Predictions of Forest Inventory Cover Type Proportions Using Landsat TM

Steen Magnussen, Paul Boudewyn, Mike Wulder and David Seemann

The feasibility of generating via Landsat TM data current estimates of cover type proportions for areas lacking this information in the national forest inventory was explored by a case study in New Brunswick. A recent forest management inventory covering 4196 km^2 in south-eastern New Brunswick (the test area) and a coregistered Landsat TM scene was used to develop predictive models of 12 cover type proportions in an adjacent 4525 km^2 region (the validation area). Four prediction models were considered, one using a maximum likelihood classifier (MLC), and three using the proportions of 30 TM clusters as predictors. The MLC was superior for non-vegetated cover types while a neural net or a prorating of cluster proportions was chosen for predicting vegetated cover types. Most predictions generated for national inventory photo-plots of 2×2 km were closer to the most recent inventory results than estimates extrapolated from the test area. Agreement between predictions and current inventory results varied considerably among cover types with model-based predictions outperforming, on average, the simple spatial extensions by about 14 %. In this region, an 11-year-old forest inventory for the validation area provided estimates that in half the cases were closer to current inventory estimates than predictions using the optimal Landsat TM model. A strong temporal correlation of photo-plot-level cover type proportions made old-values more consistent than predictions using the optimal Landsat TM model in all but three cases. Prorating of cluster proportions holds promise for large-scale multi-sensor predictions of forest inventory cover types.

1 Introduction

Statistics on the current extent of land-cover types are needed for a plethora of natural resource planning and policy purposes (Apps et al. 1995, Dale and Robinson 1996, Noss et al. 1997, Odum 1996, Scott et al. 1996). Increased demands for reporting the state of the environment has accentuated the need for statistics stating the extent of land cover types (White et al. 1992). Estimates of cover-type proportions are pivotal statistics suited to quantify strategic issues and trends. Detailed spatial knowledge about cover type distributions in the landscape is needed for efficient implementation and monitoring of specific objectives and plans (Andersen 1988, Costanza et al. 1992). Compilation of current land-cover statistics is a challenge for large countries like Canada (Leckie and Gillis 1995). Although parts of the country are inventoried each year many regions are without a current land-cover inventory (Gray and Power 1997). Estimates of cover type proportions for these regions are lacking, incomplete or dated. Remote sensing has the potential to fill inventory gaps and provide approximations for the missing estimates (Ahern et al. 1996, He et al. 1998). Yet development of a remote sensing classifier using verified ground data and implementing an independent validation/calibration process are both costly and time-consuming (Congalton and Biging 1992) activities, even within a sampling context (Bauer et al. 1994, Moody 1998). However, if the missing data is primarily cover type proportions for landscape segments several hectares in size, cheaper and more expedient shortcuts should be possible.

The objective of this study is to explore, within the framework of a national inventory, an expedient use of Landsat TM imagery and existing regional forest inventories to make predictions of forest cover type proportions for areas for which there are no or only dated estimates. We shall evaluate our predictions against current inventory estimates and against estimates from a 11-year-old inventory. Predictions are assessed on how well they match estimates that would otherwise have been forthcoming from a local inventory to the national inventory (Lowe et al. 1994). Predictions of the proposed nature would greatly expand the domain with up-to-date land-

cover information. The tested methods are both fast and cheap as they rely on existing inventories and readily available software algorithms.

Our approach differs from the now classic combination of remotely sensed data and ground inventory in a double sampling framework (Cochran 1977, Czaplewski and Catts 1992, Tenenbein 1972) in that the reference (training) data come from an area outside the area for which estimates are needed. Furthermore, estimates of cover type proportions are produced for 2×2 km area units with no attempt to produce a map. Unbiasedness is only assumed and not given by design (Gregoire 1998).

2 Material and Methods

2.1 Study Site and Photo-plots

An 8721-km^2 region in the southeastern part of New Brunswick was used as the study site (Fig. 1). Approximately one half of the area (TEST) was used to train and test various predictive models while the other half (VALID) was used for comparing final model predictions against forest inventory results. The test area was again split into two non-overlapping halves (TEST1 and TEST2) in a checkerboard fashion, one half for model fitting (TEST1) and the other (TEST2) for model comparisons. Models chosen based on their ability to predict cover type proportions in TEST2 were then refitted using all data in the test area (e.g. TEST1 + TEST2) and used to make predictions for the validation area. From general considerations of climate, soils and ecological conditions, it was assumed that the cover types found in the test area would be representative for the cover types in the validation area. This assumption proved warranted; one rare cover type (bryoids) in the test area (0.003 %) was not documented for the validation area.

Cover type proportion estimates are desired for units of area called "photo-plots". The recommended size for these photo-plots is 400 ha in a square 2×2 km area ("A Plot-based National Forest Inventory Design For Canada". Natural Resources Canada. 1999. Canadian Forest Service, Pacific For. Centre, Victoria, BC. 70 p, unpubl.) The study area was therefore subdivided
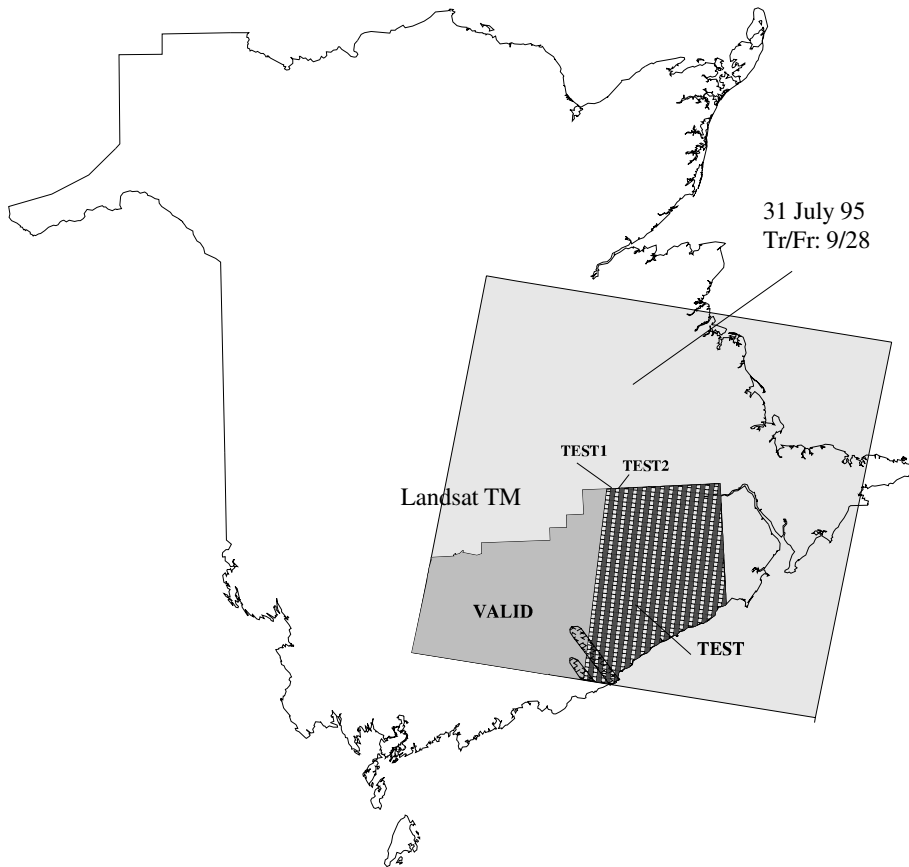
Fig. 1. Study area and Landsat TM image coverage in New Brunswick.
TEST1(dark squares): Data from this area is used to develop models.
TEST2 (light squares): Model predictions for TEST2 are compared to current inventory estimates in order to find cover type specific 'best' model.
TEST: Cover type specific 'best' models are refitted to data from the combined TEST1+TEST2 area
VALID: Area used for model validation only.

into a grid of 2×2 km photo plots, 1193 photo plots inside the validation area and 1149 photo plots in the test area. These photo plots are the units for predicting cover types.

2.2 Forest Inventory Cover Types

Forest cover type maps in a GIS format (New Brunswick Department of Natural Resources and Energy, Forest Management Branch) for the entire study area were used as benchmark data for training and fitting purposes. The cover type maps were based on photo interpretation of 1:12 500 air-photos taken during the months of July, August and early September of 1993 (test area) and 1994 (validation area). The forest inventory was completed in 1993 for the test area and in 1993–1994 for the validation area. Inventory cover types, as listed for specific polygons on the forest cover maps, were grouped into a hierarchical system compatible, as far as possible, with the classification system currently under development for the National Forest Inventory. Assignment of inventory labels to the national scheme was guided by over 30 inventory attributes associated with each polygon. Relabeling existing provincial inventory cover types to a unified national scheme is not without problems (Carpenter et al. 1999, Foody 1999) as

353

Table 1. Hierarchy of national cover types and provincial forest inventory polygon labels. Unlisted cover types are bulked as 'Other'. Level I = Vegetated / non-vegetated, Level II = treed, non-treed, land, water, Level III = hardwood, mixedwood, softwood, ..., ocean.

