INT. J. REMOTE SENSING, 2003, VOL. 24, NO. 9, 1835–1852

Taylor & Francis
Taylor & Francis Group

A non-parametric, supervised classification of vegetation types on the Kaibab National Forest using decision trees

S. M. JOY

Rocky Mountain Research Station, 2150 Centre Ave., Bldg. A, Suite 361, Fort Collins, CO 80526-1891, USA; e-mail: sjoy@fs.fed.us

R. M. REICH

Department of Forest Sciences, Colorado State University, Fort Collins, CO 80523, USA; e-mail: robin@cnr.colostate.edu

and R. T. REYNOLDS

Rocky Mountain Research Station, 2150 Centre Ave., Bldg. A, Suite 354, Fort Collins, CO 80526-1891, USA; e-mail: rreynolds@fs.fed.us

(*Received 19 March 2001; in final form 21 January 2002*)

Abstract. Traditional land classification techniques for large areas that use Landsat Thematic Mapper (TM) imagery are typically limited to the fixed spatial resolution of the sensors (30 m). However, the study of some ecological processes requires land cover classifications at finer spatial resolutions. We model forest vegetation types on the Kaibab National Forest (KNF) in northern Arizona to a 10-m spatial resolution with field data, using topographical information and Landsat TM imagery as auxiliary variables. Vegetation types were identified by clustering the field variables total basal area and proportion of basal area by species, and then using a decision tree based on auxiliary variables to predict vegetation types. Vegetation types modelled included pinyon-juniper, ponderosa pine, mixed conifer, spruce- and deciduous-dominated mixes, and openings. To independently assess the accuracy of the final vegetation maps using reference data from different sources, we used a post-stratified, multivariate composite estimator. Overall accuracy was 74.5% (Kappa statistic = 49.9%). Sources of error included differentiating between mixed conifer and spruce-dominated types and between openings in the forest and deciduous-dominated mixes. Overall, our non-parametric classification method successfully identified dominant vegetation types on the study area at a finer spatial resolution than can typically be achieved using traditional classification techniques.

1. Introduction

Habitat changes due to forest management (e.g. tree harvests, fire suppression) are thought to be responsible for declining populations of northern goshawks (*Accipiter gentilis atricapillus*; hereafter referred to as goshawk) in North America (Reynolds 1983, 1989, Speiser and Bosakowski 1984, Crocker-Bedford 1990, Reynolds *et al.* 1992). An accurate description of habitat is, therefore, paramount to understanding the relationship between goshawk demographic performance and its

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online © 2003 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/01431160210154948

changing habitat. Modelling habitat characteristics is often difficult when a study area is large and diverse and complete sampling of environmental variables is unrealistic. Remote sensing technology, however, may be used as an indirect means of obtaining information on the biophysical characteristics of large areas.

Traditional land classification techniques incorporate information derived from remotely sensed data, such as satellites (e.g. Landsat TM sensors), to develop models of land cover and are limited to the fixed spatial resolution of those data (e.g. 30 m). In addition, land cover classifications are frequently driven by (supervised classification) or defined by (unsupervised classification) the interpretation of photography or by incidental knowledge of the study area. The former requires a high level of skill to reduce interpretation error, while the latter may not provide comprehensive coverage of the study area. Ground sampling at select points, assumed to be error free, is often used as auxiliary information to validate the classification, increase the precision of the classification parameters, and/or compute a classification error rate. Disadvantages of traditional classification techniques are the expertise required to carry out and interpret the initial classification and the inability to interpolate the classified data to a scale finer than the spatial resolution of the pixel associated with the imagery.

Here, we describe a method for modelling the composition of goshawk habitat on the Kaibab National Forest (KNF) in northern Arizona to a 10-m spatial resolution using non-traditional classification techniques. Vegetative classes are first derived from a cluster analysis of the field data. Decision trees are then used to model the vegetation throughout the study area using independent variables such as Landsat TM bands, slope, aspect, elevation, and landform. Advantages of using a decision tree over traditional remote sensing classification techniques (Friedl and Brodley 1997, De'Ath and Fabricius 2000) include the non-parametric nature of the test, the test's invariance to transformations of the independent variables, ease of interpretation, and the robustness of results. Unlike many supervised and unsupervised parametric classification algorithms, decision trees can be used to describe non-linear interactions. Individual splits in a decision tree may also be checked for physical justification. In comparative analyses with equivalent linear models (analysis of variance, linear regression) (De'Ath and Fabricius 2000) and maximum likelihood estimators (Hansen *et al.* 1996, Friedl and Brodley 1997), decision trees outperformed both classification techniques. The major disadvantage associated with decision trees, and one that may hinder their widespread application for developing vegetation maps, is the need for large sample sizes (200 or more sample plots; S. M. Joy, personal observation) when dealing with complex data sets (i.e. those with non-linear or high-order interactions). Despite this limitation, decision trees have been used effectively to classify remotely sensed imagery (Michaelson *et al.* 1994, Friedl and Brodley 1997), validate cluster-driven classification algorithms (Iniguez 2000), and explore and interpret ecological data (Baker 1993, De'Ath and Fabricius 2000).

In our study, we used more than one source of data to assess the accuracy of the estimated vegetation map. Field data, considered 'error-free', were used with photo-interpreted data to independently assess model accuracy. However, combining data from several such 'sources' increased the complexity of the accuracy assessment as different sampling protocols were used. Consequently, we used a composite estimator (Green and Strawderman 1990) that combines information from our various sampling sources to estimate accuracy. We discuss the application of our multivariate composite estimator to assess the accuracy of the classification procedure used to identify the major vegetation types on our study area.

2. Study area

The Kaibab Plateau, in northern Arizona, is an oval-shaped (95 km × 55 km), limestone plateau that rises from a shrub-steppe plain at 1750 m above mean sea level (asl) to its highest point at 2800 m. Surface weathering has produced gentle drainages and moderately sloping valleys on the landform. The Plateau is bounded by escarpments of the Grand Canyon of the Colorado River on its south side, by steep slopes on the east, and gentle slopes on the north and west sides that descend to the shrub-steppe plain. The study area, which occupies the northern two-thirds of the Plateau, includes all of the Kaibab National Forest (KNF), above 2182 m asl (figure 1), or about 128 500 ha [estimated from digital elevation models (DEM) using ARC/INFO® (ESRI 1995)]. The KNF is managed by the North Kaibab Ranger District (NKRD; Fredonia, AZ).

Pinyon (*Pinus edulis*)-juniper (*Juniperus* spp.) woodlands occur at elevations below 2075 m asl. Ponderosa pine (*P. ponderosa*) forests occur between 2075 and 2450 m asl, mixed conifer (*P. ponderosa*, *Pseudotsuga mensiesii*, *Abies concolor*, and *Picea pungens*) forests between 2450 and 2650 m asl, and spruce (*Picea engelmannii*)-fir (*A. lasiocarpa*) forests between 2650 and 2800 m asl (Rasmussen 1941, White and Vankat 1993). At transition zones of forest types, associated tree species typically intermix because of differences in slope and aspect. A series of narrow meadows containing grasses and herbaceous vegetation occur on the Plateau.

