# Land cover of the BOREAS Region from AVHRR and Landsat data

by J. Cihlar • J. Beaubien • Q. Xiao • J. Chen • Z. Li


## RÉSUMÉ

*L'objectif de cette étude visait à caractériser la répartition des types de couvert dans la région BOREAS. Des données multitemporelles AVHRR ont été acquises tout au long de la saison de végétation de 1993 et traitées pour en déduire les moyennes saisonnières de pixels de 1 km² dans les canaux 1 et 2, l'indice NDVI (Normalized Difference Vegetation Index) et la superficie sous la courbe NDVI. Deux procédures différentes de classification AVHRR ont été utilisées pour identifier >30 types de couvert sur un territoire de 1.44x10⁶ km² couvrant la région BOREAS, et la légende de classification hiérarchique utilisée était compatible avec le programme IGBP (International Geosphere-Biosphere Program). La précision de la classification AVHRR a été évaluée qualitativement et quantitativement à l'aide d'une comparaison avec des images Thematic Mapper de Landsat (TM). La composition des types de couvert à l'intérieur des pixels AVHRR a été quantifiée en utilisant les classifications des deux zones d'étude BOREAS dérivées des images TM. On a pu observer que les deux classifications AVHRR donnent des estimations très similaires (entre 0.5%-1.4% de la superficie totale) des proportions de superficies au plan des classes individuelles, la précision dépendant quelque peu du niveau à l'intérieur de la hiérarchie de classification. En accord avec une étude précédente, la précision absolue des classifications AVHRR était relativement faible quand tous les pixels étaient considérés, mais élevée (>80%) quand seulement les pixels AVHRR contenant principalement un type de couvert étaient considérés. Une combinaison de données AVHRR et TM a pu être utilisée pour quantifier l'effet de couverts mixtes à l'intérieur d'un pixel. À l'intérieur de la région de 1.44x10⁶ km², 32% de la superficie a été identifiée comme forêt conifériennne, 11% comme forêt mixte et 10% était de l'eau; les autres classes occupaient individuellement <5% de la superficie. L'étude a démontré que les deux méthodologies de classification AVHRR donnaient des résultats consistants et que l'utilisation de différentes méthodes peut constituer une stratégie efficace pour accroître la consistance et la robustesse des estimations de superficies. Des améliorations à apporter ont été identifiées dans les données et méthodes requises pour optimiser les estimations de superficies des classes individuelles aux niveaux régional et national.*

## SUMMARY

*The objective of this study was to characterize the distribution of land cover types in the BOREAS Region. Multitemporal AVHRR data were obtained during the entire 1993 growing season and processed to yield seasonal means for 1 km² pixels in channels 1 and 2, the Normalized Difference Vegetation Index (NDVI), and the area under the NDVI curve. Two different AVHRR classification procedures were employed to identify >30 cover types in a 1.44x10⁶ km² area encompassing the BOREAS Region, using an IGBP-compatible hierarchical classification legend. The accuracy of the AVHRR classification was evaluated qualitatively and quantitatively through a comparison with Landsat Thematic Mapper (TM) images. The composition of land cover types within the AVHRR pixels was quantified using classifications of the two BOREAS study areas derived from TM images. It was found that the two AVHRR classifications provide closely similar estimates (within 0.5%-1.4% of the total area) of area proportions for the individual classes, the accuracy depending somewhat on the level within the classification hierarchy. Consistently with a previous study, the absolute AVHRR classification accuracy was fairly low when all pixels were considered, but was high (>80%) when only AVHRR pixels containing mostly one cover type were considered. A combination of AVHRR and TM data could be used to quantify the effect of mixed cover types within a pixel. Within the 1.44x10⁶ km² region, 32% was identified as coniferous forest, 11% as mixed forest, and 10% was water; other classes occupied <5% each. The study indicated that the two AVHRR classification methodologies provided consistent results and that using different methods may be an effective strategy for increasing the consistency and robustness of area estimates. Improvements were identified in data and methods which are needed to optimize area estimates of the individual classes at the regional and national levels.*


• J. Cihlar, J. Chen and Z. Li are with the Canada Centre for Remote Sensing, Ottawa, Ontario, K1A 0Y7
• J. Beaubien is with the Canadian Forest Service, Quebec City, Quebec
• Q. Xiao is with Intermap Technologies, Ottawa, Ontario K2E 1A2

## Introduction and Objectives

Land cover characteristics influence many of the mass and energy exchange processes at the land-atmosphere interface. They are therefore of strong interest in observation and modeling studies dealing with the ecosystem-atmosphere interactions, such as the BOReal Ecosystem-Atmosphere Study (BOREAS; Sellers *et al.*, 1995). Ideally, these studies require information on the spatial distribution of specific parameters such as plant biomass or canopy conductance. Since in practice many of these variables cannot yet be directly estimated over large areas, land cover type can be employed as a useful surrogate because the required parameters can be estimated for each cover type from surface or near-surface observations.

