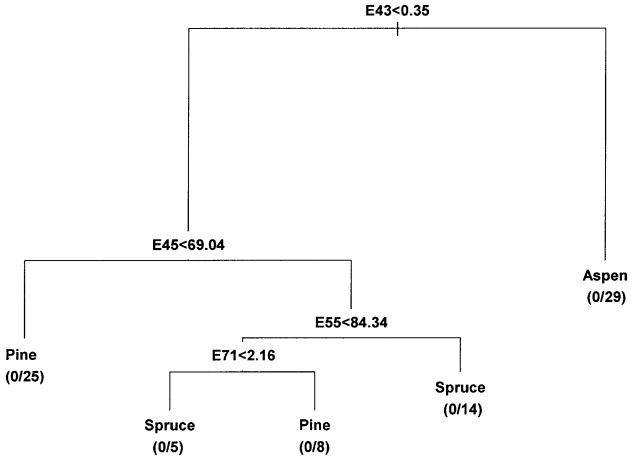

Figure 2. Classification tree: species separability. Metric codes and associated thresholds are shown above each split (refer to Table 1 for metric code definitions). Terminal nodes are labelled by class. Misclassification (error) rates are shown below each terminal node (number of incorrectly classified objects out of the total number of objects assigned to that node).

TABLE 1. KEY TO IMAGE OBJECT METRIC CODES

Metric Code	Metric Name
E1	Mean brightness
E2	Mean: Ikonos blue band
E28	Mean: Ikonos red band
E31	Mean difference to neighbors: Ikonos red
E32	Mean difference to neighbors (absolute): Ikonos red
E34	Mean difference to brighter neighbors: Ikonos red
E39	Mean of sub-objects standard deviation: Ikonos red
E41	Mean: Ikonos NIR band
E43	Ratio: Ikonos NIR band
E45	Mean difference to neighbors: Ikonos NIR band
E52	Mean of sub-objects standard deviation: Ikonos NIR band
E55	Standard deviation: Ikonos panchromatic band
E56	Ratio: Ikonos panchromatic band
E67	Mean: DEM
E68	Standard deviation: DEM
E71	Mean difference to neighbors: DEM
E87	Direction of sub-objects standard deviation

even greater separation between the aspen and conifer samples than simple mean near-infrared reflectance. The ratio operation appeared to have the effect of reducing within-class variation, due likely to effects such as topography, resulting in increased between-class separation. Twenty-five of the 33 pine stands were separated from the remaining conifer stands on the basis of lower values for *mean absolute difference to neighbors in the NIR band*. This indicates that most of the pine sample objects were surrounded by objects with similar near-infrared values; likely other pine stands. Fourteen of the 19 spruce stands were distinguished on the basis of higher *standard deviation in the panchromatic band*, which can be seen as a measure of texture based on internal spectral heterogeneity within an image object. In this case, higher internal heterogeneity within spruce stands was attributed to larger shadows cast by spruce trees versus pine trees due to differences in crown architecture. The spruce shadows were captured in the panchromatic band as groups of relatively dark pixels. Contrast between dark shadow pixels and brighter pixels representing sunlit portions of tree crowns resulted in relatively higher standard deviation for spruce image objects. The remaining 13 conifer stands were separated using *mean absolute difference to neighbors*: *DEM*, implying that of these stands, spruce were situated on flatter ground (surrounded by objects with similar DEM values) while pine stands were located on sloped or undulating terrain (greater difference in DEM values among neighboring objects). The observation that aspen was partitioned into a single terminal node using a single spectrally-derived feature suggests that this class is spectrally homogeneous, as well as distinct from the other classes. In contrast, the distinction between pine and spruce required multiple terminal nodes for each class (implying multi-modal distributions) and relied on context- and texture-related information for class separation (due to spectral overlap between classes in the training data).

Land-cover Classification

The second classification tree analysis was undertaken with the objective of generating decision rules for stratifying the study area into eight discrete land-cover classes (pine-dominated forest, spruce-dominated forest, aspen-dominated forest, shrubland, grassland, rock, sand, and water).

The original classification tree, constructed using the 60 percent training data subset, contained 13 terminal nodes. Results of the cross-validation routine, using the 20 percent pruning subset, indicated that the optimal tree size for this model was eight terminal nodes; cross-validation misclassification rates increased above this threshold. Based on the cross-validation results, the five terminal nodes with the least explanatory power were pruned from the original tree resulting in a new tree with eight terminal nodes: one node for each class of interest (Figure 3). The land-cover training data were split originally according to a *mean red band* threshold; water, spruce, pine, aspen, and shrubland absorbed more energy from the red band portion of the electromagnetic spectrum, and therefore, reflected less red band energy than grassland, sand, and rock. Aspen and shrubland were separated from water, spruce, and pine on the basis of higher *mean near-infrared* values which can be explained by the leaves of deciduous species, such as aspen and shrubs tending to reflect higher amounts of near-infrared energy than the needles of conifer species, and that water absorbs more near-infrared energy than the other targets. A second *mean near-infrared* threshold was used to separate water from spruce and pine. Spruce and pine were split subsequently based on the *ratio of the panchromatic band* to all Ikonos-2 bands. Aspen and shrubland were distinguished from one another on the basis of textural differences using