Each forest type has been altered by some form of management: livestock grazing, fire suppression, thinning, shelterwood, seed-tree and sanitation cuts, and clearcuts. Prior to the introduction of livestock grazing (late-1800s), fire suppression (beginning in the early 1900s), and intensive logging (beginning in the 1980s), many trees were in mature size classes, and occurred singly or in groups (Rasmussen 1941). Forest understories were dominated by grasses (*Poa*, *Sitanion*, and *Muhlengergia* spp.) and typically free from smaller trees (Rasmussen 1941, Merkle 1962). Currently much of the ponderosa pine type understory is dense with pine reproduction and, at higher elevations, with white fir reproduction.

3. Methods

3.1. *Field data*

Field sampling occurred between July and September 1997. A total of 272 field plots (figure 1) obtained from three sources were used to develop a vegetation map of the study area:

3.1.1. *Spectrally-derived plots*

To capture the spectral variability of the study area, we used an unsupervised classification procedure [(CLASSIFICATION; IMAGINE® version 8.3, ERDAS 1997) (ISODATA algorithm, Tou and Gonzalez 1974)] on a 1994 Landsat TM scene of the study area (16 July, centred on Path 37 and Rows 34 and 35) to identify 50 spectral classes. The ISODATA algorithm computes the minimum distance between spectral signatures to identify clusters of pixels with similar spectral characteristics. Use of additional classes might have captured more of the spectral variability in the image; however, we believed that 50 classes sampled a reasonable spectral range and that the remaining plot sources (below) would capture any residual spectral variability. To sample the corresponding vegetative characteristics in the field, we generated 100 random coordinates, two per spectral class (EVALUATE CLASSIFICATION; IMAGINE®, ERDAS 1997). Each coordinate represented the

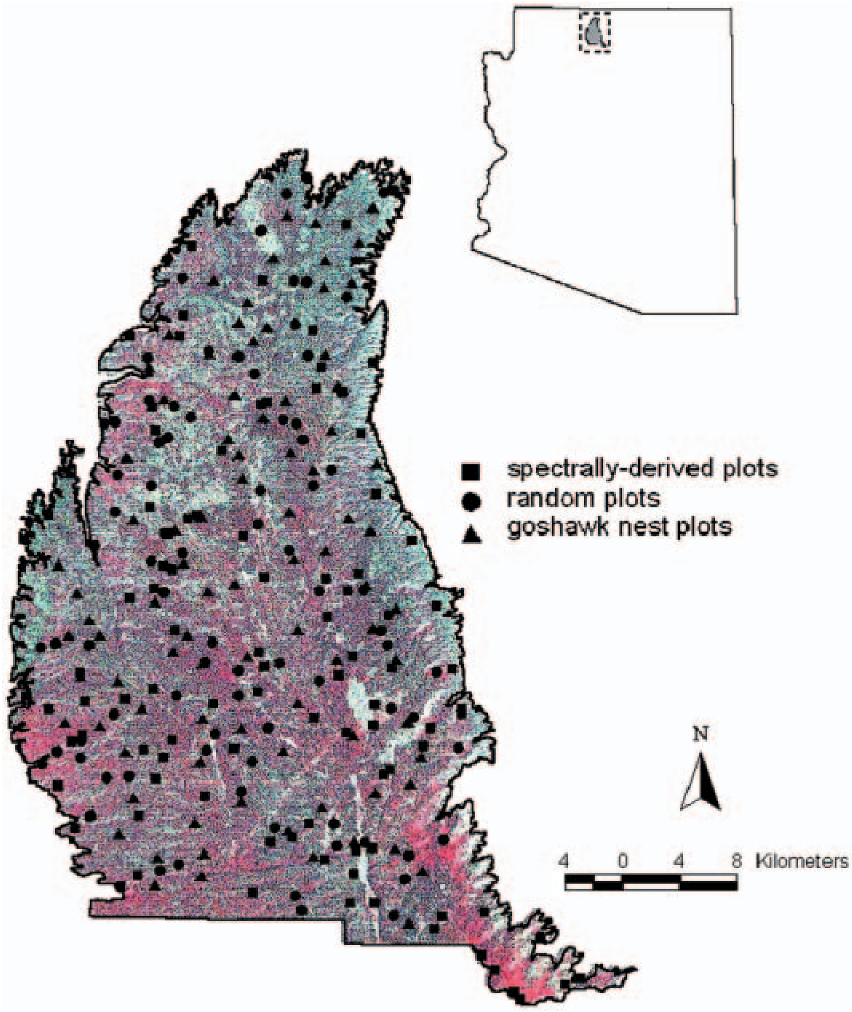

Figure 1. 1997 Landsat TM image (bands 4–red, 3–green, and 2–blue) of the Kaibab National Forest, Arizona, above 2182 m in elevation. Points indicate spectrally-derived (square), random (circle), and northern goshawk nest (triangles) plots.

centre of a 3 pixel × 3 pixel (90-m × 90-m) 'window' of like spectral class. Due to steep terrain, we sampled 98 of the 100 plots.

3.1.2. *Goshawk nest plots*

Goshawks typically lay eggs in two or more 'alternate' nests within their territories among breeding years (Reynolds *et al.* 1994). To estimate the range of vegetative characteristics at goshawk nest areas, a sample plot, centred on the nest tree, was established at one randomly chosen alternate nest in 95 goshawk territories that had been identified on the KNF between 1991 and 1997. Due to data collection errors during field sampling, 92 of the 95 nest tree plots were used in analyses.

3.1.3. *Randomly located plots*

One hundred sample plots were located randomly to capture the remaining vegetative and topographic variability on the study area. The sample plots were established irrespective of territories and nests. Due to steep terrain or the remoteness of plots and time constraints, 85 of the 100 sample plots were measured. The majority (13 of 15) of plots not measured were at the edge of the study area where access was difficult due to steep slopes or cliffs.

Each plot was oriented in a north-south/east-west direction and divided into a cluster of nine 10-m × 10-m subplots corresponding to a 30-m × 30-m pixel on a Landsat TM image. In developing this vegetation map, we used only data from the central 10-m × 10-m subplot on each plot (figure 2). The remaining eight subplots were sampled to model forest structure (canopy closure, total basal area, proportion of pine, spruce-fir, and aspen basal area, maximum height of the understory vegetation, and presence or absence of small trees in the understory) on the study area (Joy and Reich, forthcoming).