Land cover in the BOREAS Region has been mapped with Landsat data for the two study areas (approximately 7000 $km^2$ each (Hall and Knapp, 1994a,b), and with 1992 AVHRR data for most of the BOREAS Region (Steyaert and Loveland, 1995). The purpose of this paper is to characterize land cover composition in the BOREAS Region and adjacent areas ($1.44x10^6$ $km^2$) using 1993 AVHRR data, in combination with higher resolution classifications derived from Landsat Thematic Mapper (TM) data. This paper extends previous work by encompassing a larger area, using more recent data, combining medium and high resolution data to assess within pixel heterogeneity, and using a classification legend which is compatible with that used internationally but also provides a link between the BOREAS Region and land cover composition across Canada.

## Methodology

The basic methodological approach consisted of: collection and processing of AVHRR data over Canada during the 1993 growing season; extraction of land cover type information from the AVHRR data using two methods; quantification of the composition of AVHRR pixels using land classifications derived from higher resolution satellite data; and the combination of the partial results to derive land cover estimates for the BOREAS Region. Although the focus of this paper is the BOREAS Region the AVHRR processing and classifications were carried out for all Canada and are discussed elsewhere (Cihlar *et al.*, 1997a).

### AVHRR processing

Multitemporal AVHRR data have been successfully employed to provide cover type information on the biosphere (Loveland *et al.*, 1991; Brown *et al.*, 1993; Eidenshink and Faundeen, 1994). AVHRR contains bands in the red (channel 1, C1), near infrared (C2), mid-infrared (channel 3), and two in the thermal region (C4, C5) of the electromagnetic spectrum. In the past, the Normalized Difference Vegetation Index (NDVI, computed as [C1 minus C2] divided by [C1 plus C2]) has been most frequently used, among others because it is derived in a way that reduces the noise in the measurements. However, we wished to use individual channels (C1, C2) because our previous studies (Beaubien and Simard, 1993; Cihlar *et al.*, 1996) showed that these channels contain useful information on northern land cover types. The AVHRR processing thus consisted of two steps: (i) preprocessing of daily images into 10-day composites, and (ii) a series of postprocessing steps which yielded the data set for land cover analysis.

### Preprocessing and compositing

Daily afternoon NOAA-11 AVHRR images were obtained at receiving stations in Prince Albert, Saskatchewan and Dartmouth, Nova Scotia between 11 April and 31 October, 1993. Using the GEOCOMP system (Robertson *et al.*, 1992) the data were calibrated by incorporating post-launch degradation of C1 and C2 (Cihlar and Teillet, 1995), corrected for geometric distortions, and resampled to 1 km pixel size in a Lambert Conformal Conic Projection (LCC; standard parallels 49° N and 77° N, reference meridian 95° W). A nominally cloud-free composite image was reconstituted every 10 days (11 for the last period of May, July) by retaining the most cloud-free pixel for each location; the maximum NDVI value was used to make the choice. Each of the 20 composites consisted of 10 channels of data (5 reconstituted AVHRR channels for top-of-the-atmosphere (TOA) radiance, NDVI, 3 viewing geometry channels, date of acquisition). More details on the preprocessing are provided by Cihlar *et al.* (1997b,c).

### Post-processing

The ABC3 procedure (Atmospheric, Bidirectional and Cloud Contamination Corrections of CCRS; Cihlar *et al.*, 1997b) was applied to transform the TOA radiances in the composites to surface reflectance/emission with a standard viewing geometry. The steps included atmospheric corrections for C1 and C2; detection of pixels containing snow, subpixel clouds, smoke or other significant atmospheric contamination in all channels using the CECANT algorithm (Cihlar, 1996); bidirectional corrections for C1 and C2 (Wu *et al.*, 1995; Li *et al.*, 1996); NDVI corrections for only solar zenith angle (SZA) effects using the FASIR procedure (Sellers *et al.*, 1994); corrections of C4 for atmospheric and surface emissivity effects; and replacement of contaminated pixels in all channels through temporal interpolation. These steps, applied to all 20 composites, yielded normalized surface reflectance (SZA=45°, nadir) for C1 and C2; NDVI; and surface temperature for each pixel/compositing period. It should be noted that the NDVI was corrected for SZA effects only, while C1 and C2 were fully corrected; this conservative approach was used because of the early stages of the development of the ABC3.

Once the seasonal composite series was available, the growing season was identified for each pixel as that period during which the ABC3 surface temperature during satellite overpass was above +10°C, linearly interpolating for days between the centres of the 10-day periods (Cihlar *et al.*, 1997b). Mean C1, C2, and NDVI ($N_m$) values for the growing season were computed for each pixel. In addition, the area under the NDVI curve ($N_a$) during the growing season was also calculated. The data were then transformed from the original sensor precision of 10 bits (all above processing was done in 16 bits) to 8 bits by truncating the tails of the histograms. These four data channels formed the basis for land cover analysis.