Cover types	Examples of forest inventory polygon map labels
Vegetated	Vegetation covers at least 5 % of the area
Treed	Crown closure ≥ 10 %
hardwood	hardwood species percentage ≥ 80
mixedwood	hardwood species percentage 21–79
softwood	softwood species percentage ≥ 80
Non-treed	crown closure < 10 %
shrub – low	BL (barren land), WL (wetland), CB (cultivated blueberry), NP (non-productive), BB (burn)[1], CC (cut)[1], PC (partial cut)[1], PL (planted)[1], TI (thinned)[1], SR (scarification)[1]
shrub – tall	AC (alders on cut-over), AF (alders on field), AP (alders on poor site), AW (alders near water), HW (cut with hardwood regeneration), SW (cut with softwood regeneration), BB (burn)[2], CC (cut)[2], PC (partial cut)[2], PL (planted)[2], TI (thinned)[2], SR (scarification)[2]
bryoids	CP (cultivated peat bog)
crop/pasture	CL (cultivated farm land), FP (fallow pasture), AG (agriculture), F (field)
Non-Vegetated	Vegetation covers less than 5 % of the area
Land	
exposed	GP (gravel pit), RD (road), RR (railroad), AI (airstrip), AR (abandoned railway), OC (occupied land – cities, towns), IZ (military impact zone), WR (winter road), RW (winter road)
rock/rubble	MI (mine), QU (quarry), RO (rock outcrop)
Water	
lake/pond	LK (lake), PN (pond), WA (water)
river/stream	RV (river)
ocean	ON (ocean)

1. Recently disturbed polygons (disturbance after 1990) currently dominated by young and emerging vegetation. A few (maximum 143) polygons with a combined area of maximum 183 ha were labeled as partial cuts or thinning (PC, TI) with 0 % crown closure. The inventory information indicated 'low shrub' as the most likely cover type for these polygons.
2. Polygons disturbed between 1983 and 1990 and currently dominated (area) by a tree/shrub vegetation > 2m tall. A few (see note 1 above) polygons with incomplete records may have erroneously been placed in this category.

definitions prevented a perfect conversion. Discussions with provincial inventory experts helped decide a 'most likely' classification in many cases, and defaulting to past conversion rules used in Canada's National Forest Inventory (Gray and Power 1997) were used in remaining unresolved cases. Only the aggregate cover types are used in the analyses. The first three levels of the aggregated cover types are listed in Table 1 along with the complement of forest inventory cover types. Cover type proportions are listed in Table 2.

To check whether 'old' inventories could provide estimates of current cover type proportions that were in some sense superior to what could be predicted with our models (see below) we added an inventory from 1983/84 to our data sets. Due to minor changes in inventory proce-

dures we discarded about 15 % of the 'old' polygons which could not be classified with any confidence. We used simple prorating of areas to compensate for excluded areas (Little and Rubin 1987).

2.3 Landsat Image Data

Landsat 5 TM scene number (Track 9, Frame 28) from July 31 1995 covered the entire study area. This scene was deemed the better choice among alternatives from August 1993, and July/August 1994, and July 1995/96 as these had either more cloud cover or were less temporally synchronized with the ground inventories. Spatial resolution was 30 × 30 m for channels 1–5,

Table 2. Cover type distribution in the test area and model dependent mean absolute difference (MAD %) between Landsat TM model-based predictions and inventoried cover type proportions in 2×2 km photo plots within TEST2. Models were fitted to inventory land cover proportions in area TEST1. The row wise minimum of MAD % is typed in Bold. ΔMAD(%) is the relative reduction in MAD % achieved by using the best model instead of SPEX.

Cover-type			Test-area %	MAD % by model					ΔMAD (%)
				NNET	ACE	PRO	MLC	SPEX	
Vegetated			96.61	**1.12**	1.71	1.51	1.49	2.01	38
	Treed		82.84	**4.80**	9.27	6.51	5.28	13.40	66
		hardwood	21.61	**7.69**	15.34	8.96	10.17	12.08	37
		mixedwood	31.84	**8.57**	11.95	8.98	8.74	10.28	17
		softwood	29.32	**8.85**	21.10	12.10	10.64	14.86	40
	Non-treed		13.40	**4.44**	7.80	5.85	5.00	11.07	60
		shrub – low	3.87	2.24	5.31	**1.62**	2.09	3.04	47
		shrub – tall	1.98	1.31	**1.13**	1.58	1.40	2.47	46
		bryoids	–	–	–	–	–	–	–
		crop/pasture	7.55	**2.08**	8.64	4.86	3.70	9.01	77
Non-Vegetated			3.38	**1.12**	1.68	1.46	1.49	2.00	44
	Land		2.92	1.04	2.10	1.33	**0.94**	1.69	56
		exposed	2.82	1.06	1.75	1.31	**1.00**	1.95	49
		rock/rubble	0.10	0.01	0.03	0.02	**0.00**	0.06	92
	Water		0.47	0.29	0.36	0.29	**0.13**	0.51	75
		lake/pond	0.18	0.06	0.08	0.05	**0.00**	0.19	97
		river/stream	0.27	0.18	0.03	0.25	**0.00**	0.31	99
		ocean	0.02	0.02	0.04	0.00	**0.00**	0.01	99
Other			0.00	0.02	0.03	**0.01**	n.a.	0.01	13
	Other		0.38	**0.08**	0.13	0.32	–	0.26	69
		other	0.44	**0.09**	0.60	0.00	–	0.33	73

and 7 as opposed to 120×120 m for channel 6 (thermal).

The Landsat TM scene was geometrically corrected using the PCI utility "GCP Works". Vector data from New Brunswick Forest Inventory Maps (scale 1:12 500) were used to provide ground control points. A total of 37 control points were gathered throughout the image utilizing features such as road intersections, bridges, etc. An overall root mean square error of 0.46 pixels was attained using a first order nearest neighbor sampling technique. Radiance values were given in an 8-bit format. No radiometric correction was done. First, accurate correction factors were not available, and secondly, we do not expect such values to be readily available for the type of areas that we envision for the application domain of our methods. Further, it was also deemed desirable to maintain the full dynamic range of the radiance values.

Forest cover type maps were then co-registered with the Landsat TM scene and rasterized to a grid size compatible with the Landsat TM resolution. Vector data within the image was used to ensure that the co-registration was correct to within a pixel. Further details are given in Wulder (1998). Two streaks of coastal fog covering about 144 km^2 (18 photo plots in the test area, and 16 photo plots in the validation area) and clearly visible in southern part of the image (Fig. 1) were left 'as is' in the data.

2.3.1 Principal Component Scores

To reduce redundancy and potential collinearity between Landsat TM channels we extracted the first four principal components accounting for 94.5 % of the total variation in the 7 channels (Rencher 1993), and computed four principal

component scores for each of the $9.7 \cdot 10^6$ pixels in the study area ($4.7 \cdot 10^6$ in the test area and $5.0 \cdot 10^6$ in validation area). Scores of each retained principal component were multiplied by the square root of the variance explained by the component. Preliminary analyses with the six channel TM data plus various vegetation indices (Rouse et al. 1974) treated as additional channel data failed to isolate significant information in the vegetation indices not already accounted for by a linear combination of the six channels.

2.3.2 Unsupervised Clustering

A pre-stratification of principal component scores into disjoint clusters (Lillesand 1996) was subsumed to enhance our ability to predict cover types from Landsat TM data. Formed clusters are assumed associated with one or a few inventory cover type classes. If true, it would be possible to exploit this association in a regression framework as outlined in section 2.4 where we quantify the association between clusters and cover types.

In an unsupervised clustering (Hartigan 1975) of PC-scores, with a predetermined number of clusters $n_{clu} = \{3, 6, 9, ... ,33, 40, 50, 60, 70\}$ and random initial centroids, the assignment of pixels to clusters was done when the maximum change in cluster centroids, between two iterations, fell below 0.02. About 140 iterations were needed to meet this criterion for $n_{clu.} = 70$. Three scaled (to the 0–1 range) and penalized clustering criteria: i) cubic clustering criteria (Milligan and Cooper 1985), ii) R^2 (Hartigan 1975), and iii) a within-cluster relative diversity index of cover types assisted in deciding on how many clusters to maintain. Ideal clusters would represent only one cover type (a diversity index of 0). For a particular n_{clu} value the mean diversity index (D) was estimated as:
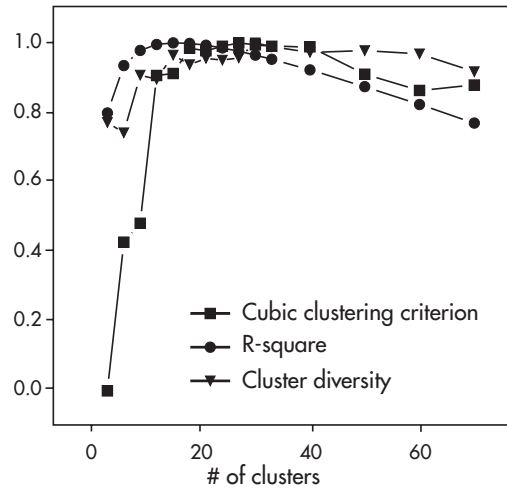