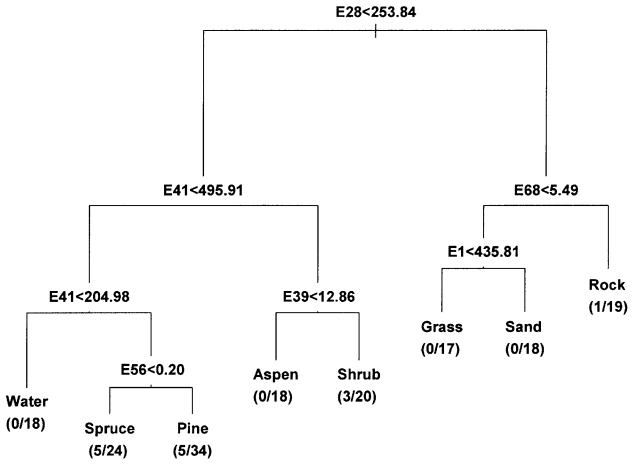

Figure 3. Classification tree: land-cover classification decision rules. Metric codes and associated thresholds are shown above each split. Terminal nodes are labelled by class. Misclassification (error) rates are shown below each terminal node.

a sub-object-based texture measure (*mean of sub-objects standard deviation: red band*). Of the remaining three non-vegetation classes, rock was separated from grass and sand based on topography (*standard deviation DEM*), and grass and sand were separated according to average tonal differences using *mean brightness*.

A comprehensive land-cover classification map of all the sub-stand-level image objects within the study area, created using decision rules derived from the image object metrics and associated thresholds identified in Figure 3, is presented in Plate 2. From a qualitative perspective, the classification map presented in Plate 2 portrays a realistic representation of the general trends observed during the field season with respect to the spatial distribution of the three main forest species and non-forest land-cover. Pine-dominated stands account for the majority of forested areas, followed by spruce- and then aspen-dominated stands. Pine stands are found on various types of sites located throughout the entire study area. Spruce occurs in patches throughout the study area as well, but especially on north-facing slopes (particularly noticeable in the south eastern corner of Plate 2) and in gullies. The association between spruce and north-facing slopes and shaded gullies is well documented, and is attributed to the relatively cool, moist conditions at those locations (Burns and Honkala, 1990). Aspen-dominated stands are located primarily along the shores of Barrier Lake in the center of the study area, on lower-elevation meadows in the northern third of the study area, and in a few recent clear-cuts.

The spatial distribution of the non-forest classes (shrubland, grassland, rock, sand, and water) depicted in Plate 2 appears realistic as well. Shrubland was defined partially by the training data as treed areas with <6 percent crown closure, and partially as woody deciduous brush. Areas classified as shrubland in Plate 2 reflect both of these definitions, namely sparsely-treed areas near the tree line at high elevation, and woody deciduous brush growing in association with aspen-dominated forest. The grassland class is restricted correctly to areas of open meadow located in the northern portion of the study area. Bare rock areas along the tops and upper slopes of the high elevation ridges are well defined, as are the sandy areas around the perime-

TABLE 2. ACCURACY ASSESSMENT RESULTS: LAND-COVER-CLASSIFICATION

Class	Pine	Spruce	Aspen	Shrub	Grass	Sand	Rock	Water	Totals
Pine	93	19	0	0	0	0	0	0	112
Spruce	7	81	0	0	0	0	0	0	88
Aspen	0	0	89	0	0	0	0	0	89
Shrub	0	0	11	94	0	0	0	0	105
Grass	0	0	0	6	87	0	0	0	93
Sand	0	0	0	0	13	100	0	0	113
Rock	0	0	0	0	0	0	100	0	100
Water	0	0	0	0	0	0	0	100	100
Totals	100	100	100	100	100	100	100	100	800
Producer's	0.93	0.81	0.89	0.94	0.87	1.00	1.00	1.00	
User's	0.83	0.92	1.00	0.89	0.94	0.88	1.00	1.00	
KIA per class	0.92	0.79	0.88	0.93	0.85	1.00	1.00	1.00	
Overall Accuracy	0.93								
Overall KIA	0.92								

ter of Barrier Lake. Barrier Lake itself is delineated correctly as water, as are several smaller lakes in the northern part of the study area and the Kananaskis River, which flows into Barrier Lake from the south, then north from the lake, and west along the top of the study area. Some shadowed areas that appear dark on the imagery were classified originally as water. These shadow objects were re-classified into a separate class based on a simple elevation threshold value.

Quantitative evaluation of the land-cover classification confirmed that classification accuracy was generally high (Table 2). Individual class accuracies were between 0.81 and 1.0, although some confusion between pine and spruce, and aspen and shrubland, was observed. This result was expected in light of the misclassification rates for these classes shown in Figure 3. Overall, however, the accuracy of the land-cover map presented in Plate 2 (i.e., producer's and user's accuracies of 0.81 to 0.93 and 0.83 to 1.0, respectively for forest classes, and an overall accuracy of 0.93) meets or exceeds the accuracy expected of most forest inventory maps (Leckie and Gillis, 1995).