Coordinates for the spectrally-derived and nest plots were assigned to the centre of the plot because of the geographical specificity of the unit (Landsat TM pixel or nest tree) on which the plots were based. Coordinates for random plots were systematically assigned to the lower left-hand corner of each plot. A Trimble Navigation Pathfinder® Asset Surveyor Global Positioning System (GPS) with an estimated accuracy of 1–3 m was used to ascertain actual plot locations and account for orienteering errors. Measurements taken on plots included canopy closure [estimated with a concave, spherical densiometer (Lemmon 1956, 1957)], overstory species, total basal area by species [estimated with a 4.6 (metric) factor prism], height of the understory vegetation, % ground cover, and presence of seedling and sapling trees.

Digital maps were generated to assist with the orienteering of sample plots on the ground. Each digital map was composed of a digital elevation model (DEM) (USGS, 1:24 000, 30-m spatial resolution) overlaid with digital line graphs of NKRD roads (R. Crawford, personal communication) and the sample plots.

3.2. *Landsat TM and GIS data*

We obtained a cloud-free, Landsat TM image (22 June 1997; Path 37, Row 35) of the study area that corresponded relatively well with the timing of field data

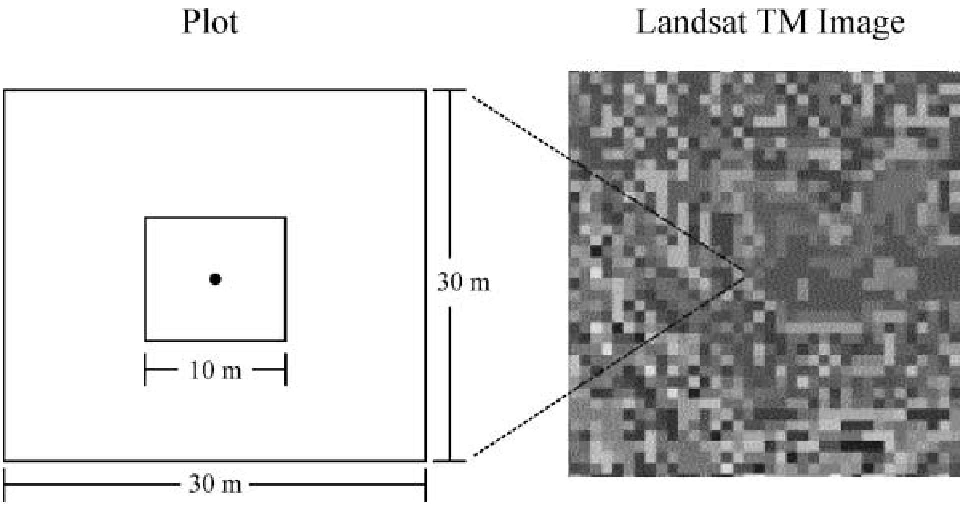

Figure 2. Plot layout used to sample vegetation types on the study area in Arizona to a 10-m spatial resolution and the relationship of the plot to Landsat TM imagery.

collection. Band layers 1–5 and 7 were exported as ARC/INFO® (ESRI 1995) grid coverages and resampled (RESAMPLE, nearest neighbour, GRID Module) to 10 m, corresponding to the spatial resolution of the field data. The value of each pixel was averaged by passing a 3×3 moving window (FOCALMEAN, GRID Module) over the resampled grids. Thus, averaging occurred every 10 m to smooth the spectral transition between features. This resulted in a grid with a piece-wise, continuous surface where every 10-m × 10-m pixel represented the average of the surrounding 30-m × 30-m pixels, including the central 10-m pixel of each original Landsat TM pixel whose value did not change. This was done to reduce potential registration errors and to reflect changes in forest structure and vegetative classes among the nine 10-m × 10-m subplots measured on the ground. Resampling was also important because not all of the forested plots fell within spectrally distinct areas; some plots were in transition zones between spectral classes. Average DN values associated with the 272 sample plots were then extracted from each of the six Landsat TM grids using a customized ArcView® (ESRI 1998) application.

Elevation, slope, aspect, and landform (McNab 1989) were also determined for each sample plot from DEMs. Landform (McNab 1989) is an index that expresses surface shape as a measure of surface concavity or convexity (computed as the mean slope gradient from the original cell to adjacent cells in four directions) creating a continuous variable. Prior to extracting the cell values, each grid was resampled to 10 m as described above.

3.3. *Photographic data*

True colour aerial photography (1:12 000, 1991, NKRD), infrared National High Altitude Photography (NHAP) (1:58 000, 1980, USGS), Digital Orthophoto Quadrangles (DOQs) (1:24 000, 1992, USDA Forest Service Geometric Service Center, USGS), and photographs taken during field sampling were used in the assessment of classification accuracy. Information on forest management activities [Resource Information System (RIS) data] aided in identifying management treatments that occurred subsequent to the acquisition of the photography and was provided by the NKRD (K. Fuelling, D. Steffensen, personal communication).

3.4. *Image classification*

A hierarchical clustering algorithm (HCLUST; S-PLUS©, Statistical Sciences 1995), based on group averaging, was used to group the field data into clusters with similar vegetative characteristics (i.e. species composition, basal area, canopy closure, understory vegetation, etc.). The algorithm produces a two-dimensional dendrogram representing the fusion of clusters at each successive stage of the analysis. At each stage, the two 'nearest' clusters are combined to form one bigger cluster (initially each cluster contains a single sample plot). This process continues until only one group containing all of the observations remains. Empirical studies have shown that when there are an unequal number of observations per cluster, group averaging is more successful at identifying clusters than other hierarchical clustering methods (Everitt 1974). Group averaging also has superior performance when the data contain evidence of a distinct cluster structure (Everitt 1974).

The resulting field data clusters were then assigned to modified Anderson Level II ('opening' class only) or III (all other forest types) (Anderson *et al.* 1976) land cover classes (table 1). Anderson Levels II and III were assumed to be sufficient for modelling dominant goshawk habitat. S-PLUS© (TREE; Statistical Sciences 1995)

Table 1. Modified land cover classification system (Anderson *et al.* 1976) for vegetation types on the Kaibab National Forest, Arizona. Characteristics at higher levels are nested within the lower level.

Land cover		
Level I	Level II	Level III
Forest	Coniferous	Ponderosa pine
		Mixed Conifer
↓	↓	↓
	Deciduous	Deciduous-dominated mix
	↓	↓
	Mixed	Pinyon-juniper
	↓	Spruce-dominated mix
Non-forested	Opening	

was used to generate a stepwise decision tree (Breiman *et al.* 1984, Friedl and Brodley 1997, De'Ath and Fabricius 2000) that identified independent variables (Landsat TM bands, elevation, slope, aspect, or landform) that were important in discriminating among vegetation types. The decision tree uses a binary partitioning algorithm that maximizes the dissimilarities among groups to compare all possible splits among the independent variables and splits within each independent variable to partition the data into new subsets. Once the algorithm partitions the data into new subsets, new relationships are developed to split the new subsets. The algorithm recursively splits the data in each subset until either the subset is homogeneous or the subset contains too few observations (<5) to be split further. To prevent over fitting the data, a pruning algorithm (Friedl and Brodley 1997) was used to eliminate subsets that were fit to noise in the data. Decision tree criteria were then used as 'training' statistics to classify the 1997 Landsat image. The classification was achieved using Arc Macro Language [AML (CON; ARC/INFO®, ESRI 1995)] to produce the final vegetation map.