Fig. 2. Three indices of cluster separation/homogeneity (y-axis) versus number of clusters (x-axis). Indices are constrained to the interval [0,1]. See text for details on indices.

where E stands for an expectation (average) and S is Simpson's diversity index (Patil 1982). Level refers to the three classification levels (Table 1). Penalty factors ($n_{plots} - n_{cover\,types}$ x $n_{clusters} - 1$) / ($n_{plots} - n_{covertypes} - 1$) were levied to account for the decline in degrees of freedom of the residual terms of any prediction model as a result of adding more clusters as predictors. Fig. 2 illustrates the trends in the three indicators. Fifteen to 35 clusters appear to perform equally well. Our final choice was 30, coincident with the maximum of the cubic clustering criterion. Fig. 3 presents an example of the clustered TM-image (30 gray tones) and the corresponding forest inventory cover type map.
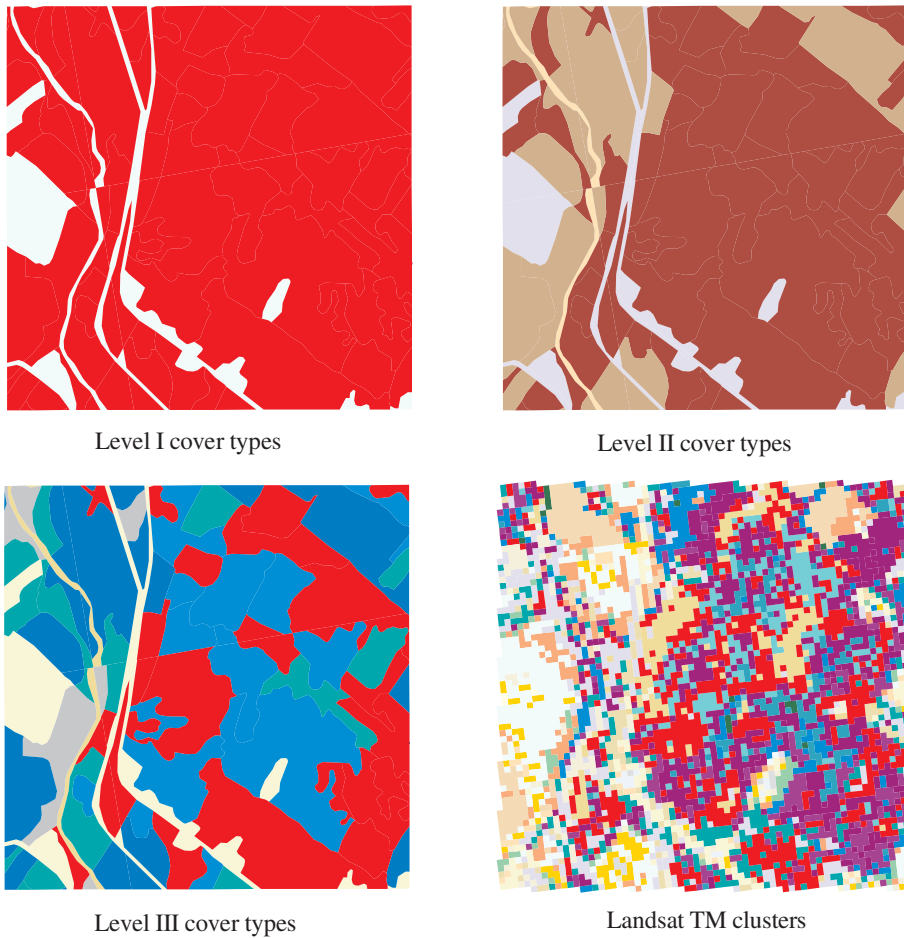
$$\hat{D} = 1 - \frac{\underset{level}{E}\ \underset{clusters}{E}\ (S \,|\, \text{level})}{\underset{level}{E}\ \underset{clusters}{E}\ (S \,|\, \text{random assoc. of level cover types to clusters})} \qquad (1)$$

Level I cover types   Level II cover types

Level III cover types   Landsat TM clusters

Fig. 3. Examples of forest inventory cover type maps and clustered (1–30) Landsat TM pixels of a randomly selected 2×2 km National Forest Inventory photo-plot.

2.3.3 *Comparing TM Data from the Test and Validation Areas*

Statistical tests of the null hypotheses of no differences between PC scores and cluster proportions in the test and the validation areas were done with either Hotelling's T^2 tests or Chi-square goodness of fit tests (Rencher 1993).

2.4 Predicting Cover Type Proportions from Cluster Proportions

Clusters generated by the unsupervised clustering are assumed associated with one or a few inventory cover types. If so, we would also expect proportions of clusters in an area (here 2×2 km) to be associated with the proportions of cover types in this area (Metternicht and Fermont 1998). We do not expect this association to be perfect nor simple. To explore what could be a very complex relationship we tested two flexible non-linear models: neural net (NNET), and alternating conditional expectations (ACE) as the models most likely to succeed (Breiman and Friedman 1985, Carpenter et al. 1997, Lek et al. 1996, Moody et al. 1996, Warner and Misra 1996).

Should the relationship between cluster proportions and cover type proportions, however,

turn out to be linear, then linear models should perform well. Two linear models were tested, one that simply exploits conditional probabilities of a cover type given the cluster (PRO), and a spatial extension model (SPEX) which simply extrapolates results from one area to adjacent areas.

The generic model applicable to all tested models is:

$$g(\vec{P}_{li}) = f(\vec{C}_i), \quad \vec{P}_{li} \cdot \vec{J}_l = 1, \quad \vec{C}_i \cdot \vec{J}_{clu} = 1 \quad (2)$$

where $\vec{P}_{li}$ is a $1 \times n_l$ vector of proportions of the n_l cover types within classification level l ($l = I, II, III$, see Table 1) in the ith photo-plot, $\vec{C}_i$ is a $1 \times n_{clu}$ vector of cluster proportions (n_{clu} is the number of defined clusters), J_{clu} is a $n_{clu} \times 1$ vector of ones, J_l a $n_l \times 1$ vector of ones and g and f are two monotone functions of proportions. The last two equations in (2) are sum-to-one restrictions.

In the NNET model f is replaced by a feed-forward neural net (Lek et al. 1996) with one hidden layer taking $\ln(\vec{C}_i + 10^{-9})$ as input and with $\ln(\vec{P}_{li} + 10^{-9})$ as the target output (during training only). The decay rate for back-propagated residuals through the network was a constant 0.05 (Venables and Ripley 1994). NNET predictions were back-transformed to proportions. Trials with several other network configurations (number of layers, weights, and decay rates) on TEST1 data yielded inferior results.

In the ACE model f and g are monotone transformations of cluster and cover type proportions maximizing the correlation between the transformed proportions (Breiman and Friedman 1985).

In the PRO model f is a matrix of conditional probabilities $P(cover\ type \mid cluster)$ obtained from the training area and pre-multiplied to the cluster proportions in a given photo-plot, and g is an identity matrix pre-multiplication of $\vec{P}$. $P(cover\ type \mid cluster)$ is thus the proportion of pixels of a given cover type within the training area classified to a specific cluster. In other words a PRO prediction is the probability $P(cover\ type \mid photo\text{-}plot) = P(cover\ type \mid cluster) \cdot P(cluster \mid photo\text{-}plot)$. Thus f is essentially a 'confusion' matrix (Steele et al. 1998) and predictions are compatible with the classical 'inverse' method (Magnus-

sen 1997a, Walsh and Burk 1993). $P(cover\ type \mid cluster)$ is obtained from the training area and $P(cluster \mid photo\text{-}plot)$ is merely the 'observed' cluster composition of a photo-plot in the area for which predictions are to be made.

The SPEX model used the mean cover-type proportions of the training area as the predictions for the testing area; specifically, $\vec{P}_{li} = E(\vec{P}_l \mid training\ area)$. In other words, the area for which predictions are needed is assumed to be identical to the training area.

Linear combinations of NNET, ACE, and PRO predictions were also tested as a viable hybrid prediction model (Leblanc and Tibshirani 1996, Peña 1997). However, no tangible improvements could be obtained by any weighted linear combination of individual model predictions (weights assigned to each model prediction were proportional to the median absolute residuals obtained in the training area with NNET, ACE and PRO).

All predictions from NNET, ACE, and PRO were proportionally adjusted and calibrated to satisfy a sum-to-one constraint and to ensure that all predictions were between 0 and 1. Predictions at a specific classification level were harmonized by iterative proportional fitting (Bishop et al. 1975) to match predictions at the next higher level.