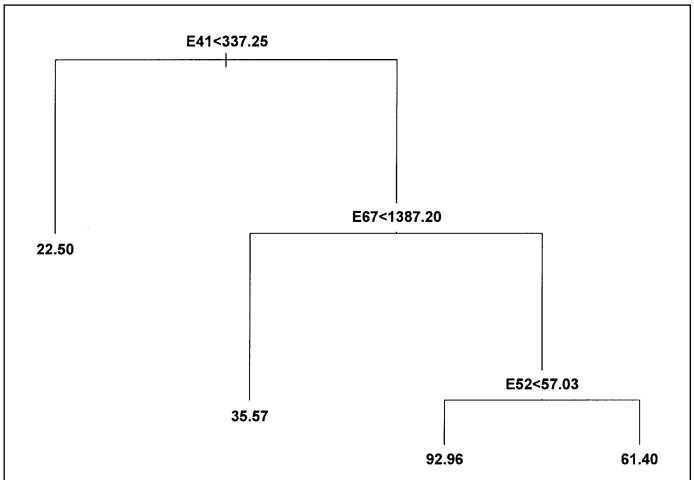

Figure 4. Regression tree: percent pine composition. Metric codes and associated thresholds are shown above each split. Values at terminal nodes represent the mean percent pine composition value for that node.

Regression Tree Results

Species Composition

The regression tree analysis of percent species composition for pine resulted in a tree with six terminal nodes. The original tree was pruned to four nodes based on cross-validation (Figure 4). The values below each of the four terminal nodes shown in Figure 4 represent the means of data clusters distinguished by the metrics identified above each split. These per-node means were used to create four percent species classes: 10 to 29 percent, 30 to 49 percent, 50 to 79 percent, and 80 to 100 percent. Class boundaries were selected based on the mid-points between per-node mean values. The 10 to 29 percent class was distinguished from the other classes based on lower *mean near-infrared* values. Explanation of this finding may be found in the forest species composition patterns observed within the study area; where pine accounts for just 10 to 29 percent of a stand's species composition, spruce was found to account for most of the remainder. Relatively pure stands of spruce tend to have lower near-infrared values than pine-dominated stands due to greater shadowing by spruce trees (Chen *et al.*, 1999). Thus, the observation of relatively low near-infrared values associated with stands with low pine concentrations can be explained partially by the presence of relatively high concentrations of spruce in the same stands. The 30 to 49 percent class was distinguished from the remaining data according to a *mean DEM* threshold, suggesting that most of the training samples with percent pine measurements in this range were located below 1,387 m, and possibly reflecting elevation-related growing conditions affecting species composition. The final two classes were defined based on *mean of sub-objects standard deviation:* NIR, an object-based texture measure. The utility of DEM- and texture-based features for estimation of forest species composition is well known, thus, the finding that these features are important for distinguishing species composition classes is consistent with previous research (Franklin, 1994; Franklin *et al.*, 2000; Hay *et al.*, 1996).

The results of the accuracy assessment of the percent pine composition regression tree model are presented in Table 3. The overall accuracy was 0.86. The highest accuracy was achieved for the 10 to 29 percent and 80 to 100 percent composition classes. These were also the classes that encompassed the majority of the continuous percent pine composition measurements in the training data; relatively fewer field observations were made for pine in proportions $\geq$30 percent but $<$80 percent. It should also be noted that the four percent pine classes identified by the model output do not correspond with the species composition classes used by most forest inventories; species composition typically is estimated to the nearest ten percentile. Delineation of ten percentile species composition classes using the methods