3.5. *Accuracy assessment*

A sample-based assessment of accuracy (i.e. that based on the same data used to generate the classification) for the decision trees was calculated by weighting the classification error associated with a given vegetation type proportional to its area, a process referred to as post-stratification (Cochran 1977:134–135). To assess a model's classification accuracy independently, field data alone may be used if the data were not involved in the classification procedure or came from a sufficiently large sample. However, if the sample of field data is small or the data were used in developing the classification procedure (such as here), independent auxiliary information is needed in conjunction with the field data to estimate accuracy. When data from different sources are combined in this manner, the complexity of the accuracy assessment increases due to different sampling protocols. A composite estimator (Green and Stawderman 1990) may be used to combine the data from the different sources and improve estimates of model accuracy. Czaplewski (2000) provides detailed discussion on the use of auxiliary information to assess the accuracy of vegetation maps derived from remotely sensed data (see also Kalkhan *et al.* 1996, 1998).

To assess independently the accuracy of our classification using the multivariate composite estimator, we generated 498 random points as auxiliary data using simple random sampling. The location of the 498 random points and 269 of the 272 field plots (photographic data were unavailable for three of the field plots) were then transferred to aerial photographs, NHAP, and/or DOQs. An independent photo interpreter (i.e. one with no knowledge of the classification) identified the dominant vegetation type corresponding to each $10\,\mathrm{m} \times 10\,\mathrm{m}$ sample location (random photo points, field plots). Grid values were also then extracted from the vegetation map for each location using a customized ArcView® (ESRI 1998) application.

In this study, we used three selection criteria (spectrally-derived, nest-based, and random plots) to sample the field data, increasing the complexity of our analysis as formulas used in summarizing these data depended on sampling design (Stehman 1999). When vegetation types associated with the sample data are proportional to the true values, the sample data should provide good estimates of the overall accuracy, as a simple random sample distributes itself approximately proportionally among strata (Cochran 1977: 134–135). However, because of the nesting requirements of the goshawk, the vegetation types associated with nest-tree plots are not likely to be in proportion to those observed on the study area outside of nest areas. Also, because spectrally-derived plots were designated based on reflectance values, they may not be in proportion to the vegetation types observed on the study area. As a result, the combination of these data may lead to poor estimates of the overall model accuracy. Stehman (1996) points out, however, that if simple random sampling is combined with a post-stratified estimator, the estimator is nearly as efficient as a proportionally-allocated stratified design. A post-stratified estimator combined with simple random sampling should also perform nearly as efficiently as an optimally-allocated stratified design for estimating overall accuracy (Stehman 1999), assuming a large enough sample, because an optimally-allocated stratified design is usually only slightly more precise than a proportionally-allocated stratified design (Cochran 1977: 327–335).

To compensate for differences in the selection probabilities associated with our three sets of sample data, we adjusted the joint probability matrix associated with the three sampling procedures using post-stratification. This adjusted error matrix was used to calculate a Kappa statistic to estimate the difference between the classified and ground-verified themes, and the agreement contributed by chance. Because of the complexity of our sampling design, the usual formula for calculating the variance of the Kappa is not appropriate, as the data do not follow a multinomial distribution. Under these conditions, bootstrapping techniques can be used to construct confidence bands around the estimates of Kappa (Kalkhan *et al.* 1997). As we were not testing hypotheses or making comparisons, this was not done.

4. Results

4.1. *Image classification*

Preliminary analysis indicated that the accuracy of the classification could be increased by modelling pure ponderosa pine forests separately from the other vegetation types. Sample plots with 100% of basal area in ponderosa pine were assigned a value of 1 and all other plots were assigned a value of 0 when fitting the decision tree (figure 3). The final decision tree had 25 terminal nodes. Discriminating variables included elevation, slope, aspect, landform, and Landsat TM bands 1–5 and 7. The initial split, which maximized the distance between the response variables

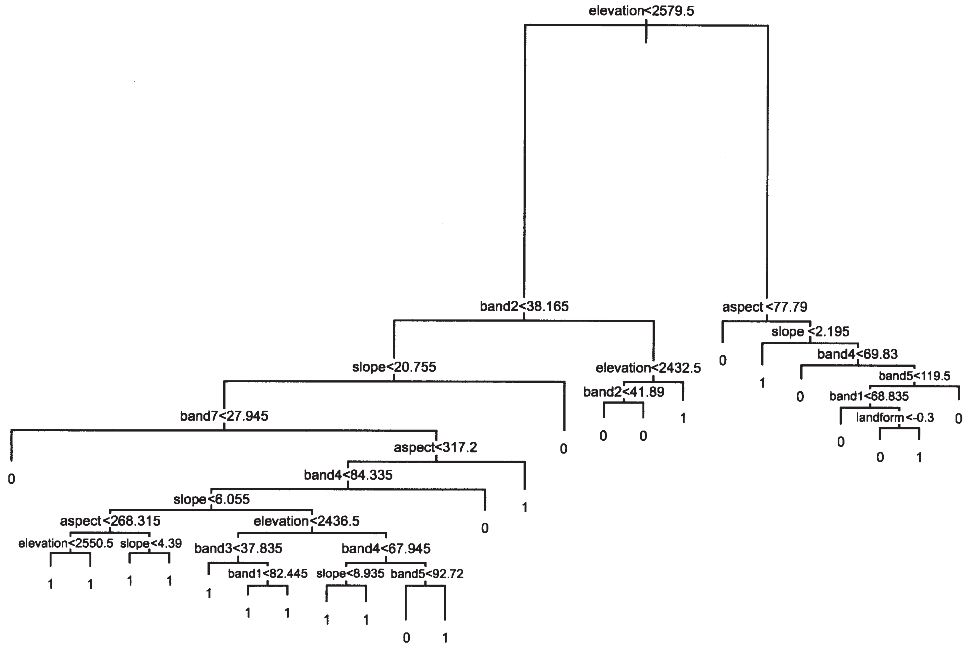

Figure 3. Decision tree classifier for ponderosa pine (1) and other dominant vegetation types (0) on the study area in Arizona. Independent variables (and units) used in the classification procedure include elevation (m), slope (%), aspect (degrees), landform (index), and Landsat TM bands 1–7 (digital number).