## Land cover classification

Two classification methodologies were employed to classify land cover on the AVHRR data. The first procedure, the Enhancement-Classification Method (ECM), is based on extensive previous work with TM images and uses analyst's knowledge to optimize the classification process. The second methodology, Classification by Progressive Generalization (CPG; Cihlar *et al.*, 1997d) consists of a series of operations intended to identify the actual spectral content of the data set while setting a minimum of restrictions and automating most of the steps. Two methods have been used because of the differences in the input data and to provide a partial check on the consistency of the area estimates. Both procedures were applied to all Canada but only the BOREAS results are discussed here. The two procedures are briefly described in the following sections.

The IGBP classification legend (Belward, 1996) was adopted as the basic scheme for the study. Consisting of 17 possible cover types, this legend was developed specifically for use at continental to global scales and with AVHRR and similar data. The ECM and CPG methodologies resulted in the same IGBP classes (termed IGBP level L 0 in **Table 1**). These classes were further subdivided in 1 or 2 levels to accommodate the more detailed differentiation of cover types we felt was possible with the data and the classification methods employed. There were some differences between ECM and CPG at the level L -2 (**Table 1**), in part due to the classification approach used (e.g., urban areas) and in part to reflect the different information content in the data (four compared to three input channels, e.g. see cropland).

Because of the AVHRR processing methodology employed, open water could not be mapped with the AVHRR data (the compositing tends to select cloudy pixels over water). A water mask (U.S. Department of Commerce, 1977) was therefore used to mask out major water bodies. However, small (including subpixel) water bodies could not be masked and were thus classified by both methods as one of the land cover types.

### *Enhancement-Classification Method (ECM)*

The ECM methodology has been developed on the basis of previous studies involving Landsat Thematic Mapper classification (Beaubien, 1994; Beaubien *et al.*, 1997). The procedure is intended to capture most of the information visible in an enhanced image but converted into a classification. It has been designed to be reproducible as much as possible for an interactive classification. In summary, it consists of the following steps:

(i) contrast enhancement in three channels (C1, C2, $N_m$), using values for selected cover types as the histogram extremes;
(ii) quantizing each stretched channel into 10 equal gray levels between the minimum and the maximum values;
(iii) filtering each quantized channel using a 5x5 median filter;
(iv) identifying and locating significant spectral clusters/colours (pixels representing distinguishable cover types or conditions and present in significant amounts) on the filtered image, and compiling their digital values; 92 such colours were found on the 1993 images of Canada;
(v) using a minimum Euclidean distance decision rule, the full resolution contrast-stretched image (from (i)) is classified according to one of the clusters in (iv);
(vi)grouping the clusters from (iv) into classes specified by the classification legend based on ground-truth knowledge (mainly TM colour composites). To facilitate the labeling process, the pseudo-colour table applied to the resulting classification is derived from the image in (ii) above.

In some cases (e.g., built-up areas, cropland) there was a significant confusion with other classes, especially in the northern areas. This problem was dealt with by creating bitmaps using ancillary information, and relabeling the classes in these areas accordingly.

### *Classification by Progressive Generalization (CPG)*

A new classification methodology was developed to take advantage of the features of the above classification but ensuring reproducibility and minimizing analyst bias. This is achieved by automating the classification process until the number of remaining spectral clusters is not too much higher than the number of the desired thematic classes, i.e. low enough that it is practical to label all the clusters. CPG consists of three phases: identification of important combinations of spectral values to serve as seed clusters; clustering process which assigns each pixel to one of the seed clusters using minimum Euclidean distance as the criterion; and a progressive grouping of clusters through a combination of spectral and spatial proximity measures. The individual steps are illustrated in **Figure 1** and described in detail by Cihlar *et al* (1997d). Four channels of data were used (C1, C2, $N_m$, and $N_a$). The classification legend shown in **Table 1** was used. The resulting AVHRR spectral clusters were labeled based on their correspondence to the cover types obtained through the classification described in section on the ECM (in turn derived by visual inspection of many TM images).

### *Accuracy assessment*

Qualitative accuracy assessment of the methodology was carried out by comparing the classifications with Landsat TM images (analogue form) over many areas, with an increased density over Quebec. In general, very good correspondence with the TM images was observed, including changes in land cover since the TM acquisition date such as forest burns and regrowth in burned areas. Qualitative assessment was also made by comparing the classification results with the input data, after colouring the classes with colours derived from the original data. Quantitative accuracy assessment was performed for the two BOREAS study areas using 30m TM classifications of Hall and Knapp (1994 a,b). The diagonal classification accuracy DiAc was derived from confusion matrices between the TM classification ('truth') and the AVHRR-derived classification:

$$DiAc(i,i) = \frac{100 * \sum_{i=j}^{q} P(AVHRR)_i}{P(TM)_i} \qquad (1)$$

where: $P(TM)_i$ is the number of 30m-pixels in the AVHRR classification correctly labeled as true class i (i.e., diagonal entry); $P(AVHRR)_j$ is the number of 30m-pixels in all the AVHRR clusters labeled as TM class i (i.e., column total); and j to q are the AVHRR clusters labeled as class i.

DiAc thus measures the proportion of ground cover that was 'positively identified' as that cover through AVHRR classification, assuming that the TM classification is 100% accurate.