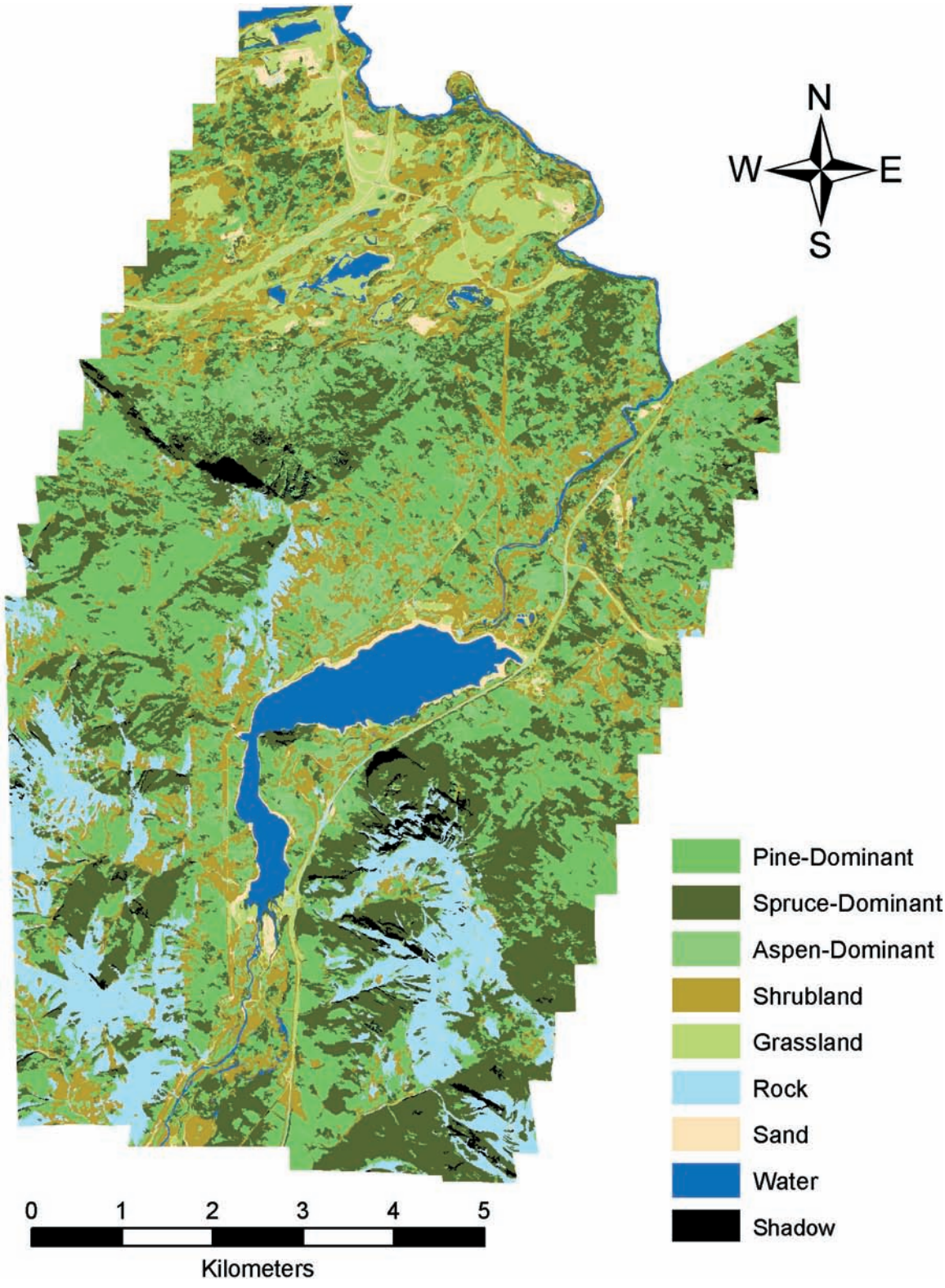

Plate 2. Land-cover classification results.

applied in this study would likely be possible, but would require a more comprehensive training data set encompassing a more complete range of measurements.

Crown Closure

Analysis of percent crown closure for pine-dominated stands resulted in a final regression tree with three terminal nodes

TABLE 3. ACCURACY ASSESSMENT RESULTS: PERCENT PINE COMPOSITION

Class	10 to 29%	30 to 49%	50 to 79%	80 to 100%	Totals
10–29%	93	0	0	0	93
30–49%	7	75	21	0	103
50–79%	0	25	79	5	109
80–100%	0	0	0	95	95
Totals	100	100	100	100	400
Producer's	0.93	0.75	0.79	0.95	
User's	1.00	0.73	0.72	1.00	
KIA per class	0.91	0.66	0.71	0.93	
Overall Accuracy	0.86				
Overall KIA	0.81				

TABLE 4. ACCURACY ASSESSMENT RESULTS: PERCENT CROWN CLOSURE (PINE)

Class	6–47%	48–66%	67–100%	Totals
6–47%	83	0	0	83
48–66%	17	84	10	111
67–100%	0	16	90	106
Totals	100	100	100	300
Producer's	0.83	0.84	0.9	
User's	1.00	0.75	0.85	
KIA per class	0.75	0.75	0.85	
Overall Accuracy	0.85			
Overall KIA	0.78			

(Figure 5). Three percent crown closure classes were defined based on the model output: 6 to 47 percent (the AVI definition of *forest* requires >5 percent crown closure), 48 to 66 percent, and 67 to 100 percent. The 6 to 47 percent class represents open canopy forest and corresponds roughly to AVI crown closure classes "A" (6 to 30 percent) and "B" (31 to 50 percent) combined. The remaining two classes represent closed canopy forest (generally defined as >50 percent crown closure), and correspond roughly to AVI crown closure classes "C" (51 to 70) and "D" (71 to 100). The inability of the model to distinguish more than one open canopy crown closure class was attributed to the lack of data collected for open stands, particularly for stands with <30 percent crown closure. A *mean blue band* threshold provides the initial discrimination between classes, with the 6 to 47 percent class characterized by higher blue band reflectance. The correlation between low percent crown closure values and high blue band reflectance is in agreement with findings from another recent study, which reported a negative correlation between percent crown closure and blue band reflectance due to dry background vegetation with greater reflectivity than the tree crowns (Xu *et al.*, 2003). The two closed-canopy classes were separated on the basis of textural differences using *direction of sub-objects standard deviation*. The selection of a texture-based metric for differentiating between crown closure classes supports the conclusions from other studies which have found image-based texture measures important for forest crown closure mapping from high spatial resolution digital imagery (St-Onge and Cavayas, 1997). Overall accuracy for the crown closure model was 0.85 (Table 4). As was the case with the percent species analysis discussed above, the highest accuracy was achieved for the classes that encompassed the highest proportion of the training data, the 6 to 47 percent and 67 to 100 percent classes in this case.