2.5 Maximum Likelihood Classification (MLC)

MLC is generally considered a standard method for classification (McLachlan 1991) and we used MLC as an alternative to the above cluster-based models. Specifically we applied a quadratic Gaussian maximum likelihood classifier of principal component scores (see 2.3.1). The MLC used separate within-cover-type covariance matrices as a hypothesis of homogeneous covariance matrices was rejected ($P = 0.001$, χ^2-test, Rencher(1993) page 272–273). The MLC classifier was trained on principal component scores of the pixels in the TEST1 area and tested on the pixels in TEST2. TEST2-predictions were then summarized for the 588 photo plots in this area and compared to test results from four alternatives (see below). Given the objective of predicting cover type proportions for a validation area

assumed similar to the training area we used the appropriate cover type proportions from the TEST area(s) as priors for the classifier (Frigessi and Stander 1994). Incidentally, the assumption of normality was violated for the first principal component ($P = 0.001$) due to an excess kurtosis and right skewed distributions of principal component scores (D'Agostino et al. 1990). No simple transformation achieved normality; consequently we accepted that classifications may be less than optimal (McLachlan 1991). Classification accuracy of pixels in the TEST2 area varied from a high of 97 % for separating vegetated from non-vegetated to 92 % for the second classification level (Table 1) and 51 % in the third level.

2.6 Temporal Extension of Old Inventory Results (OLD)

One of our subsidiary objectives was to compare TM-based predictions with estimates from dated inventories. In areas of little change an old inventory might still produce estimates of broad cover type proportions that are as good as TM-based model predictions. Cover type proportions as per the old inventory were simply used as estimates of current proportions. Estimates are, in all cases, only applied to the area (photo plot) from which they were derived.

2.7 Evaluating Cover Type Specific Prediction Models

In practical applications the choice of model to predict cover type proportions rests with its ability to predict consistently and with minimum divergence from established benchmarks. To establish which of the tested models is likely to produce the most consistent and least divergent predictions for a specific cover type, we first fitted all models to data from TEST1 and then compared their predictions for the TEST2 area to corresponding inventory estimates. Under practical circumstances this would complete the model evaluation and a decision on preference. For each cover type, the best model would then be refitted using data from both TEST1 and TEST2

and used to make predictions for an area with no or only dated inventory information. Everything else being equal, the model producing the lowest mean absolute deviation for a given cover type is also expected to produce the minimum absolute deviation when used to generate predictions. Note, however, that this selection does not imply that the chosen model is statistically superior. From the estimated mean absolute differences it is possible to conclude that NNET and MLC, on average, produce predictions of equal quality.

Predictions of cover type proportions in TEST2 photo plots from MLC, NNET, ACE, PRO, and SPEX were compared against current inventory results. Differences simply mean a lack of agreement as both the prediction as well as the inventory could be in error. A lack of agreement simply quantifies how much a Landsat-based summary of cover types would deviate from a summary based on a provincial inventory. The model producing the minimum average absolute deviation (Tauer 1983) for a given cover type was chosen to predict cover type proportions for photo-plots in the validation area.

Predictions made for the photo-plots in the validation area with the cover-type-specific chosen (best) models were compared to current inventory estimate by computing the mean absolute deviation.

3 Results

3.1 Model Comparisons Based on TEST2 Predictions

Results from the comparison of model predictions with inventory estimates for TEST2 photo-plots are shown in Table 2. NNET produced the lowest mean absolute deviations for all but two of the vegetated cover-types. MLC, on the other hand, was superior for predicting non-vegetated types. MLC was in all but two cases among the best two models. ACE, despite strong correlation (>0.8) between transformed cluster and cover-type proportions, resulted in five times more divergence compared to a simple extension of the test results (SPEX). PRO was, on average, at par with MLC. The SPEX model produced, in general, the largest discrepancies between pre-

dictions and inventory estimates. From this we expected that relative to SPEX, the best model for each cover type would lower the mean absolute difference between a predicted cover type proportion and the actual cover type proportion of an inventory by 38 % for level I cover types, 64 % for level II, and 36 % for level III cover types. Table 2 shows the cover type specific models chosen to best predict cover type proportions for the validation area. NNET was chosen ten times, MLC seven times, and PRO and ACE were each chosen once. The SPEX model was not chosen.

3.2 Model Predictions for the Validation Area

With the cover-type-specific preferred models predictions of inventory cover type proportions for photo-plots in the validation area were, on average, significantly closer to the inventory estimates than predictions obtained by spatially extending results from the TEST area (SPEX). Predictions were also, on average, closer than the mean cover type proportions of the validation area ($P = 0.001$, test: Chi-square goodness of fit, Miller 1980). Predictions improved, in general, with an increased aggregation of cover types as one moves from level III to level I in the classification system. At the top level the reduction of differences amounted to 2/3 while at the most detailed the reduction was no more than about 1/6. Similarly, correlation coefficients between predicted and observed cover type proportions declined from 0.82 ($P = 0.001$) at the top level to an average of 0.49 ($P = 0.001$) at the third level. Non-vegetated predictions were exclusively based on the maximum likelihood classifier while the vegetated predictions, with two exceptions were based on neural net fitting involving unsupervised clustering. A simulation study indicated that the quality of our predictions was commensurate with that expected from an actual inventory of 6–8 % of the validation area.

Table 3 details the predicted and actual cover type proportions and the mean absolute difference of the predictions (PRED). Results obtained by spatial extension of mean proportions from

the test area (SPEX) and a temporal extension of the 'old' inventory (OLD) are included for comparison. Predictions using the optimal Landsat TM model of vegetated area were, on average, within 2 percentage points of the actual proportion while the individual photo-plot predictions were, on average, within about 9 percentage points of the comparable inventory figures. At this coarse level the model predictions were about 50 % closer to the target than spatially extended averages. As the shifts from vegetated to non-vegetated were rather limited during the last decade, the 11-year-old inventory is closer to the current state than our predictions. Had we used the 'old' inventory for our predictions, the mean absolute deviations would have been about 1/4 as large. A strong temporal correlation of photo-plot proportions (~0.98) explains this observation.

At the next more detailed level (level II; Table 1) the benefit of predicting via Landsat TM images is striking. Average predictions fall within 1–2 percentage points of the inventory values while SPEX and OLD mostly miss them by 5–11 and 3–8 percentage points, respectively (Table 3). The TM-based predictions appear especially promising for the treed, land, and water portions. Photo-plot level disagreements in predicted cover type proportions were generally smallest for the 'old' inventory, followed by PRED, with SPEX a close third. Correlation coefficients between old and new inventory attributes at this level averaged 0.92.

Predictions of hardwood, mixedwood, and softwood proportions in the treed group were prone to consistent and substantial deviations from the inventory results (Table 3). Only mixedwood predictions matched the inventory estimates. SPEX and OLD had also apparent problems in predicting softwood and mixedwood proportions. Agreement between 'old' and new inventory estimates of softwood, mixedwood, and hardwood as epitomized by a correlation coefficient was 0.78, 0.71, and 0.72, respectively.

In the non-treed group it was clearly difficult to discriminate between low-shrubs and crop/pasture. Spatial extension was only marginally better while the 'old' inventory was disadvantaged due to the aforementioned large shift in the non-treed group between the two inventories;

Table 3. Inventory (INV) and model-based predictions of cover type proportions (in %). Mean absolute differences (MAD) between inventory estimates and a prediction (PRED, SPEX, OLD) are in percentages. PRED are predictions made with the model chosen as best for a specific cover type. SPEX are extension of test results to the validation area. OLD are temporal extension of 'old' inventory results for the validation area. Displayed numbers may not add to 100 % due to rounding.

Cover-type			Estimate/Prediction			MAD (%)		
		INV	PRED	SPEX	OLD	PRED	SPEX	OLD
Vegetated		85.4	87.4	96.6	86.6	8.9	13.1	2.2
Treed		71.4	72.5	82.8	64.9	10.3	20.0	8.3
	hardwood	16.3	24.5	21.6	15.3	13.5	13.0	8.2
	mixedwood	33.0	33.2	31.8	28.4	12.7	14.2	10.5
	softwood	21.9	14.8	29.3	21.2	12.8	15.0	7.7
Non-treed		13.8	14.9	13.4	21.7	11.8	10.6	9.4
	shrub – low	7.0	4.0	3.9	4.1	6.1	6.1	3.5
	shrub – tall	1.7	1.7	2.0	10.2	2.7	2.1	9.0
	bryoids	0.0	–	0.0	0.0	–	0.0	0.0
	crop/pasture	5.1	9.2	7.6	7.4	7.6	7.9	1.5
Non-Vegetated		14.5	12.6	3.4	13.1	8.6	13.1	2.4
Land		8.1	6.9	2.9	3.8	6.8	7.3	4.2
	exposed	8.0	6.7	2.8	3.8	6.9	7.3	1.2
	rock/rubble	0.0	0.2	0.1	0.0	0.2	0.0	0.0
Water		6.4	5.7	0.5	9.2	4.6	6.5	3.1
	lake/pond	2.9	3.0	0.2	9.1	4.1	3.0	5.6
	river/stream	3.5	2.7	0.3	0.1	4.3	3.6	4.3
	ocean	0.0	0.1	0.0	0.0	0.0	0.0	0.0
Other		0.1	0.1	0.0	0.3	0.1	0.1	0.2
Other		0.4	0.1	0.4	0.7	0.9	0.6	0.3
	other	0.5	0.1	0.4	0.7	0.8	0.4	0.2

changes that are a mix of real changes and method-dependent artifacts. Disagreement between predictions and inventory estimates for this group is, in general, about the same magnitude as the average predicted proportion. The exception is again the 'old' inventory that benefits from the strong temporal correlation of cover types within photo plots (avg. 0.72).