The overall classification accuracy for a given data set was computed as

$$DiAc = \frac{100 * \sum_{i=j}^{n} P(i,i)}{NP} , \qquad (2)$$

where P(i,i) is the diagonal entry in the confusion matrix (number of pixels), and NP is the total number of pixels in the matrix. DiAc is thus the most stringent measure, counting only positive matches as correctly classified.

The Khat distance was also computed for each confusion matrix (Congalton, 1991), as follows:

$$Khat = \frac{N \sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r} (x_{i+} * x_{+i})}{N^2 - \sum_{i=1}^{r} (x_{i+} * x_{+i})} \qquad (3)$$

where: $x_{ii}$ is the total number of pixels in row i, column i; N is the total number of pixels; r is the number of rows (columns); and $x_{i+}$ and $x_{+i}$ are row and column totals, respectively.

As seen from Eqation 3 Khat measures the dispersion outside of the diagonal in the confusion matrix in relation to the concentration along the diagonal axis. For a perfect classification, Khat=1.0.

**Table 1.**

**Land cover classes employed in the two AVHRR classification methods.**

| Land cover category* | Level* | ECM | CPG |
|---|---|---|---|
| ***A - Natural Vegetation*** | | | |
| **1. Forest Land (tree crown density >10%)** | **L 1** | | |
| **1.1 Evergreen needleleaf forest (>80% coniferous trees)** | **L 0** | | |
| <u>High crown density</u> (>60%) | L -1 | x | x |
| <u>Medium crown density</u> (40-60%) | L -1 | x | x |
| <u>Low crown density</u> (25-40%) | L -1 | | |
| Southern boreal forest | L -2 | x | x |
| Northern** boreal forest, heath-and-moss cover | L -2 | x | x |
| Northern boreal forest, lichens cover | L -2 | x | x |
| <u>Very low crown density</u> (10-25%) | L -1 | | |
| Southern boreal forest | L -2 | x | x |
| Northern boreal forest, lichens covers | L -2 | x | x |
| Northern boreal forest, heath-and-moss cover | L -2 | x | x |
| 1.2 Deciduous broadleaf forest (>80% broadleaf trees) | **L 0** | | |
| Broadleaf forest | L -1 | x | x |
| 1.3 Mixed forest | **L 0** | | |
| <u>60-80% needleleaf trees</u> | L -1 | | |
| More regular forest | L -2 | x | x |
| Younger, more irregular | L -2 | x | x |
| Low crown density | L -2 | x | x |
| <u>60-80% broadleaf trees</u> | L -1 | | |
| More regular forest | L -2 | x | x |
| Younger, more irregular | L -2 | x | x |
| <u>40-60% broadleaf or needleleaf trees</u> | L -1 | | |
| Young, irregular | L -2 | x | x |
| **2. Open Land (tree crown density <10%)** | **L 1** | | |
| **2.1 Shrubland** | **L 0** | | |
| Closed shrubland (>50% shrubs), particularly wetlands | L -1 | x | x |
| Open shrubland (25-50% shrubs) | L -1 | x | x |
| Shrubs (10-25%) and lichens (<50%) | L -1 | x | x |
| Recent burns patches, may be mixed pixels including water (other classes in burns as well) | L -2 | x | |
| **2.2 Grassland** | **L 0** | | |
| Grassland | L -1 | x | x |

continued on page 168.

***Analysis and integration***

Ideally, land cover estimates should be obtained from 100% coverage of the area of interest with high resolution data. This approach is not used over large areas because of the costs and amount of work involved. However, it may be feasible to combine 100% coverage at a coarse resolution with selected coverage of high resolution data. We employed the TM and AVHRR classifications to estimate the total area of the TM classes in the BOREAS area.

To obtain information on the land cover composition, the number of 1 km² pixels within the BOREAS Region was counted. The actual area employed is the rectangle bounding the original BOREAS Region in the LCC projection; its coordinates were NW corner at -115°23′52.8″, 59°21′50.4″, NE corner at -93°17′21.8″, 61°0′35.6″, SW corner at -110°14′ 48.1″, 48°50′22.6″ and SE corner at -93°44′28.7″, 50°1′52.7″. To estimate the proportions of cover types within individual 1 $km^2$ pixels the TM-derived cover type map was registered to the AVHRR classification, and the AVHRR pixels were resampled to 30 m TM-equivalent pixels. Within each TM-mapped area, the composition of a given AVHRR cover type could then be described in terms of the fractions of the individual TM cover types. Using the AVHRR vs. TM classes confusion matrix the total area of a given class was estimated as:

$$A(i) = \sum_{j=1}^{n} A_{AVHRR}(j)^* \left[ \frac{A_{TM}(i,j)}{\sum_{k=1}^{m} A_{TM}(k,j)} \right] \quad (4)$$

where: A(i) = total area of cover type i (in TM classes) in the BOREAS Region (covered by the AVHRR classification); $A_{AVHRR}$ (j) = total area of cover type j in the BOREAS Region (AVHRR classification); $A_{TM}$ (i,j) = area classified as cover type i (TM) and type j (AVHRR) in the region covered by both classifications; m= number of TM classes; n= number of AVHRR classes.