(pine/other), was based on elevation (2579 m) and separated the higher elevation spruce-fir forests from the pine and pinyon-juniper forests that occurred at lower elevations. Ninety-eight per cent (148/151) of the pure pine plots and 86% (104/121) of 'other' forest type plots (i.e. ponderosa pine mixed with other species, spruce, fir and deciduous vegetation, openings) were correctly classified for an overall resubstitution model accuracy of 92%.

After we developed the decision tree for pure ponderosa pine, the pure pine sample plots were removed from the data set. We then identified six additional vegetation classes from the remaining data by clustering the proportion of basal area associated with the remaining species. Vegetation classes included pinyon-juniper, mixed conifer, fir-dominated mix, spruce-dominated mix, deciduous-dominated mix, and openings. We later consolidated mixed conifer and fir-dominated mix into the mixed conifer class to improve the model. The final decision tree (figure 4) for the 'other' (i.e. non-pure pine) vegetation classes had 20 terminal nodes. Important variables in the partitioning criteria included elevation, slope, aspect, and Landsat TM bands 1, 2, 4, 5, and 7. The initial split was again based on elevation (2418 m) and separated pinyon-juniper and some deciduous-dominated classes and openings from higher elevational forest mixes. Landsat TM band 7, representing the mid-infrared (2.08–2.35M) spectral range, further distinguished openings from the majority of higher elevational forest mixes. Eighty-seven per cent (14/16) of the pinyon-juniper sites were correctly classified (table 2). Similarly, 88% (29/33) of the mixed conifer sites, 87% (27/31) of the spruce-dominated sites, 90% (18/20) of the deciduous-dominated sites, and 76% of the openings (16/21) were correctly

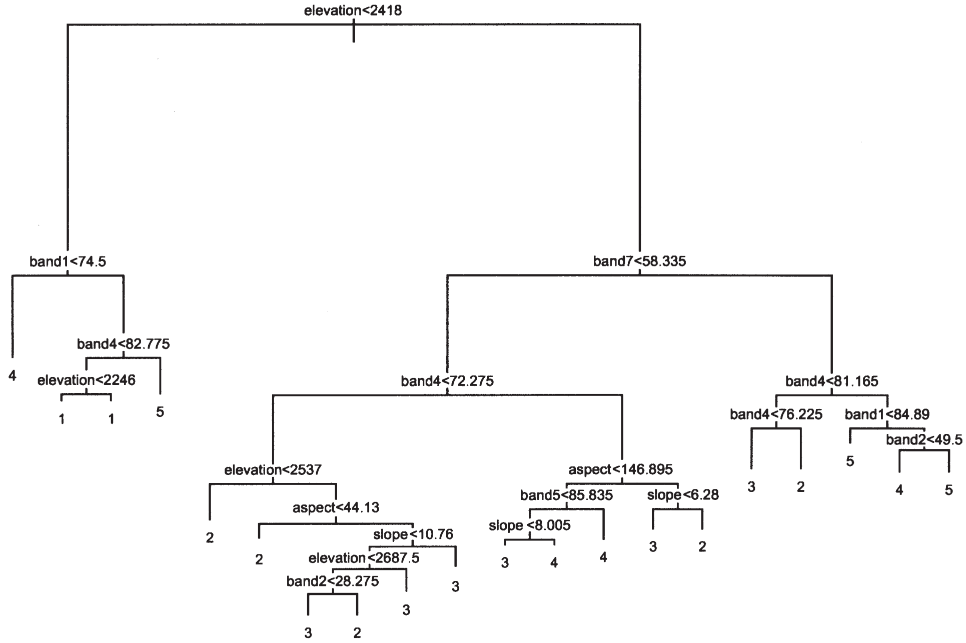

Figure 4. Decision tree classifier for dominant vegetation types other than pure ponderosa pine on the study area in Arizona. Vegetation types include pinyon-juniper (1), mixed conifer (2), spruce-dominated mixes (3), deciduous-dominated mixes (4), and openings (5). Independent variables (and units) used in the classification procedure include elevation (m), slope (%), aspect (degrees), and Landsat TM bands 1, 2, 4, 5 and 7 (digital number).

Table 2. Resubstitution classification error rates of the decision tree for the vegetation types on the Kaibab National Forest, Arizona. Error rates were adjusted (post-stratified) for the proportions of the vegetation classes on the study area.

Vegetation type	Misclassified plots	Total plots	Classification error rate	Post-stratified classification error rate
Pinyon-juniper	2	16	0.125	0.010
Ponderosa pine	3	151	0.020	0.011
Mixed conifer	4	33	0.121	0.014
Spruce-dominated mix	4	31	0.129	0.013
Deciduous-dominated mix	2	20	0.100	0.009
Openings	5	21	0.238	0.014
Column totals	20	272		
Overall error			0.125	0.071

classified. The overall resubstitution accuracy of this decision tree was 86%. The combined model (pure ponderosa pine and other vegetation classes) correctly classified 93% (20/272) of the sample plots. Overall post-stratified resubstitution accuracy for the model was also 93% (table 2). In both models (pure ponderosa pine and all other vegetation types), the use of additional variables (e.g. canopy closure, height of the understory vegetation, presence of seedling and sapling, and proportion of ground covered) tended to obscure the cluster structure and did not improve our ability to identify major vegetation types.

On the final vegetation map (figure 5), ponderosa pine occupied the majority (55.5%) of the study area and occurred at lower elevations, with mixed conifer (11.3%) and spruce-dominated mixes (10.1%) occurring at higher elevations (table 3). Pinyon-juniper (8.2%) was found predominantly along lower elevational edges of the study area and where crown-destroying fire and intensive management (e.g. shelterwood and seed-tree cuts) had occurred at lower elevations. Deciduous-dominated mixes (8.8%) and openings (6.1%) occurred throughout the study area. All field plots, except goshawk nests, were sampled in relatively close proportion to the occurrence of vegetation type in the model (table 3).

4.2. *Accuracy assessment*

The overall accuracy of the classification model when compared to the photo interpreted data was 62.8%, with a Kappa of 32.9% (table 4). The low Kappa statistic suggests a poor relationship between the classification derived from the decision tree and the photo interpreted classification. Low accuracy estimates associated with mixed conifer, spruce-dominated mixes, and openings contributed largely to the low overall accuracy (table 4). Errors associated with coniferous species were attributed to difficulties in distinguishing among mixed stands of pine, fir, and spruce on the photographs, especially in small ($<100\,m^2$) patches of individual species.