Predictions of third-level non-vegetated cover-types were superior to either spatial and temporal extensions. The maximum-likelihood-based pixel-level predictions were significantly ($P = 0.01$) closer to estimates from the current inventory than the two alternatives. Yet, individual photo plot predictions were associated with relatively large departures from the inventory.

All predictions for the 16 photo plots covered by at least 30 % coastal fog were significantly ($P = 0.01$) worse than other predictions. Predictions for partially obscured photo plots were, on aver-

age, about four times further away from the inventory estimates than unobscured photo plots.

The distribution of differences between model-based and inventory estimates of cover type proportions is an important aspect of model performance. Fig. 4 highlights eight typical examples. Long tails of 'gross' differences and lack of symmetry was evident in most distributions. Also, the split between variance and bias differs between PRED, SPEX and OLD. If the concern is to minimize excessive disagreement then it is almost always 'safer' to use the decade-old inventory as it limits the disagreement to the amount of actual change in a photo plot plus any artifacts caused by changes in inventory procedures. With both PRED and SPEX the chance of an excessive photo-plot-level disagreement (defined as a disagreement that is twice the predicted value) is between 30 and 40 %.
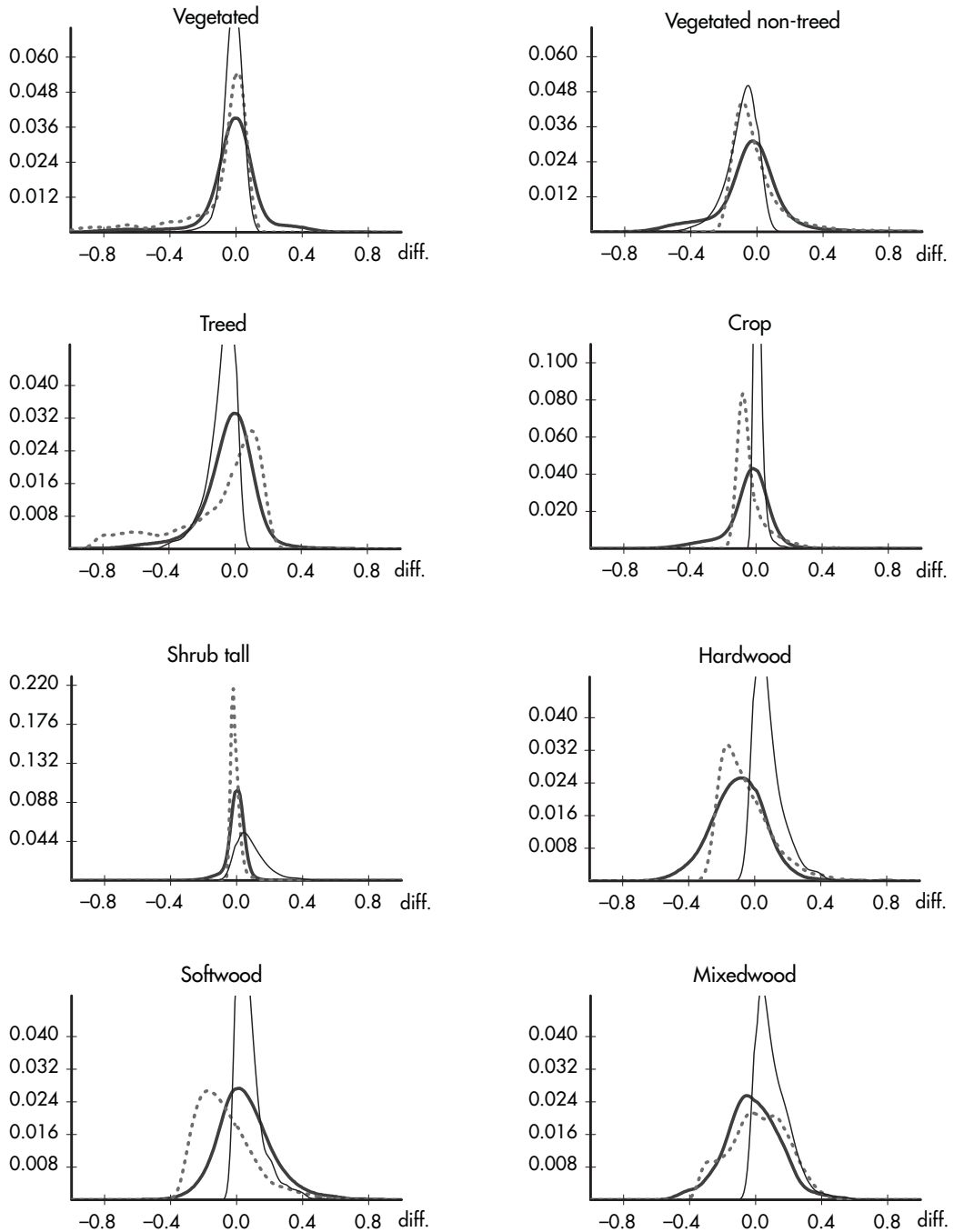
Fig. 4. Distributions of the difference between model-based predictions and inventory estimates of cover type proportions (in %) for eight cover type in the validation area. Thick line: predictions made with neural nets or maximum likelihood classifiers. Dashed line: differences between the mean proportions in the test area and current inventory estimates for the validation area. Thin line: differences between 'old' and current inventory estimates.
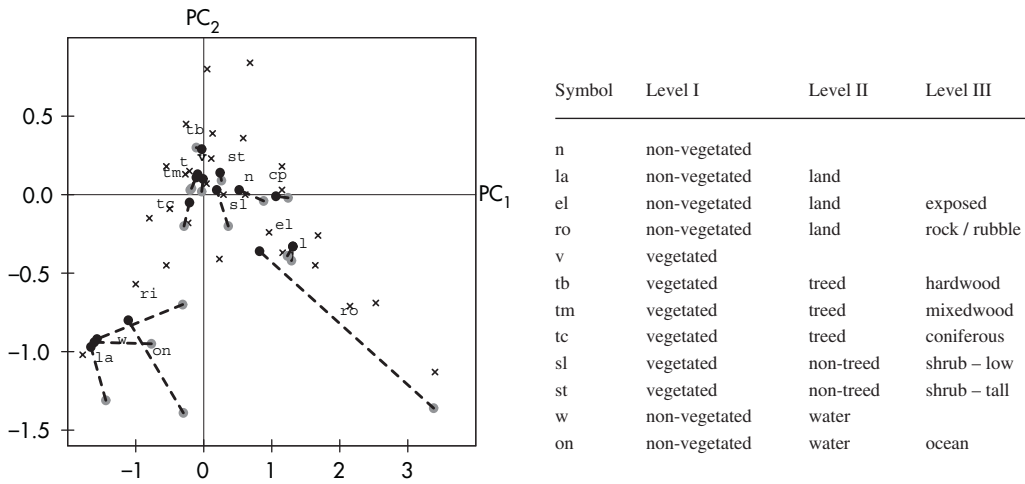
Symbol	Level I	Level II	Level III
n	non-vegetated		
la	non-vegetated	land	
el	non-vegetated	land	exposed
ro	non-vegetated	land	rock / rubble
v	vegetated		
tb	vegetated	treed	hardwood
tm	vegetated	treed	mixedwood
tc	vegetated	treed	coniferous
sl	vegetated	non-treed	shrub – low
st	vegetated	non-treed	shrub – tall
w	non-vegetated	water	
on	non-vegetated	water	ocean

Fig. 5. Bi-plot of the means of the first two principal components for various cover types in the test and the validation area. Scores on first principal components are along the horizontal axis and scores on the second principal component are along the vertical axis. Gray circles: test area means. Black circles: validation means. Corresponding means have been connected by a dashed line.

3.3 Comparing Results from Testing and Validation

Discrepancies between predicted and inventory estimates of cover type proportions in the validation area were substantially higher than those reported when the best model was tried on the TEST2 area. In the vegetated parts the deviations were 2–4 times as high as comparable figures from the test area. In non-vegetated parts the deviations were even higher (6–8 times). Several factors contribute to this phenomenon. Spectral feature differences between the test area and the validation area were apparent. For example, identically labeled cover types in the two areas had different average spectral reflectance values. Fig. 5 captures the essence of this by plotting and connecting the means of the first two principal components in the TEST area and the validation area for the most important cover types. Hotelling's T^2-tests of equal area means of the four principal components rejected the null hypothesis of no differences for all cover types ($P = 0.001$). The same test for a single photo plot failed to identify significant differences ($P = 0.52$). The shifts in mean principal component values were for the most part large enough to modify the association between cover types and clusters. Classification of most non-vegetated cover types by maximum likelihood is very sensitive to such shifts in the centroids of the principal components. As well, the variances of principal component scores for pixels in the cover types rock, river/stream, lake/pond, water and ocean changed significantly ($P = 0.001$) between the test and the validation sites and compromised predictions.