Stand Height

Analysis of stand height resulted in a final regression tree model with four terminal nodes (Figure 6). From these results, four height classes emerged: <15 m, 15 to 17 m, 18 to 19 m, and >19 m. Partitioning of the training data was based on metrics calculated from the red band (*relative and absolute mean difference to neighbors*) and the near-infrared band (*mean of sub-objects standard deviation*). Identification of metrics derived from the red and near-infrared bands by the regression tree analysis was expected, since red and near-infrared reflectance have been correlated to stand height in the past (Gemmell, 1995; Gerylo *et al.*, 2002). A biophysical explanation for the correlation between red and near-infrared reflectance with stand height is that as trees grow taller, foliage increases; consequently, red reflectance decreases (due to absorption by the foliage) and near-infrared reflectance

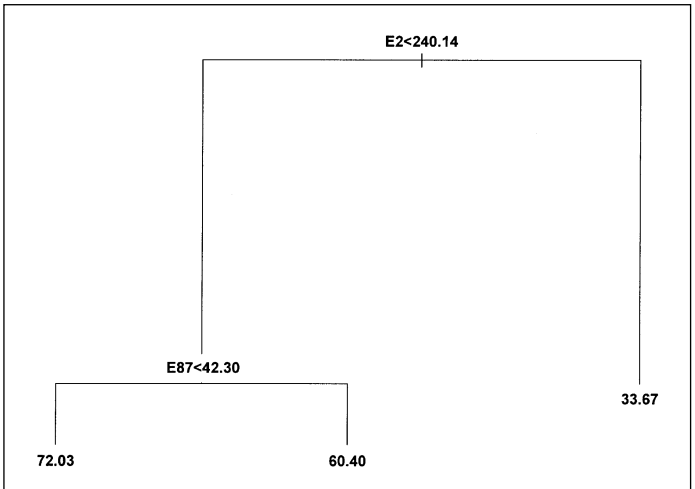

Figure 5. Regression tree: percent crown closure (pine). Metric codes and associated thresholds are shown above each split. Values at terminal nodes represent the mean percent crown closure value for that node.

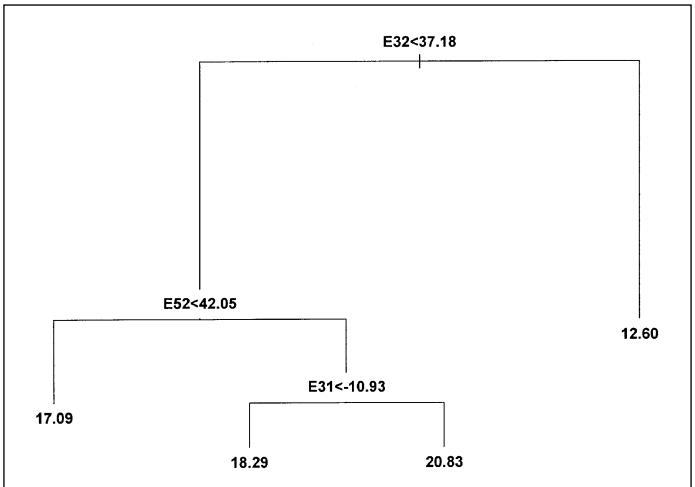

Figure 6. Regression tree: stand height (pine). Metric codes and associated thresholds are shown above each split. Values at terminal nodes represent the mean height value for that node.

TABLE 5. ACCURACY ASSESSMENT RESULTS: STAND HEIGHT (PINE)

Class	<15 m	15 to 17 m	18 to 19 m	>19 m	Totals
<15m	47	11	0	0	58
15–17m	53	43	0	9	105
18–19m	0	46	47	31	124
>19	0	0	53	60	113
Totals	100	100	100	100	400
Producer's	0.47	0.43	0.47	0.60	
User's	0.81	0.41	0.38	0.53	
KIA per class	0.38	0.23	0.23	0.44	
Overall Accuracy	0.49				
Overall KIA	0.32				

TABLE 6. ACCURACY ASSESSMENT RESULTS: STAND AGE, YEARS (PINE)

Class	<90	90 to 100	101 to 110	111 to 120	>120	Totals
<90	67	34	31	0	0	132
90 to 100	33	46	0	0	0	79
101 to 110	0	0	69	28	0	97
111 to 120	0	20	0	35	32	88
>120	0	0	0	37	68	105
Totals	100	100	100	100	100	500
Producer's	0.67	0.46	0.69	0.35	0.68	
User's	0.51	0.58	0.71	0.40	0.65	
KIA per class	0.55	0.35	0.62	0.21	0.60	
Overall Accuracy	0.57					
Overall KIA	0.46					

increases (foliage is highly reflective in the near-infrared). In addition, the amount of shadow cast by a tree will differ with height: taller trees cast longer shadows. Shadow effects are captured particularly well by the red and near-infrared bands (Chen *et al.*, 1999). Despite these relationships, tree height estimates from digital satellite and airborne imagery have been unreliable (Franklin, 2001). Height estimates based on the regression tree model from this study proved to be unreliable also; overall accuracy of model results was assessed at just 0.49 (Table 5). A promising approach for improving remote sensing-based forest height estimates may be to incorporate lidar data into the analysis of segmented image data (Wulder and Seemann, 2003).