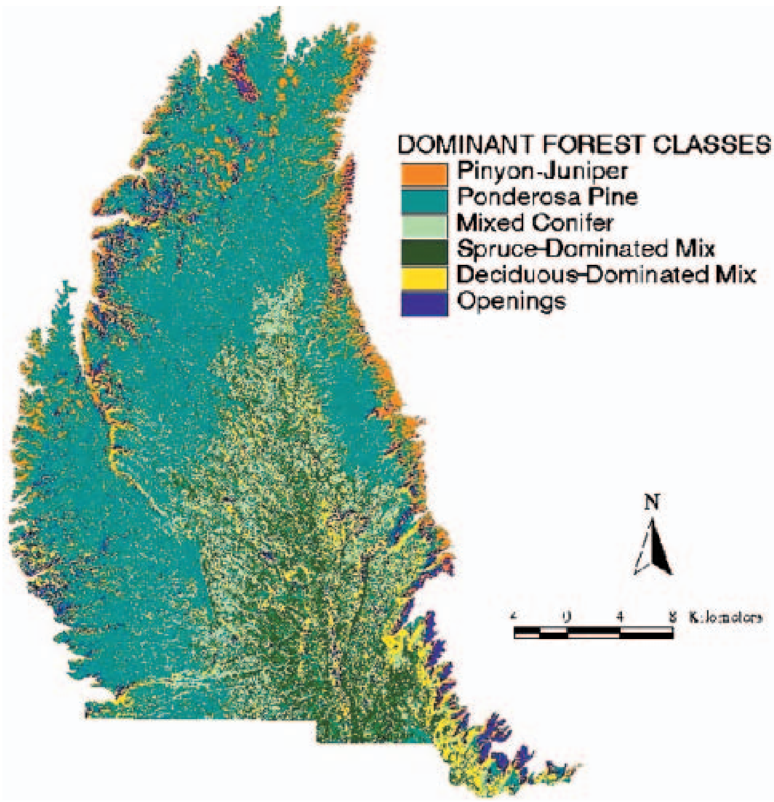

Figure 5. Dominant vegetation types on the study area in Arizona.

Table 3. Distribution of sample plots by vegetation type on the Kaibab National Forest, Arizona.

Vegetation type	Number of pixels (10 m × 10 m)	Proportion of total area	Spectrally-derived plots	Goshawk nest plots	Randomly located plots	All field plots	Photo-interpretation plots
Ponderosa pine	7 137 947	0.56	0.45	0.70	0.59	0.58	0.58
Mixed conifer	1 454 207	0.11	0.14	0.12	0.10	0.12	0.11
Spruce-dominated mix	1 300 085	0.10	0.11	0.12	0.09	0.11	0.11
Deciduous-dominated mix	1 124 891	0.09	0.09	0.06	0.07	0.07	0.10
Pinyon-juniper	1 055 726	0.08	0.10	0.00	0.06	0.05	0.07
Openings	778 865	0.06	0.11	0.00	0.09	0.07	0.03
Column totals	112 851 721	1.00	1.00	1.00	1.00	1.00	1.00

Table 4. Joint probability error matrix for the vegetation types and photo-interpretation of 767 sample plots.

Vegetation type (Photographs)	PJ	PP	MC	SD	D	OP	Row totals	Accuracy
Pinyon-juniper (PJ)	0.0313	0.0222	0.0000	0.0000	0.0000	0.0078	0.0613	0.5106
Ponderosa pine (PP)	0.0222	0.4003	0.0117	0.0104	0.0261	0.0078	0.4785	0.8365
Mixed conifer (MC)	0.0026	0.0561	0.0469	0.0130	0.0143	0.0039	0.1368	0.3429
Spruce-dominated mix (SD)	0.0000	0.0209	0.0326	0.0717	0.0169	0.0026	0.1447	0.2687
Deciduous-dominated mix (D)	0.0013	0.0378	0.0104	0.0052	0.0235	0.0091	0.0873	0.4955
Openings (OP)	0.0052	0.0430	0.0091	0.0078	0.0117	0.0143	0.0911	0.1571
Column totals	0.0626	0.5803	0.1107	0.1081	0.0925	0.0455	0.9997	
Overall accuracy = 62.82%	Standard error = 0.99%							
Kappa = 32.90%								

Photo interpretation of these forest types using mid-summer, true colour aerial photography and black and white DOQs was difficult. The large error associated with openings was a result of forest regeneration following tree harvests or fire subsequent to the acquisition of photography and to the misclassification of openings containing a small arboreal component. Spectral similarities between deciduous and open forest vegetation also contributed to the error estimate for openings. Narrow (<800 m) meadows comprised of grasses and forbs were often classified as deciduous-dominated vegetation, rather than as openings.

After using the multivariate composite estimator (Appendix) to adjust for differences in the photo-interpretation of random points and field plots, and adjusting for the proportion of vegetation types (i.e. post-stratification) on the study area, the performance of the classification procedure improved to 74.5% (table 5). All vegetation types except openings had an estimated accuracy greater than 50%. This level of accuracy is comparable to results achieved with traditional classification techniques that use a larger (typically 900 m^2) per-pixel area. Our Kappa statistic for the classification model was approximately 50%, which indicated good agreement between the classification model and observations on the ground.

The upward adjustment in the overall accuracy and Kappa statistic after correcting for classification errors in photographic interpretation indicated a larger classification error associated with the independent photo-plots compared to the error of classifying field plots. There was no indication of any subjective bias associated with the photo-interpretation of the sample plots to suggest a loss of efficiency in estimating the weights of the composite estimator. If a bias were to exist, an alternative composite estimator proposed by Green and Strawderman (1990) could be employed. This estimator behaves as well as the usual precision-weighted multivariate composite estimator when auxiliary data are unbiased, yet is superior when the auxiliary information is severely biased.

5. Discussion and conclusions

Goshawks occur in a wide range of forest types and structures throughout their geographical range in North America. Qualitative differences among the habitats occupied may be measured in terms of the fitness of individuals who occupy the range of habitat conditions (Fretwell and Lucas 1970, Van Horne 1983). To understand how goshawks are influenced by their environment, the relationships between their demographic performance on territories and the composition and spatial arrangement of vegetation within their territories need to be determined. That is, we need to accurately model the vegetation to adequately characterize goshawk habitat use on the study area.

Dominant vegetation types on the study area included pinyon-juniper, ponderosa pine, mixed conifer, spruce-dominated mixes, deciduous-dominated mixes, and openings. Stands of pure ponderosa pine were identified with the highest accuracy. Differentiating between mixed conifer and spruce-dominated forest types and between deciduous-dominated mixes and openings was difficult due to their spectral and/or physical similarities. Auxiliary variables (canopy closure, understory vegetative species and height, proportion of ground covered, and presence of seedling/sapling) did not improve the accuracy of the model because these variables tended to reflect differences in stocking levels, and not differences in species composition. Modifying our sampling to include additional field plots in these vegetation classes might have mitigated these classification dilemmas. Our efforts to model vegetation type,

Table 5. Joint probability error matrix for the decision tree for vegetation types and the field data. Probabilities were corrected for differences in the photo-interpretation of random points and field plots using the multivariate composite estimator and adjusted to the proportions of the vegetation classes on the Kaibab National Forest, Arizona.