Equation 4 thus assumes that the proportions of classes i in the overlapping area (mean for BOREAS Southern Study Area, SSA and Northern Study Area, NSA in this paper) represent those in the entire BOREAS Region covered by the AVHRR classification.

## Results and Discussion

### Accuracy of AVHRR classifications

Two sources of error need to be distinguished, errors due to input data and those due to the classification methodology. Potential errors in input data arise due to the large amount of the initial raw data processed and the range of conditions under which these data were obtained. The residual imperfections in the input data will influence the classification results, in addition to any effects peculiar to the classification procedure itself. Although the ABC3 procedure strongly mitigates or removes many of these, the resulting data are not as sound as if they were obtained in one day under uniform conditions (Cihlar *et al.*, 1997b). ABC3 also does not compensate for occasional errors in geometric registration. However, since it is virtually impossible to identify all the residual errors in the input data (Cihlar, 1996; Cihlar *et al.*, 1997b) the data to be classified must necessarily be regarded as containing no errors, and the resulting inaccuracies dealt with in the post-classification process (e.g., through labeling using an independent information).

Qualitatively, the accuracy of the classifications can be assessed by a visual comparison with the input data. This can be readily accomplished by displaying individual classes in colours which correspond to the spectral values of the cluster means in the unclassified images. **Figure 2a** shows an example of a part of the BOREAS Region with

**Table 1. (continued)**

| | | | |
|---|---|---|---|
| ***B - Developed and Mosaic Land*** | | | |
| **3.1 Cropland (>60% cropland)** | **L 0** | | |
| Well watered or irrigated | L -2 | x | |
| Subhumid in the Plains | L -2 | x | |
| Semiarid in the Plains | L -2 | x | |
| High biomass (well-watered or irrigated) | L -2 | | x |
| Medium-High biomass | L -2 | | x |
| Medium biomass | L -2 | | x |
| Medium-Low biomass | L -2 | | x |
| Low biomass (semiarid) | L -2 | | x |
| **3.2 Cropland/Woodland Mosaics** | **L 0** | | |
| Cropland and woody areas | L -2 | x | |
| Cropland/Woodland mosaic (cropland dominant) | L -2 | | x |
| Woodland/Cropland mosaic (woodland dominant) | L -2 | | x |
| **3.3 Urban and built-up areas** | **L 0** | | |
| Built-up areas | L -1 | x | |
| Urban areas | L -1 | x | |
| *C - Non-Vegetated Land* | | | |
| **4.1 Barren** | **L 0** | | |
| Patches of stunted black spruce | L -2 | x | x |
| Lichens and shrubs | L -2 | x | x |
| Low vegetation cover (<25%) | L -2 | x | x |
| Very low vegetation cover (<10%), particularly rocks, boulders | L -1 | x | x |
| Barren | L -1 | x | x |
| **4.2 Water and other** | **L 0** | | |
| Water bodies, snow, ice, unclassified | L -1 | x | x |

Notes:
* Categories at the levels L 1 and L 0 are those defined by the IGBP (Belward, 1996); categories L -1 and L -2 were defined to fall within the IGBP classes
** As defined by Rowe (1972).