Stand Age

Stand age cannot be remotely sensed directly, since age is a descriptor not a physical property. It may be possible, however, to correlate stands of particular ages to unique combinations of measurable properties, such as variables related to structure and composition that vary with maturity (Lefsky and Cohen, 2003). The goal of the regression tree analysis of stand age was to identify such age-correlated associations. The result of the regression tree analysis of stand age was a tree with five terminal nodes, from which five age classes were defined: <90, 90 to 100, 101 to 110, 111 to 120, and >120 (Figure 7). Ten-percentile groupings, such as those derived from the age model output, are typical of standard forest inventory age classes. The metrics used by the regression tree model to partition the stand age training data included derivatives of the DEM (*mean DEM*; *standard deviation*: *DEM*), near-infrared band (*mean NIR*), and red band (*mean difference to brighter neighbors: red*). Once again, correlation of red and near-infrared band derivatives with stand age was expected; red and near-infrared reflectance have been linked to tree height (as discussed above), and tree height has been positively correlated to stand age (Cohen *et al.*, 1995). Accuracy assessment results indicated that the associations between stand age and the model-identified variables were weak, highlighting the difficulties related to identifying relationships between age class through indirect associations with remotely sensed data (Table 6).

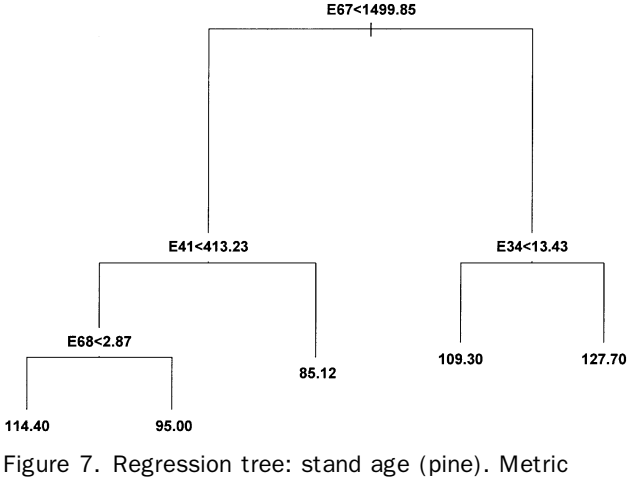

Figure 7. Regression tree: stand age (pine). Metric codes and associated thresholds are shown above each split. Values at terminal nodes represent the mean age value for that node.

Future Work

Decision tree models, like all supervised classification approaches, rely heavily on the quality and representation of the training data (Friedl *et al.*, 1999). It is well known that the accuracy of decision tree models tends to increase with training sample size until a threshold is reached, but this threshold is ambiguous and depends on the application (Pal and Mather, 2003). An association between training data abundance and class accuracy was indeed observed in the present study. For example, in the analyses of species composition and crown closure, the highest accuracies were achieved for the classes that encompassed the majority of the training data (i.e., more field data were collected for certain measurement ranges than for others). Therefore, it is possible that relatively lower accuracy statistics for some classes were due, in part, to under-representation in the training data. An important topic for future research is to determine the thresholds for the decision tree analysis of individual forest inventory parameters in terms of the minimum amount of training data required for optimal model performance.

The classes of individual forest inventory parameters identified through regression tree analyses were based entirely on model output and do not correspond precisely with typical forest inventory classes in all cases. It is likely that a larger training data set, covering a more complete range of measurements, would enable modelling of narrower and more comprehensive classes for individual parameters. Calibration of decision tree models to output more familiar forest inventory class definitions, such as 10-percentile groupings for species composition and crown closure, and 5 to 10 year age classes, would be desirable from a forest management viewpoint. Requirements for large training data sets may be a concern from an operational perspective; however, this limitation could likely be overcome through selective use of photo-interpretation for the purpose of training data collection.

Conclusions

The object-based image analysis approach developed in this study was successful in extracting forest inventory information from Ikonos-2 imagery. Our results suggest that image objects delineated through segmentation are carriers of important forest-related information in the form of image object metrics derived from the inherent spectral and spatial characteristics of forest stand components. We find also that decision trees are effective for identifying relationships between image object metrics and classes of standard forest inventory parameters. The strongest relationships were observed for classes of individual forest species and non-forest land-cover, percent species composition, and percent crown closure, based on quantitative and qualitative assessment of classification results. Comparatively weaker relationships were identified for stand height and age classes. Although many of the relationships between explanatory variables and forest inventory parameters identified by the decision tree models in this study lend themselves to ready biophysical interpretation, these relationships are empirically-derived, and therefore, may not be applicable to other data sets.

We feel that the results obtained in this study, coupled with the relative simplicity of the methods, indicate a growing utility of digital processing of high spatial resolution satellite imagery to meet some operational forest inventory needs. Digital processing approaches that enable a minimization of interpretation time (e.g., through object-based stand boundary delineation) may prove to reduce overall production costs. Additionally, processing of appropriate attributes from satellite image data may also play a role in forest inventory update, by confirming the update results, or through provision of updated species composition as well as disturbance information. The findings of this study indicate the utility of an object-based approach to the processing of high spatial resolution imagery to extract select forest inventory attributes.