	Vegetation types (classified map)							
Vegetation type (field data)	PJ	PP	MC	SD	D	OP	Row totals	Accuracy
Pinyon-juniper (PJ)	0.0374	0.0133	0.0000	0.0000	0.0000	0.0115	0.0622	0.6014
Ponderosa pine (PP)	0.0149	0.4956	0.0100	0.0035	0.0118	0.0102	0.5460	0.9076
Mixed conifer (MC)	0.0000	0.0308	0.0667	0.0192	0.0046	0.0000	0.1213	0.5499
Spruce-dominated mix (SD)	0.0000	0.0179	0.0197	0.0736	0.0139	0.0010	0.1261	0.5017
Deciduous-dominated mix (D)	0.0000	0.0261	0.0100	0.0038	0.0402	0.0000	0.0801	0.5838
Opening (OP)	0.0000	0.0163	0.0055	0.0089	0.0057	0.0279	0.0643	0.4333
Column totals	0.0523	0.6000	0.1119	0.1090	0.0762	0.0506	1.0000	

Overall accuracy = 74.47% (Overall standard error = 1.56%)
Kappa = 49.87%

however, were incidental to the primary sampling objective, which was to model forest structure on the study area to a 10-m spatial resolution (Joy and Reich, forthcoming) and so we were limited by the protocol thereby established.

Use of independent data (i.e. those not used in the classification procedure) to assess the accuracy of a classification of remotely sensed data is now common practice. The assessment procedure should take into consideration the sampling design, sample size, and the classification scheme (Congalton 1991). Our method of accuracy assessment uses photo-interpretation as a relatively inexpensive source of independent reference data to assess the accuracy of our vegetation map. Use of the multivariate composite estimator also takes into account our multiple sampling schemes, by combining the independent set of photo-interpreted sample data with our field data to obtain a more efficient estimate of the overall classification accuracy (74.5%). This method corrected for photo-interpretation errors such as misclassification and incorrect plot locations, which can confound the assessment (Czaplewski 2000). However, the overall accuracy of our map may be slightly higher or lower than this because of registration errors between the Landsat TM image and the field data (Reimann *et al.* 2000). We feel that registration errors were negligible in our case.

Overall, the non-parametric classification method presented here successfully identified dominant vegetation types at a finer spatial resolution than is typically achieved using traditional classification techniques. We used this model to aid in the description of forest structure on the study area (Joy and Reich, forthcoming). We are also currently correlating our model of vegetative composition with a ranking of goshawk territories based on the hawk's long-term demographic performance to identify forest characteristics of goshawk territory quality (Joy 2002).

Acknowledgments

We gratefully acknowledge the assistance of Carrie M. Erickson, Matthew A. Gavin, Luke J. H. Hunt, Amy M. Iñiguez, Jose M. Iñiguez, Mohammed A. Kalkhan, Donna C. Laing, Jeffrey S. Lambert, Joanna L. Nelson, Susan S. Salafsky, John C. Seyfried, Geroge S. Stamatellos, Rebecca A. Steffensen, Vernon L. Thomas, J. David Wiens, and Laura E. Williams for collecting field data. Vernon L. Thomas also provided customized ArcView® code and guidance on GIS manipulations. We thank Donna C. Laing for the interpretation of photographic images. Rudy M. King, William J. Krausmann, and Michael S. Williams provided helpful reviews of the manuscript. The North Kaibab Ranger District (NKRD, Fredonia, AZ) kindly provided housing and logistical support. We also thank Rick Crawford, Karl Fuelling, and Dave Steffensen of the NKRD for support in acquiring RIS data, DOQs, and road coverages of the KNF. This study was funded by Region 3 (Albuquerque, NM) and the Rocky Mountain Research Station (Fort Collins, CO), USDA Forest Service.

Appendix

Estimators that combine information from various sources are referred to as composite estimators (Green and Strawderman 1990). To apply the multivariate composite estimator to assess the accuracy of our classification procedure, we collected the data in two phases. In phase one, n' (296 field and 498 random) plots were located on aerial photographs. Each of the n' photo plots were classified by photo interpretation into one of k classes. This information was used to construct a $k \times k$ matrix of joint probabilities of all possible classification outcomes associated with the interpretation of plots on the photographs and the classification procedure under evaluation. To implement the composite estimator, we stacked the columns of the $k \times k$ error matrix of joint probabilities on top of each other to create a $k^2 \times 1$ vector (x'), along with a $k^2 \times k^2$ variance-covariance matrix of $V(x')$.

In phase two, the field plots (n, a sub-sample of n') were classified into one of k classes associated with the classification outcome. This information was combined with the photo interpretation and classification outcome on our map to construct a $k \times k \times k$ matrix of joint probabilities associated with the three classification procedures. This matrix of joint probabilities was re-arranged to create the $k^3 \times 1$ vector, y_1, and $k^3 \times k^3$ variance-covariance matrix $V(y_1)$. Given an appropriately structured $k^2 \times k^3$ matrix of zeros and ones (H_x), the vector y_1 can be collapsed to provide the $k^2 \times 1$ vector, $x = H_x y_1$, of joint probabilities associated with the photo interpretation of field plots and the classification procedure being evaluated. Next, the composite estimator was used to combine these two sources of information to obtain a more efficient $k^3 \times 1$ vector:

$$y_c = y_1 + K(x' - x) \tag{A1}$$

of joint probabilities associated with the three classification procedure, where

$$K = V(y_1)H_y'[H_y V(y_1)H_y' + V(x')]^{-1} \tag{A2}$$

is the $k^3 \times k^2$ weighting factor, H_y is a $k^2 \times k^3$ matrix of zeros and ones used to sum the joint probabilities within the appropriate groups to form the desired $k^2 \times k^1$ vector $y = H_y y_c$ of joint probabilities associated with the classification of the field plots and the classification procedure, and H_y' is the transpose of H_y. In equation (A1), $(x' - x)$ is an estimate of the difference in the classification errors association with the photo interpretation in the first and second phases of sampling. If there are no differences (i.e. $x' = x$), the most efficient estimation of the classification error is based on the data collected in the second phase of sampling. If there are differences, the classification errors are adjusted accordingly (equation (A1) to take these differences into consideration.

References

Anderson, J. R., Hardy, E., Roach, J., and Witmer, R., 1976, A land use and land cover classification system for use with remote sensor data. Geological Survey Professional Paper 964 (Washington, DC: US Government Printing Office), pp. 1–28.

Baker, F. A., 1993, Classification and regression tree analysis for assessing hazard of pine mortality caused by *Heterobasidion annosum*. *Plant Disease*, **77**, 136–139.