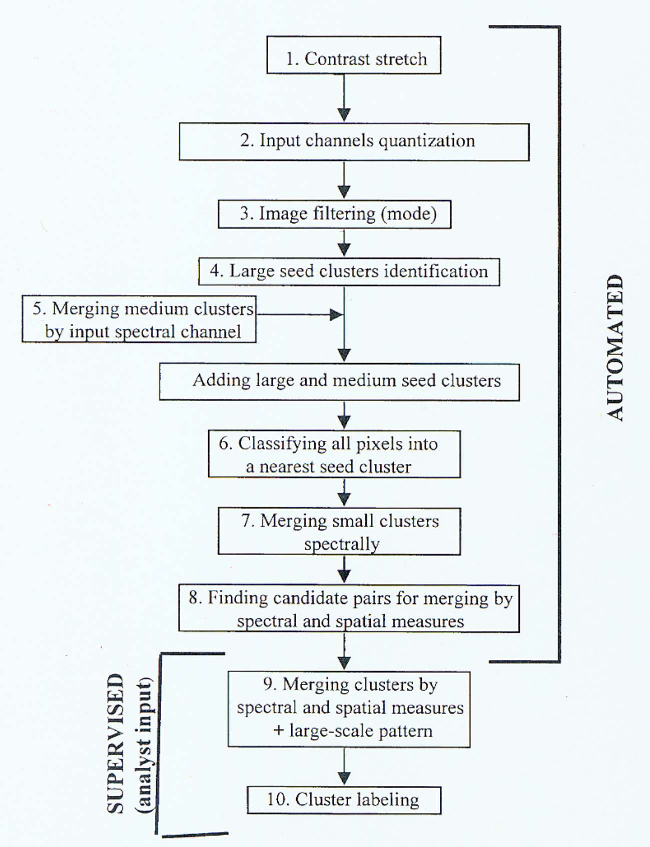


**Figure 1.**
**Flowchart for the CPG method.**

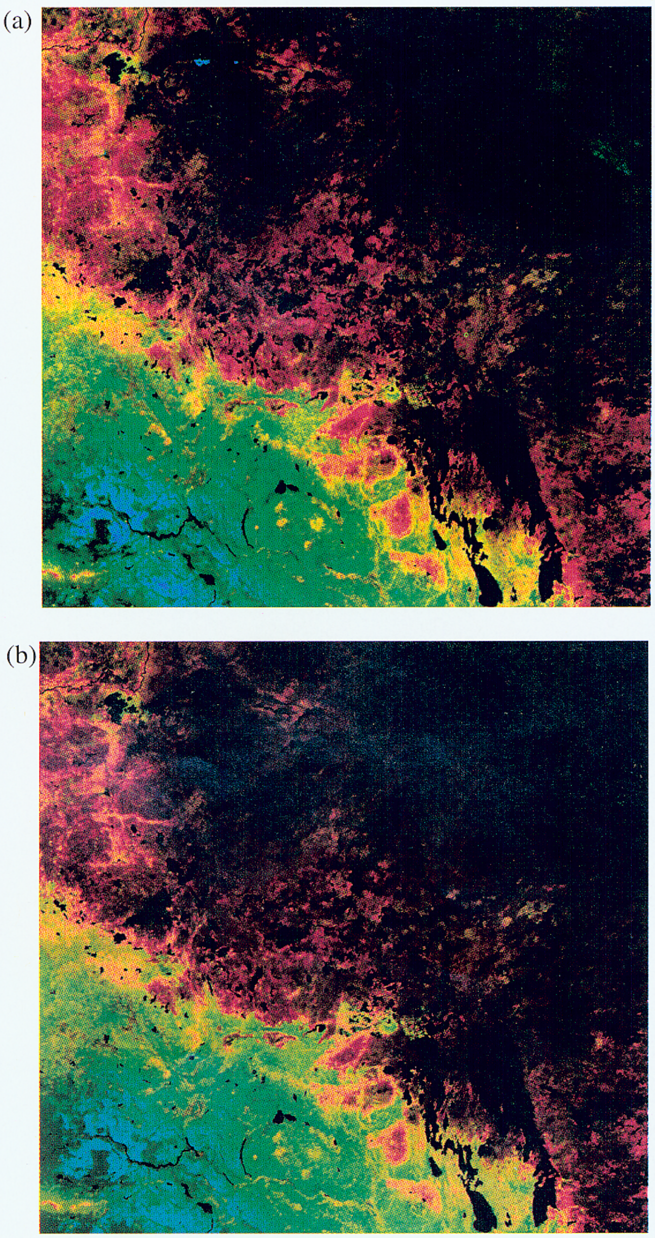


**Figure 2.**
**Qualitative accuracy assessment of the AVHRR classification of the BOREAS Region. Figure 2a: original input data (C1=blue, C2=green, Nm=red). Figure 2b: CPG classification based on four channels (C1,C2,Nm,Na); the class codes were replaced by the means of C1, C2, and Nm to facilitate visual comparison. BOREAS northern (NSA) and southern (SSA) study areas are outlined.**

three of the input channels displayed (C1, C2, $N_m$). The same area is shown classified using the CPG in **Figure 2b** where each class is displayed in the colour of its cluster mean (as it appeared in the original data). The close correspondence of the original and classified images is clearly evident, indicating that the classification retained most of the content of the input data. Very similar results were obtained in the case of the ECM.

The AVHRR accuracies were evaluated quantitatively using classifications derived from TM data for the southern (SSA) and northern (NSA) study areas (Hall and Knapp, 1994a,b; **Table 2**). **Table 3** shows the confusion matrices for both study areas and the two methods, ECM and CPG. To minimize the impact of the different classification schemes the AVHRR classes were grouped to correspond to the TM classes as closely as possible and assigned the corresponding TM class number (refer to the second column of **Table 3**). It is evident that the accuracies of the two classification methods were very similar for both study areas. Secondly, the AVHRR class 1 (corresponding to wet conifers in the TM classification, **Table 3**) also contained much of the remaining classes. This was true in all four cases and suggests that the classes occurring in smaller patches are obscured in the 1 $km^2$ pixels and classified as the dominant cover type. Thirdly, TM class 2 (dry conifers) could not be identified in any of the cases, probably because of insufficient spectral contrast and/or small size of the stands.

**Table 3** shows that most of the AVHRR classes consist of mixed pixels at the TM scale. This is true even for the fairly generalized classes and was also observed in other areas (Cihlar *et al.*, 1996). The classification accuracy should thus depend on the land cover homogeneity of the AVHRR pixels, i.e. the degree to which various classes defined by the AVHRR classification legend are mixed within the 1 $km^2$ pixels. This expectation was confirmed by assessing the accuracies of pixels with different degrees of purity, defined as the proportion of

Table 2.