The tight clustering of cluster centroids in the upper part of the illustrated feature space (Fig. 5) generates disproportionately large shifts in the association between cluster and cover-type proportions for even modest shifts in the cover type centroids between the test and validation areas. This is illustrated in Fig. 6 where the composition of each cover type in terms of image clusters is given for the test and the validation area. Only few cover types (water, lake/pond/, ocean, and rock) are strongly associated with one or two clusters. Most cover types are represented more or less equally by 4–12 clusters. Shifts between the two areas are legion. A global test of difference in proportions between the test and validation area led to a rejection ($P = 0.001$) of the null hypothesis while only non-vegetated,
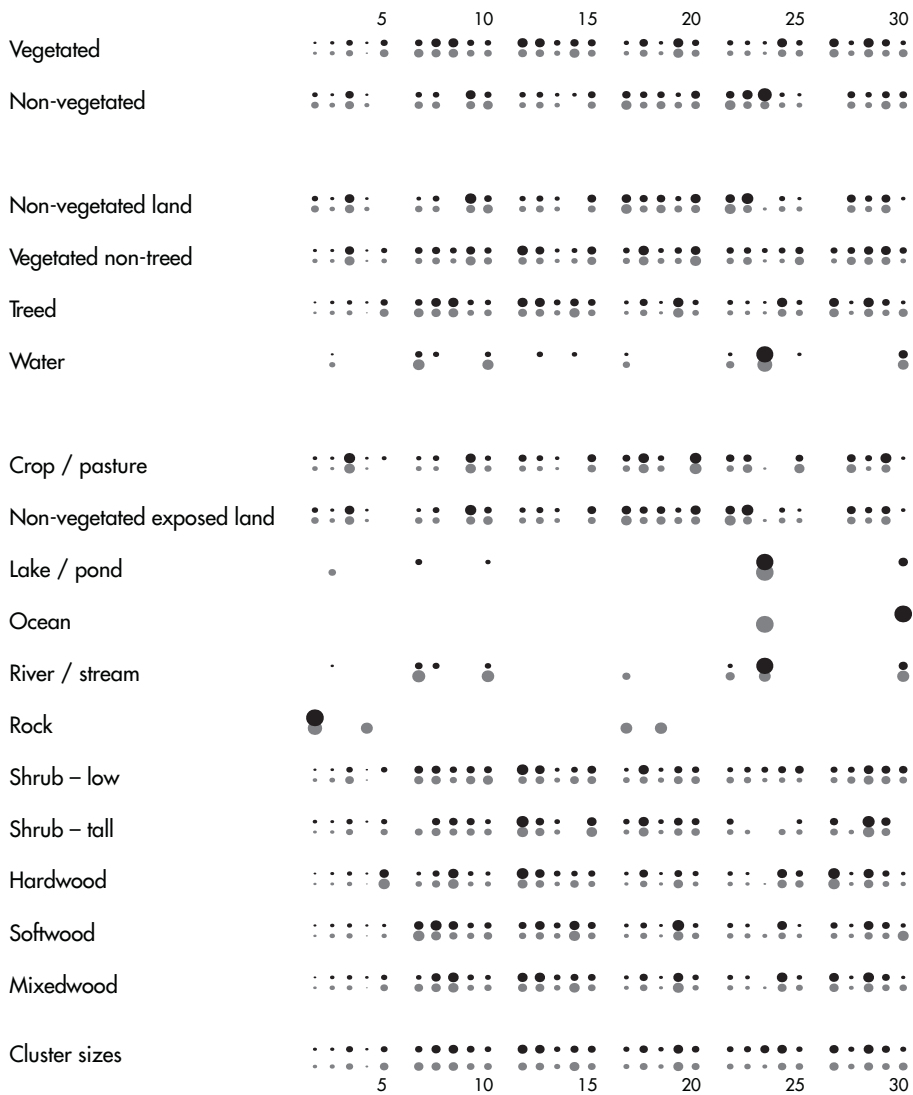
363

Fig. 6. Cluster representation of inventory cover types. Rows: Inventory cover types. Columns: Clusters 1,..., 30. Each inventory cover type is represented by a maximum of 30 clusters. The relative cluster composition of a cover type is illustrated by a series of 30 dots along each row. The size of a dot representing one cluster is proportional to the fraction of pixels within a given inventory cover type that were assigned by unsupervised classification to the cluster. Gray dots: TEST area. Black dots: VALID area. Relative overall cluster sizes for the test and validation area are displayed in the last row.

lake/pond, ocean and rock differed significantly at the photo plot level.

Inflation of discrepancies between model predictions and inventory estimates when going from a test to a validation area is also attributable to statistical problems with the models. For the neu-

ral-net-based predictions, overfitting and collinearity of cluster proportions were deemed to be the most significant factors. For example, the mean lack-of-fit obtained within the TEST1 area was increased by about 35 % when predictions were generated for the other half of the test area

(TEST2). Clearly, making predictions for areas outside the spatial domain of the model development results in a significant deterioration of performance over and above what can be deduced from testing the model inside the same area that generated the data for model development.

Neural net and maximum likelihood models were preferred for all but two cover types based on their expected performance in the validation area. While the very simple prorating scheme (PRO) only qualified for the best low shrub predictions in the testing phase it would, in hindsight, have made an appealing alternative due to simplicity and potential for fitting this procedure into a multi-phase prediction system. Prorating did significantly better in predicting the hardwood:mixedwood:softwood split and crop/pasture. Conversely it performed no better than the spatial extension when it came to vegetated/non-vegetated, and water. Mean predictions of a prorating scheme would have been, on average, about half a percentage point closer to the inventory estimates than predictions obtained with a mixture of neural nets and maximum likelihood classifiers.

4 Discussion and Conclusions

The demonstrated method for using Landsat TM images to obtain estimates of forest cover type proportions for sampling units in a national forest inventory is both fast, simple, and low-cost. Although the use of a single datum Landsat TM image naturally limits the number of distinguishable cover types well below the numbers possible with a photointerpretation of cover-types or multi-date images (Aldrich 1979, Drieman 1993, Lachowski and Bowlin 1988, Lillesand 1996, Wang et al. 1998, Wolter et al. 1995) the attraction remains for areas where no reliable cover type information exist. Enriching the predictions by provision of expected associations with more detailed inventory information remains an option (He et al. 1998).

In areas where the rate of net change in broadly defined cover type proportions is deemed low (0.5 % · yr^{-1}) a cover type mapping completed within the last 20 years may still give estimates of the current cover type distribution as good as

possible with the methods tested in this study. Although roughly 45 % of all pixels changed their level III cover type class between 1983 and 1994, the median net change in a cover type proportion of a 2×2 km photo plot was close to 0.2 %. Unless prediction differences can be reduced by an order of magnitude an update of inventories less than about 20 years old by use of current Landsat imagery appears, in situations similar to that of the case study, problematic (Sachs et al. 1998, Smiatek 1995, Van Deusen 1994), at least when the benchmark is a forest cover inventory based on interpreted aerial photos. Only a combined analysis of misclassification rates in the Landsat-based models and the forest inventory (Czaplewski 1999) can resolve the absolute error rates of our models. However, this is beyond the scope of this study. For an evaluation of how close we can predict estimates that would otherwise enter the national forest inventory our analyses remain valid.

Expedient use of existing forest inventory cover type information to develop allocation and classification rules for image data guarantees, everything else being equal, suboptimal results (Cohen et al. 1995, Dobbertin and Biging 1996, Franklin et al. 1997, Wulder et al. 1996). Cover-type polygons delineated and classified by photo-interpretation into a single type usually contains a considerable amount of variation, both within and among the polygons of the class. A closer inspection of, say, a mixed-wood polygon would show a mosaic of pure hardwood, pure softwood pixels interspersed with bona-fide mixes. As well, photo-interpretation is not without errors. Agreement between ground and photo-based interpretation of cover-types can diverge considerably (Biging et al. 1995, Eid and Næsset 1998, Hall and Fent 1996, Holmgren et al. 1997, Holopainen and Wang 1998, Magnussen 1997ab). Translation of provincial polygon classes into a set of national classes, possibly compounds the issue further. The price for this expediency is an increased frequency of disagreement and possibly biases. The significant deterioration of model performance experienced during the validation process confirmed that identically labeled polygons may indeed mask substantial differences in spectral reflectance. Subtle shifts in rock-bed, landscape features, agri-

cultural practices, and species composition between the test area and the extension area could trigger the effect (Lillesand 1996, Slaymaker et al. 1996).

Differences between Landsat-based model predictions and inventory estimates of cover type proportions reported in this study appear to be only slightly higher than results obtained with a two-phase sampling and verified training pixels (Bauer et al. 1994). Direct comparisons have been hampered by a scarcity of published records on the performance of Landsat TM-based predictions in regions outside the area used to develop the classifier. Studies by Efron (1986) indicate that an appreciable deterioration is common. Spatial heterogeneity in the performance of a classifier also increases the chance of mistakes when a classifier is used outside the domain from which it was derived (Steele et al. 1998). Our reported differences compare favorably with results obtained with AVHRR (Moody and Woodcock 1996, Turner et al. 1993).