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J., 1984, *Classification and Regression Trees* (Belmont, California: Wadsworth Ind. Group).

Cochran, W. G., 1977, *Sampling Techniques* (New York: John Wiley & Sons).

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

Crocker-Bedford, C., 1990, Goshawk reproduction and forest management. *Wildlife Society Bulletin*, **18**, 262–269.

Czaplewski, R. L., 2000, Accuracy assessment and areal estimates using two-phase stratified random sampling, cluster plots, and the multivariate composite estimator. In *Quantifying Spatial Uncertainty in Natural Resources: Theory and Applications for GIS and Remote Sensing*, edited by H. T. Mowrer and R. G. Congalton (Chelsea, Michigan: Ann Arbor Press), pp. 79–100.

De'ath, G., and Fabricius, K. E., 2000, Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology*, **81**, 3178–3192.

ERDAS, 1997, *IMAGINE® 8.3 Software and On-line Help Manual* (Atlanta, Georgia: ERDAS, Inc.).

ESRI, 1995, *ARC/INFO® Software and On-line Help Manual* (Redlands, California: Environmental Research Institute, Inc.).

ESRI, 1998, *ArcView® 3.1*. Environmental Research Institute, Inc., Redlands, CA.

Everitt, B. S., 1974, *Cluster Analysis*, 3rd edn (London: Edward Arnold).

Fretwell, S. D., and Lucas, H. L., 1970, On territorial behavior and other factors influencing habitat distribution in birds. I. Theoretical development. *Acta Biotheoretica*, **19**, 16–36.

Friedl, M. A., and Brodley, C. E., 1997, Decision tree classification of land cover from remotely sensed data. *Remote Sensing of Environment*, **61**, 399–409.

Green, E. J., and Strawderman, W. E., 1990, A multivariate composite estimator. *State-of-the-Art Methodology of Forest Inventory: a Symposium Proceeding; 1989 30 July–5 August, Syracuse,* NY. General Technical Report PNW-GTR-263 (Pacific Northwest Research Station, Portland, Oregon: US Department of Agriculture, Forest Service), pp. 181–188.

Hansen, M., Dubayah, R., and DeFries, R., 1996, Classification trees: an alternative to traditional land cover classifiers. *International Journal of Remote Sensing*, **17**, 1075–1081.

Iniguez, J. M., 2000, *Cover Type and Structural Classification in the Sky Islands of Southeastern Arizona*. (Flagstaff, Arizona: Northern Arizona University).

Joy, S. M., 2002, Northern goshawk habitat on the Kaibab National Forest in Arizona: Factors affecting nest locations and territory quality (Fort Collins: Colorado State University) Ph.D. Dissertation.

Kalkhan, M. A., Reich, R. M., and Czaplewski, R. L., 1996, Statistical properties of measures of association and the Kappa statistic for assessing the accuracy of remotely sensed data using double sampling. General Technical Report RM-GTR-277 (Rocky Mountain Forest and Range Experiment Station, Fort Collins, Colorado: US Department of Agriculture, Forest Service), pp. 467–476.

Kalkhan, M. A., Reich, R. M., and Czaplewski, R. L., 1997, Variance estimates and confidence intervals for the Kappa measure of classification. *Canadian Journal of Remote Sensing*, **23**, 246–252.

Kalkhan, M. A., Reich, R. M., and Stohlgren, T. J., 1998, Assessing the accuracy of Landsat Thematic Mapper classification using double sampling. *International Journal of Remote Sensing*, **19**, 2049–2060.

Lemmon, P. E., 1956, A spherical densiometer for estimating forest overstory density. *Forest Science*, **2**, 314–320.

Lemmon, P. E., 1957, A new instrument for measuring forest overstory density. *Journal of Forestry*, **55**, 667–668.

McNab, W. H., 1989, Terrain shape index: quantifying effect of minor landforms on tree height. *Forest Science*, **35**, 91–104.

Merkle, J., 1962, Plant communities of the Grand Canyon area, Arizona. *Ecology*, **43**, 698–711.

Michaelson, J., Schmiel, D. S., Friedl, M. A., Davis, F. W., and Dubayah, R. O., 1994, Regression tree analysis of satellite and terrain data to guide vegetation sampling and surveys. *Journal of Vegetation Science*, **5**, 673–686.

Rasmussen, D. I., 1941, Biotic communities of the Kaibab Plateau, Arizona. *Ecological Monographs*, **11**, 230–274.

Reimann, R., Hoppus, M., and Lister, A., 2000, Using arrays of small ground sample plots to assess the accuracy of Landsat TM-derived forest cover maps. *Proceedings, Accuracy 2000, 4th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, 12–14 July 2000* (Amsterdam: The Netherlands).

Reynolds, R.T., 1983, Management of western coniferous forest habitat for nesting Accipiter hawks. General Technical Report RM-102, US Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station, Fort Collins, Colorado.

Reynolds, R. T., 1989, Accipiters. *Proceedings of the Western Raptor Management Symposium and Workshop, 26–28 October 1987, National Wildlife Federation Scientific and Technical Series No. 12* (Washington, DC: National Wildlife Federation), pp. 92–102.

Reynolds, R. T., Graham, R. T., Reiser, M. H., Bassett, R. L., Kennedy, P. L., Boyce, D. A., Jr., Goodwin, G., Smith, R., and Fisher, E. L., 1992, Management recommendations for the northern goshawk in the southwestern United States. General

Technical Report RM-217, US Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station, Fort Collins, Colorado.

Reynolds, R. T., Joy, S. M., and Leslie, D. G., 1994, Nest productivity, fidelity, and spacing of northern goshawks in northern Arizona. *Studies in Avian Biology*, **16**, 106–113.

Speiser, R., and Bosakowski, T., 1984, History, status, and future management of goshawk nesting in New Jersey. *Records of New Jersey Birds*, **10**, 29–33.

Statistical Sciences, 1995, *S-PLUS© 3.3Statistical Software Package for Personal Computers* (Seattle, Washington: StatSci Division, MathSoft, Inc.).

Stehman, S. U., 1996, Estimating the kappa coefficient and its variance under stratified random sampling. *American Society of Photogrammetry and Remote Sensing*, **62**, 401–402.

Stehman, S. U., 1999, Basic probability sampling designs for thematic map accuracy assessment. *International Journal of Remote Sensing*, **20**, 2423–2441.

Tou, J. T., and Gonzalez, R. C., 1974, *Pattern Recognition Principles* (Reading, Massachusetts: Addison-Wesley).

Van Horne, B., 1983, Density as a misleading indicator of habitat quality. *Journal of Wildlife Management*, **43**, 893–901.

White, M. A., and Vankat, J. T., 1993, Middle and high elevation coniferous forest communities of the North Rim region of the Grand Canyon National Park, Arizona, USA. *Vegetation*, **109**, 161–174.