Characteristics of the TM reference classifications.

| Site | Landsat Path/Row (Date) | Scene centre lat/long (°) (Size, km²) | Classification procedure |
|---|---|---|---|
| SSA (BOREAS SSA) | 37/22-23 (90/08/06) | 53.8/105.4 (9850) | Note 1) |
| NSA (BOREAS NSA) | 33/21 (88/08/20) | 55.8/98.2 (9045) | Note 2) |

**Note 1**): Classification of TM bands 1 though 5 and 7 prepared for the BOREAS project was used as the reference (Hall and Knapp, 1994a). It is based on supervised classification of an image from August 6, 1990 image and supported by extensive field checking. The following original classes were identified (the class accuracy based on ground data is given in the parentheses, where reported): 1=wet conifers (primarily black spruce; 96.8%), 2=dry conifer (jack pine; 55.6%), 3=mixed (coniferous and deciduous; 14.8%), 4=deciduous (95.2%), 5=disturbed, 6=fen (wetland; 0%), 7=water, 8=regeneration (medium age), 9=regeneration (younger), 10=regeneration (older; 44.4%), 11=burn (visible). The overall diagonal accuracy was 66.6%, Khat=0.56 (Hall and Knapp, 1994a).

**Note 2**): Classification of TM bands 1 though 5 and 7 prepared for the BOREAS project was used as the reference (Hall and Knapp, 1994b). It is based on supervised classification of an image from August 20, 1988 and supported by extensive field checking. The following classes were identified (the class accuracy based on ground data is given in the parentheses, where reported): 1=wet conifers (primarily black spruce; 77.8%), 2=dry conifer (jack pine; 55.6%), 3=mixed (coniferous and deciduous; 66.7%), 4=deciduous (80.2%), 5=disturbed, 6=fen (wetland; 100%), 7=water, 8=regeneration (medium age), 10=regeneration (older). The overall diagonal accuracy was 73%, Khat=0.63 (Hall and Knapp, 1994b).

AVHRR pixel covered by one type (an AVHRR pixel is 100% pure if all the TM pixels within that AVHRR pixel belong to one TM class). **Table 4** shows the classification results for ECM and CPG for the two study areas and purity levels of 50, 60 and 80%. It is evident that classification accuracy increases with pixel purity, reaching 90% for pixels >80% pure. Note that the lower accuracies occur even at the IGBP-classes level, thus highlighting the fundamental limitation of the AVHRR spatial resolution. On the other hand, the overall patterns generally correspond well on the AVHRR and TM data which is also evident from the increasing accuracy as the AVHRR pixels become more pure.

**Figure 2** and results in **Tables 3** and **4** indicate that the accuracies of both classifications are satisfactory, given the characteristics and limitations of the input data. They also show that the AVHRR absolute accuracy will not approach that of the TM because of the mixed pixel effect, and another strategy must therefore be found to compensate for the resolution difference. It should be noted that the TM classifications were from older images (3 years for SSA, 5 years for NSA) and thus changes could have occurred in these areas, especially due to disturbances such as fire and insect damage. The differences in classes between TM (**Table 2**) and AVHRR (**Table 1**) and the <100% TM accuracy (**Table 2**) also obscured the comparison somewhat. Nevertheless, the basic pattern is not expected to change once these differences are accounted for.

## Comparison of ECM and CPG

**Table 5** shows the results of AVHRR classifications using ECM and CPG methods. The mean area, computed as the average value between the two methodologies, was highest for needleleaf forest (low and very low crown density); the areas of the other types (at level L -2) were much smaller. Overall, coniferous forest covered 48% of the area, deciduous forest 0.2%, and mixed forest 7%. The absolute difference in the area estimated by the two classification methods was small (< 1-2%) in most cases where the same classes were used. The relative difference between the areas of the classes (level L -2) varied over a wider range, from 0.3% for deciduous forest to 177% for non-vegetated land with <25% cover. Generally, the differences were highest for small classes (note that the 200% value for cases where different class definitions were employed is not meaningful). These differences generally occurred at the L -2 level, indicating different spectral class limits. Given the data available, it is difficult to determine the correct value or causes of the differences but the additional use of $N_a$ by the CPG was undoubtedly a contributing factor. At the L -2 level, fairly small spectral differences are sometimes available to make the discrimination. The varying area estimates therefore are also likely due to the differences in quantization used by the ECM and CPG which has an effect on the mean spectral values of the individual clusters, even if the classification is carried out on non-quantized data. In the case of the ECM, the selection of the pixels representing individual categories by the ECM was an additional factor. The somewhat different area totals are due to differences in labeling water-related categories.