Acknowledgments

This research was supported by the Alberta Ingenuity Fund and the Natural Sciences and Engineering Research Council of Canada.

References

Alberta Sustainable Resource Development, 1991. *Alberta Vegetation Inventory Standards Manual*, Version 2.1, Resource Information Division, Resource Information Branch, Edmonton, Alberta.

Baatz, M., and A. Schäpe, 2000. Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation, *Angewandte Geographische Informationsverarbeitung XII* (J. Strobl and T. Blaschke, editors), Wichmann, Heidelberg, pp. 12–23.

Bergen, K., J. Colwell, and F. Sapio, 2000. Remote sensing and forestry: Collaborative implementation for a new century of forest information solutions, *Journal of Forestry*, 98(6):4–9.

Breiman, L., J.H. Friedman, R.A. Olshen, and C.J. Stone, 1984. *Classification and Regression Trees*, Chapman & Hall, CRC Press, New York.

Brown de Colstoun, E.C., M.H. Story, C. Thompson, K. Commisso, T.G. Smith, and J.R. Irons, 2003. National Park vegetation mapping using multitemporal Landsat 7 data and a decision tree classifier, *Remote Sensing of Environment*, 85(3):316–327.

Burns, R.M., and B.H. Honkala, 1990. *Silvics of North America*, Agricultural Handbook 654, U.S. Department of Agriculture, Forest Service, Washington, D.C., 877 p.

Caylor, J., 2000. Aerial photographs in the next decade, *Journal of Forestry*, 98(6):17–19.

Chen, J.M., S.G. LeBlanc, J.R. Miller, J. Freemantle, S.E. Loechel, C.L. Walthal, K.A. Innanen, and H.P. White, 1999. Compact airborne spectrographic imager (CASI) used for mapping biophysical parameters of boreal forests, *Journal of Geophysical Research*, 104(D22):27945–27958.

Cohen, W.B., T.A. Spies, and M. Fiorella, 1995. Estimating the age and structure of forests in a multi-ownership landscape of Western Oregon, U.S.A., *International Journal of Remote Sensing*, 16(4):721–746.

Congalton, R., 1991. A review of assessing the accuracy of classifications of remotely sensed data, *Remote Sensing of Environment*, 37:35–46.

Crawley, M.J., 2002. *Statistical Computing: An Introduction to Data Analysis using S-Plus*, John Wiley & Sons, West Sussex, England, 761 p.

Culvenor, D.S., 2003. Extracting individual tree information: a survey of techniques for high spatial resolution imagery, *Remote Sensing of Forest Environments: Concepts and Case Studies* (M.A. Wulder and S.E. Franklin, editors), Kluwer Academic Publishers, Boston/Dordecht/London, pp. 255–277.

Definiens Imaging, 2002. eCognition, Version 2.1. Definiens Imaging GmbH, München, Germany.

DeFries, R.S., M.C. Hansen, J.R.G. Townshend, and R. Sohlberg, 1998. Global land cover classifications at 8 km spatial resolution: The use of training data derived from Landsat imagery in decision tree classifiers, *International Journal of Remote Sensing*, 19(16):3141–3168.

Flanders, D., M. Hall-Beyer, and J. Pereverzoff, 2003. Preliminary evaluation of eCognition object-based software for cut block delineation and feature extraction, *Canadian Journal of Remote Sensing*, 29(4):441–452.

Franklin, S.E., 1994. Discrimination of subalpine forest species and canopy density using digital CASI, SPOT PLA, and Landsat TM data, *Photogrammetric Engineering & Remote Sensing*, 60(10):1233–1241.

Franklin, S.E., 2001. *Remote Sensing for Sustainable Forest Management*, Lewis Publishers, Boca Raton, Florida, 407 p.

Franklin, S.E., R.J. Hall, L.M. Moskal, A.J. Maudie, and M.B. Lavigne, 2000. Incorporating texture into classification of forest species composition from airborne multispectral images, *International Journal of Remote Sensing*, 21(1):61–79.

Friedl, M.A., C.E. Brodley, and A.H. Strahler, 1999. Maximizing land cover classification accuracies produced by decision trees at continental to global scales, *IEEE Transactions on Geoscience and Remote Sensing*, 37(2):969–977.

Gemmell, F.M., 1995. Effects of forest cover, terrain, and scale on timber estimation with Thematic Mapper data in a Rocky Mountain site, *Remote Sensing of Environment*, 51:291–305.

Gerylo, G.R., R.J. Hall, S.E. Franklin, and L. Smith, 2002. Empirical relations between Landsat TM spectral response and forest stands near Fort Simpson, Northwest Territories, Canada, *Canadian Journal of Remote Sensing*, 28(1):68–79.

Green, K., 2000. Selecting and interpreting high-resolution images, *Journal of Forestry*, 98(6):37–39.

Hall, R.J., 2003. The roles of aerial photographs in forestry remote sensing image analysis, *Remote Sensing of Forest Environments: Concepts and Case Studies* (M.A. Wulder and S.E. Franklin, editors), Kluwer Academic Publishers, Boston/Dordecht/London, pp. 47–75.