The modeling approach that we adopted was inspired by recent work in linear 'de-mixing' models (Foody et al. 1996, Moody and Woodcock 1996, Turner et al. 1989) where progress towards removing resolution bias has been steady. For large spatial units, such as the 2×2 km photo plots, it appears that the mixing patterns of landscape elements within a fairly homogenous region are stable enough to make useful predictions of cover type proportions. Neural nets appear particularly apt at finding patterns that can be generalized (Carpenter et al. 1997, Lippmann 1987, Moody et al. 1996). The best number of image clusters to form remains contentious (Hartigan 1975, Milligan and Cooper 1985). Our approach to let the number be guided by a combination of three indicators appears reasonable. Preliminary analyses with 13 and 18 clusters gave poorer results. We conjecture from our experience and from that of others (Beaubien 1994, Benjamin et al. 1996) that it is a better strategy to include many rather than few clusters. Furthermore, the employed models lend themselves well to a multistage inventory design (Köhl and Kushwaha 1994).

Enhancement of remote sensing classification by inclusion of auxiliary enduring landscape features has been exploited with positive results (Binaghi et al. 1997, Frigessi and Stander 1994, Sader et al. 1995). Notably, the Finnish multiresource inventory makes extensive use of such techniques (Tomppo et al. 1999). Adding soil-type (43 types), elevation class (0–25 m, 26–50 m, ... , 426–450 m), and eco-site (26 classes, "Ecological land classification for New Brunswick: Ecoregion, Ecodistrict and Ecosite levels", New Brunswick Department of Natural Resources and Energy, Forest Management Branch, 1998) to the 30 Landsat image clusters as auxiliary predictors of cover-type proportions, however, did little to improve our predictions; the main reasons for this were the coarse resolution of soils maps (~300 m), elevation models (100 m) and the absence of ecotones at the photo plot level (Cox et al. 1997). Apart from a few significant gratuitous correlation stating the obvious, the ratios of posterior likelihood of class membership actually declined in most cases when these auxiliary variables were added (see also McLachan (1991) for an exposé on the 'curse of dimensionality').

Attempts to improve predictions for single photo plots in the validation area by means of a composite average based on a set of k "most similar" photo plots in the test did not succeed despite its intuitive appeal (Moeur et al. 1995, Stroup and Mulitze 1991). Similarities were indexed by Chi-square statistics of differences in pc-score cluster proportions. A recent testing of the Finnish multi-source k-nearest-neighbor methodology in Norway points in the same direction (Gjertsen et al. 1999).

An application of the presented methodology to the large fraction of Canada with dated (>20 yr.) inventory information (Gray and Power 1997) has immediate appeal. Recent inventories in nearby regions provide for a fast and low-cost updating of key resource cover types. Parallel arguments apply to the dispersed small scale (<20 000 ha) gaps in regional inventories. Success will, to a large degree, depend on the quality and availability of suitable forest inventories. Remote and unmanaged areas may lack current inventories within a radius deemed safe for model-based extrapolation.

Acknowledgement

References

Ahern, F., McKirdy, I. & Brown, 1996. Boreal forest information content of multiseason, multi-polarization c-band SAR data. Canadian Journal of Remote Sensing 22: 456–472.

Aldrich, R.C. 1979. Remote sensing of wildland resources: A state-of-the-art review. USDA Forest Service. General Technical Report RM-71: 1–56.

Andersen, F. 1988. Regional economic analysis. A Canadian perspective. Hartcourt Brace and Javanovich, Toronto. 228 p.

Apps, M., Prentice, D.T. & Wisniewski, J. 1995. Boreal forests and global change. Kluwer, Dordrecht. 548 p.

Bauer, M.E., Burk,T.E., Ek, A.R., Lime, S.D., Walsh, T.A., Walters, D.K., Befort, W. & Heinzen, D.F. 1994. Satellite inventory of Minnesota forest resources. Photogrammetric Engineering and Remote Sensing 60(3): 287–298.

Beaubien, J. 1994. Landsat TM satellite images of forests: from enhancement to classification. Canadian Journal of Remote Sensing 20: 17–26.

Benjamin, S., White, M., Argiro, D. & Lowell, K. 1996. Land cover mapping with spectrum. In: Gap analysis. A landscape approach to biodiversity planning. p. 279–288. American Society of Photogrammetry and Remote Sensing.

Biging, G.S., Dobbertin, M. & Murphy, E.C. 1995. A test of airborne multispectral videography for assessing the accuracy of wildlife habitat maps. Canadian Journal of Remote Sensing 21: 357–366.

Binaghi, E., Madella, P., Montesano, M.G. & Rampini, A. 1997. Fuzzy contextual classification of multisource remote sensing images. IEEE Geoscience and Remote Sensing 35: 326–339.

Bishop,Y.M., Fienberg, S.E. & Holland, P.W. 1975. Discrete multivariate analysis: theory and practice, MIT Press, Cambridge, Massachusetts.

Breiman, L. & Friedman, H. 1985. Estimating optimal transformations for multiple regression and correlation. Journal American Statistical Association 80(391): 580–619.

Carpenter, G.A., Gjaja, M.N., Gopal, S. & Woodcock, C.E. 1997. ART neural networks for remote sensing: vegetation classification from Landsat TM and terrain data. IEEE Geoscience and Remote Sensing 35: 308–325.

— , Busch, W.-D.N., Cleland, D.T., Gallegos, J, Harris, R., Holm, R., Topik, C. & Williamson, A. 1999. The use of ecological classification in management. In: Szaro, R.C., Johnson, N.C., Sexton, W.T. & Malk, A. (eds.), Ecological stewardship. Elsevier, Amsterdam p. 395–430.

Cochran, W.G. 1977. Sampling techniques. John Wiley. New York. 380 p.

Cohen, W.B., Spies, T.A. & Fiorella, M. 1995. Estimating the age and structure of forests in a multiownership landscape of western Oregon, U.S.A. International Journal of Remote Sensing 16: 721–746.

Congalton, R.G. & Biging, C.S. 1992. A pilot study evaluation ground reference data collection efforts for use in forest inventory. Photogrammetry Engineering and Remote Sensing 58: 1669–1671.

Costanza, R., Norton, B.G. & Haskell, B.D. 1992. Ecosystem Health. New goals for environmental management. Island Press, Washington DC.

Cox, D.D., Cox, L.H. & Ensor, K.B. 1997. Spatial sampling and the environment: some issues and directions. Environmental and Ecological Statistics 4: 219–233.

Czaplewski, R.L. 1999. Accuracy assessments and areal estimates using two-phase stratified random sampling, cluster plots, and the multivariate composite estimator. In: Mowrer, H.T. & Congalton, R.G. (eds.) Quantifying spatial uncertainty in natural resources: Theory and applications for GIS and remote sensing. Ann Arbor Press, Chelsea, MI. p. 79–100.

— & Catts, C.P. 1992. Calibration of remotely sensed proportion or area estimates for misclassification error. Remote Sensing and Environment 39: 29–43.

D'Agostino, R.B., Balanger, A. & D'Agostino, R.B. Jr. 1990. A suggestion for using powerful and informative tests of normality. American Statistician 44(4): 316–321.

Dale, A. & Robinson, B. 1996. Achieving sustainable development. Univ. British Columbia, Vancouver.

Dobbertin, M. & Biging, G.S. 1996. A simulation study of the effect of scene autocorrelation, training sample size and sampling method on classification accuracy. Canadian Journal of Remote Sensing 22(4): 360–367.

Drieman, A. 1993. The forest resource of Labrador: Landsat thematic mapper imagery provides a unique and current perspective. Forestry Chronicle 69: 667–671.

Efron, B. 1986. How biased is the apparent error rate of a prediction rule. Journal American Statistical Association 81(394): 461–470.

Eid, R. & Næsset, E. 1998. Determination of stand volume in practical forest inventories based on field measurements and photo-interpretation: the Norwegian experience. Scandinavian Journal of Forest Research 13: 246–254.

Foody, G.M. 1999. The continuum of classification fuzziness in thematic mapping. Photogrammetry Engineering and Remote Sensing 65(4): 443–451.

— , Lucas, R.M., Curran, P. & Honzak, M. 1996. Estimation of the areal extent of land cover classes that only occur at a sub-pixel level. Canadian Journal of Botany 22: 428–432.

Franklin, S.E., Lavigne, M.B., Deuling, M., Wulder, M.A. & Hunt, E.R. Jr. 1997. Landsat TM derived forest cover-types for modeling net primary production. Canadian Journal of Remote Sensing 23: 243–252.

Frigessi, A. & Stander, J. 1994. Informative priors for the Bayesian classification of satellite images. Journal American Statistician Association 89: 703–709.

Gjertsen, A.K., Tomppo, E. & Tomter, S. 1999. National forest inventory in Norway: Using sample plots, digital maps, and satellite images. In: IGARSS'99 IEEE Conference Proceedings Vol. 1, ISBN 0-7803-5208-4, Hamburg, GE June 28–July 3. p. 729–731.

Gray, S. & Power, K.1997. Canada's forest inventory 1991: the 1994 version – technical supplement. Canadian Forestry Service, Pacific Forestry Center. Information Report BC-X-363: 1–73.

Gregoire, T. 1998. Design-based and model-based inference in survey sampling: Appreciating the difference. Canadian Journal of Forest Research 28: 1429–1447.

Hall, R. & Fent, L. 1996. Influence of aerial film spectral sensitivity and texture on interpreting images of forest species composition. Canadian Journal of Remote Sensing 22: 350–356.

Hartigan, A. 1975. Clustering algorithms. John Wiley, New York.

He, H.S., Mladenoff, D., Radeloff, V.C. & Crow, T.R. 1998. Integration of GIS data and classified satellite imagery for regional forest assessment. Ecological Applications 8(4): 1072–1083.

Holmgren, P., Thuresson, T. & Holm, S. 1997. Estimating forest characteristics in scanned aerial photographs with respect to requirements for economic forest management planning. Scandinavian Journal of Forest Research 12: 189–199.

Holopainen, M. & Wang, G. 1998. Accuracy of digitized aerial photographs for assessing forest habitats at plot level. Scandinavian Journal of Forest Research 13: 499–508.

Köhl, M. & Kushwaha, S.S. 1994. A four-phase sampling method for assessing standing volume using landsat-TM-data aerial photography and field assessments. Commonwealth Forest Review 73: 35–41.

Lachowski, H.M. & Bowlin, H.L. 1988. Forest inventory and assessment with satellite imagery in the western states. Engineering Field Notes 20: 43–50.

Leblanc, M. & Tibshirani, R. 1996. Combining estimates in regression and classification. Journal American Statistician Association 91: 1641–1650.

Leckie, D.G. & Gillis, M.D. 1995. Forest inventory in Canada with emphasis on map production. Forestry Chronicle 71(1): 74–78.

Lek, S., Delacoste, M., Baran, P., Dimopoulos, I., Lauga, J. & Aulagnier, S. 1996. Application of neural networks to modelling nonlinear relationships in ecology. Ecological Modelling 90: 39–52.

Lillesand, T.M. 1996. A protocol for satellite-based land cover classification.In: Scott, M., Tear, T.H. & Davis, F.W. (ed.), Gap analysis: A landscape approach to biodiversity planning. American Society Photogrammetry Remote Sensing. Bethesda, MD. p. 103–118.

Lippmann, R.P. 1987. An introduction to computing with neural nets. IEEE Transactions on Neural Networks 5–23.

Little, R.A. & Rubin, D.B. 1987. Statistical analysis with missing data. John Wiley and Sons, New York.

Lowe, J.J., Power, K. & Gray, S. 1994. Canada's forest inventory 1991: The 1994 version. Natural Resources Canada, Canadian Forestry Service Information Report BC-X-362.

Magnussen, S. 1997a. A method for enhancing tree species proportions from aerial photos. Forestry Chronicle 73: 479–487.

— 1997b. Calibrating photo-interpreted forest cover types and relative species compositions to their ground expectations. Canadian Journal of Forest Research 27: 491–500.

McLachlan, G. 1991. Discriminant analysis and statistical pattern recognition. Wiley, New York. 517 p.

Metternicht, G.I. & Fermont, A. 1998. Estimating surface features by linear mixture modeling. Remote Sensing of Environment 64: 254–265.

Miller, G. 1980. Simultaneous statistical inference. 2nd Ed. Springer, New York.

Milligan, G.W. & Cooper, M.C. 1985. An examination of procedures for determining the number of clusters in a data set. Psychometrika 50: 159–179.

Moeur, M., Crookston, N.L. & Stage, A.R. 1995. Most similar neighbor: an improved sampling inference procedure for natural resource planning. Forest Science 41: 337–359.

Moody, A. 1998. Using landscape spatial relationships to improve estimates of land-cover area from coarse resolution remote sensing. Remote Sensing of Environment 64: 202–220.

— , Gopal, S. & Strahler, A.H. 1996. Artificial neural network response to mixed pixels in coarse-resolution satellite data. Remote Sensing of Environment 58: 329–343.

— & Woodcock, C.E. 1996. Calibration-based models for correction of area estimates derived from coarse resolution land-cover data. Remote Sensing of Environment 58: 225–241.

Noss, R.F., O'Connell, M.A. & Murphy, D.D. 1997. The science of conservation planning: Habitat conservation under the endangered species act. Island Press, Washington DC. 246 p.

Odum, H.T. 1996. Environmental accounting – EMERGY and environmental decision making. John Wiley, New York. 370 p.

Patil, G.P. 1982. Diversity as a concept and its measurement. Journal American Statistical Association 77: 548–561.

Peña, D. 1997. Combining information in statistical modeling. American Statistician 51: 326–332.

Rencher, A.C. 1993. The contribution of individual variables to Hotelling's T^2, Wilks' Lambda, and R^2. Biometrics 49: 479–489.

Rouse, W. Jr., Haas, R.H., Schell, A. & Derring, D.W. 1974. Monitoring vegetation systems in the Great plains with ERTS. In: Earth Resources Technology Satellite-1 Symposium. Goddard Space Flight Center, Wash. D.C. p. 309–317.

Sachs, D.L., Sollins, P. & Cohen, W.B. 1998. Detecting landscape changes in the interior of British Columbia from 1975 to 1992 using satellite imagery. Canadian Journal of Forest Research 28: 23–36.

Sader, S.A., Ahl, D. & Liou, W.S. 1995. Accuracy of Landsat-TM and GIS rule-based methods for forest wetland classification in Maine. Remote Sensing of Environment 53: 133–144.

Scott, M., Tear, T.H. & Davis, F.W. 1996. GAP analysis: A landscape approach to biodiversity planning. American Society of Photogrammetry and Remote Sensing, Bethesda, MD.

Slaymaker, D.M., Jones, K.M.L., Griffin, C.R. & Finn, T. 1996. Mapping deciduous forests in southern New England using aerial videography and hyper-clustered multi-temporal Landsat TM imagery. In: Scott, M., Tear, T.H. & Davis, F.W. (eds.), Gap analysis: A landscape approach to biodiversity planning. American Society of Photogrammetry and Remote Sensing. Bethesda, MD. p. 87–101.

Smiatek, G. 1995. Sampling thematic mapper imagery for land use data. Remote Sensing of Environment 52: 116–121.

Steele, B.M., Winne, C. & Redmond, R.L. 1998. Estimation and mapping of misclassification probabilities for thematic land cover maps. Remote Sensing of Environment 66: 192–202.

Stroup, W.W. & Mulitze, D.K. 1991. Nearest neighbor adjusted best linear unbiased prediction. American Statistician 45: 195–200.

Tauer, L.W. 1983. Target MOTAD (minimization of total absolute deviations). American Journal of Agricultural Economy 65: 606–610.

Tomppo, E., Goulding, C. & Katila, M. 1999. Adapting Finnish multi-source forest inventory techniques to the New Zealand preharvest inventory. Scandinavian Journal of Forest Research 14(2): 182–192.

Turner, D.P., Koerper, G., Gucinski, H. & Peterson, C. 1993. Monitoring global change: comparison of forest cover estimates using remote sensing and inventory approaches. Environment Monitoring and Assessment 26: 295–305.

Turner, M., O'Neill, R., Gardner, R.H. & Milne, R. 1989. Effects of changing spatial scale on the analysis of landscape pattern. Landscape Ecology 3(3–4): 153–162.

Van Deusen, P. 1994. Correcting bias in change estimation from thematic maps. Remote Sensing of Environment 50: 67–73.

Venables, W.N. & Ripley, B.D. 1994. Modern applied statistics with S-Plus, Springer-Verlag, New York.

Walsh, T.A. & Burk, T.E. 1993. Calibration of satellite classifications of land area. Remote Sensing of Environment 46: 281–290.

Wang, G., Poso, S., Waite, M.-L. & Holopainen, M. 1998. The use of digitized aerial photographs and local operation for classification of stand development classes. Silva Fennica 32(3): 215–225.

Warner, B. & Misra, M. 1996. Understanding neural networks as statistical tools. American Statistician 50: 284–299.

White, D., Kimerling, A. & Overton, S. 1992. Cartographic and geometric components of a global sampling design for environmental monitoring. Cartographic Geographic Information Systems 19: 5–22.

Wolter, P.T., Mladenoff, D., Host, G.E. & Crow, T.R. 1995. Improved forest classification in the northern lake states using multi-temporal Landsat imagery. Photogrammetry Engineering and Remote Sensing 61(9): 1129–1143.

Wulder, M.A. 1998. The Prediction of leaf area index from forest polygons decomposed through the integration of remote sensing, GIS, UNIX, and C. Computing and Geoscience 24(151): 157.

— , Franklin, S.E. & Lavigne, M.B. 1996. High spatial resolution optical image texture for improved estimation of forest stand leaf area index. Canadian Journal of Remote Sensing 22: 441–449.

Total of 83 references