Hansen, M.C., R.S. DeFries, J.R.G. Townshend, R. Sohlberg, C. Dimiceli, and M. Carroll, 2002. Towards an operational MODIS continuous field of percent tree cover algorithm: Examples using AVHRR and MODIS data, *Remote Sensing of Environment*, 83:303–319.

Hastie, T., R. Tibshirani, and J. Friedman, 2001. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer-Verlag, New York.

Hay, G.J., T. Blaschke, D.J. Marceau, and A. Bouchard, 2003. A comparison of three image-object methods for the multiscale analysis of landscape structure, *ISPRS Journal of Photogrammetry and Remote Sensing*, 57:327–345.

Hay, G.J., K.O. Niemann, and G.F. McLean, 1996. An object-specific image-texture analysis of H-resolution forest imagery, *Remote Sensing of Environment*, 55:108–122.

Huang, C., L. Yang, B. Wylie, and C. Homer, 2001. A strategy for estimating tree canopy density using Landsat 7 ETM+ and high spatial resolution over large areas, *Proceedings of the Third International Conference on Geospatial Information in Agriculture and Forestry*, Denver, Colorado, 05–07 November 2001.

Insightful Corporation, 2002. *S-PLUS 6.1 for Windows*.

Jensen, J.R., 2000. *Remote Sensing of the Environment: An Earth Resource Perspective*, Prentice-Hall, Upper Saddle River, NJ, 544 p.

Ketting, R.L., and D.A. Landgrebe, 1976. Classification of multispectral image data by extraction and classification of homogeneous objects, *IEEE Transactions on Geoscience Electronics*, GE-14(1): 19–26.

King, D.J., 2000. Airborne remote sensing in forestry: sensors, analysis, and applications, *The Forestry Chronicle*, 76(6): 859–876.

Leckie, D.G., and M.D. Gillis, 1995. Forest inventory in Canada with emphasis on map production, *The Forestry Chronicle*, 71(1): 74–88.

Lefsky, M.A., and W.B. Cohen, 2003. Selection of remotely sensed data, *Remote Sensing of Forest Environments: Concepts and Case Studies* (M.A. Wulder and S.E. Franklin, editors), Kluwer Academic Publishers, Boston/Dordecht/London, pp. 13–46.

Lobo, A., 1997. Image segmentation and discriminant analysis for the identification of land cover units in ecology, *IEEE Transactions on Geoscience and Remote Sensing*, 35(5):1136–1145.

Pal, M., and P.M. Mather, 2003. An assessment of the effectiveness of decision tree methods for land cover classification, *Remote Sensing of Environment*, 86:554–565.

Pekkarinen, A., 2002. Image segment-based spectral features in the estimation of timber volume, *Remote Sensing of Environment*, 82(2–3):349–359.

Pitt, D.G., R.G. Wagner, R.J. Hall, D.J. King, D.G. Leckie, and U. Runesson, 1997. Use of remote sensing for forest vegetation management: A problem analysis, *The Forestry Chronicle*, 73(4):459–477.

Prince, S.D., and M.K. Steininger, 1999. Biophysical stratification of the Amazon basin, *Global Change Biology*, 5(1):1–22.

St-Onge, B.A., and F. Cavayas, 1997. Automated forest structure mapping from high resolution imagery based on directional semivariogram estimates, *Remote Sensing of Environment*, 61(1):85–95.

Sugumaran, R., M.K. Pavuluri, and D. Zerr, 2003. The use of high resolution imagery for identification of urban climax forest species using traditional and rule-based classification approach, *IEEE Transactions on Geoscience and Remote Sensing*, 41(9):1933–1939.

Thomas, N., C. Hendrix, and R.G. Congalton, 2003. A comparison of urban mapping methods using high-resolution digital imagery, *Photogrammetric Engineering & Remote Sensing*, 69(9):963–972.

Venables, W.N., and B.D. Ripley, 2002. *Modern Applied Statistics with S*, Springer-Verlag, New York.

Visual Learning Systems Inc., 2002. *Feature Analyst*, Version 3.1. Visual Learning Systems Inc., Missoula, Montana.

Waring, R.H., and S.W. Running, 1999. Remote sensing requirements to drive ecosystem models at the landscape and regional scale, *Integrating Hydrology, Ecosystem Dynamics, and Biogeochemistry in Complex Landscapes* (J.D. Tenhunen and P. Kabat, editors), John Wiley & Sons, New York, pp. 23–38.

Wulder, M., 1998. Optical remote sensing techniques for the assessment of forest inventory and biophysical parameters, *Progress in Physical Geography*, 22(4):449–476.

Wulder, M.A., R.J. Hall, N.C. Coops, and S.E. Franklin, 2004. High spatial resolution remotely sensed data for ecosystem characterization, *BioScience*, 54(6):511–521.

Wulder, M.A., and D. Seemann, 2003. Forest inventory height update through the integration of lidar data with segmented Landsat imagery, *Canadian Journal of Remote Sensing*, 29(5):536–543.

Xu, B., P. Gong, and R. Pu, 2003. Crown closure estimation of oak savannah in a dry season with Landsat TM imagery: Comparison of various indices through correlation analysis, *International Journal of Remote Sensing*, 24(9):1811–1822.

(Received 10 August 2004; accepted 16 December 2004; revised 08 April 2005)