The class definitions were mostly identical for ECM and CPG but some differences existed. In general, the approach taken in the CPG was to let the spectral content of the data define the distinct classes. On the other hand, in the ECM the representative samples were selected by the analyst, and the classes of interest could thus be given special attention. For example, the CPG classified burned forest areas mostly as various mixtures of sparse or regenerating vegetation and barren, while they formed a separate class in the ECM (**Table 1**). Similar situation occurred in urban areas, although even the ECM approach was not entirely successful and the known locations of large cities were therefore outlined interactively during post-classification. Conversely, cropland was divided

**Table 3.**

**Confusion matrices for AVHRR classifications of two study areas, BOREAS SSA and NSA, and two classification methodologies, ECM and CPG*.**

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **SSA-ECM** | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| | 1 | 0.91 | 0.92 | 0.68 | 0.39 | 0.67 | 0.91 | 0.23 | 0.83 | 0.82 | 0.86 | 0.91 |
| | 3 | 0.04 | 0.01 | 0.12 | 0.11 | 0.02 | 0.04 | 0.01 | 0.02 | 0.01 | 0.03 | 0.00 |
| | 4 | 0.03 | 0.02 | 0.17 | 0.49 | 0.25 | 0.04 | 0.02 | 0.10 | 0.12 | 0.09 | 0.01 |
| | 5 | 0.00 | 0.00 | 0.00 | 0.00 | 0.02 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | 0.00 |
| | 7 | 0.02 | 0.04 | 0.03 | 0.02 | 0.03 | 0.01 | 0.75 | 0.05 | 0.03 | 0.02 | 0.08 |
| | | **Accuracy (DiAc) = 0.56796** | | | | **Khat=** | **0.3025** | | | | | |
| **SSA-CPG** | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| | 1 | 0.95 | 0.94 | 0.80 | 0.49 | 0.68 | 0.94 | 0.26 | 0.85 | 0.84 | 0.86 | 0.95 |
| | 3 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 4 | 0.04 | 0.02 | 0.17 | 0.49 | 0.25 | 0.05 | 0.02 | 0.12 | 0.12 | 0.12 | 0.01 |
| | 5 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.00 | 0.00 | 0.00 | 0.02 | 0.00 | 0.00 |
| | 7 | 0.01 | 0.03 | 0.02 | 0.02 | 0.02 | 0.01 | 0.73 | 0.02 | 0.02 | 0.01 | 0.04 |
| | | **Accuracy (DiAc) = 0.56671** | | | | **Khat=** | **0.2807** | | | | | |
| NSA-ECM | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 10 | | |
| | 1 | 0.96 | 0.99 | 0.96 | 0.94 | 0.89 | 0.99 | 0.51 | 0.99 | 0.98 | | |
| | 3 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | |
| | 4 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | |
| | 7 | 0.04 | 0.01 | 0.04 | 0.04 | 0.11 | 0.01 | 0.49 | 0.01 | 0.01 | | |
| | | **Accuracy (DiAc) = 0.48736** | | | | **Khat=** | **0.1095** | | | | | |
| NSA-CPG | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 10 | | |
| | 1 | 0.96 | 0.98 | 0.96 | 0.91 | 0.89 | 0.96 | 0.51 | 0.98 | 0.96 | | |
| | 3 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | |
| | 4 | 0.00 | 0.00 | 0.00 | 0.04 | 0.00 | 0.03 | 0.00 | 0.00 | 0.02 | | |
| | 5 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | | |
| | 7 | 0.04 | 0.01 | 0.04 | 0.04 | 0.10 | 0.01 | 0.49 | 0.01 | 0.02 | | |
| | | **Accuracy (DiAc) = 0.48772** | | | | **Khat=** | **0.1138** | | | | | |

**Notes:** The columns represent TM classes (refer to Table 2). Column 2 shows the equivalent TM classes for the AVHRR classes. The assignment was made automatically based on the closest correspondence of the AVHRR and TM classes (see also Cihlar *et al.*, 1996).

into more subcategories in the CPG because of strong spectral differences, possibly due to the additional channel ($N_a$).

The overall difference between ECM and CPG was computed as the mean weighted difference per pixel (MWDP) which accounts for the varying areas (Table 5). As expected, the MWDP depended on the classification level. The overall average was highest (2.0%) at L -2 level, 0.5% at L -1, and 0.7% at L 0 (the standard deviation had values comparable to the average, **Table 5**).

## Land cover composition in relation to pixel size

**Table 6** shows the results of the calculations when using each study area separately and in combination. When taking each study area separately, the area estimated by the two AVHRR classification methods was consistent for most classes but large differences occurred in some. Among the 20 cases (2 TM classes were absent in the NSA) the difference was <1% in 7 cases and <10% in 8. The largest difference in both areas occurred in TM class 5 (disturbed). Similarly, the difference in the area estimated on the basis of the two study areas (SSA, NSA) was relatively small, ranging from 0 to 5% (median value 1%).

When the four area estimates (SSA/NSA, ECM/CPG) are combined the proportion covered by each type assumes an intermediate value (3rd part of **Table 6**). As expected, the maximum difference in the estimated area per class also increases. The 'worst case' difference was computed as the highest minus the lowest among the four estimates divided by the mean value per class. It ranged from 1.2% to 74.9% (median 20%), and it reflects the combined effects of differences

Table 4.
The dependence of the AVHRR classification accuracies on pixel purity*

| Area | Classification Method | Pixel purity (%) | Diagonal accuracy (%) | Khat (unitless) |
|---|---|---|---|---|
| BOREAS SSA | ECM | 0 | 56.8 | 0.303 |
| | | 50 | 67.2 | 0.401 |
| | | 60 | 75.6 | 0.510 |
| | | 80 | 89.3 | 0.783 |
| | CPG | 0 | 56.7 | 0.281 |
| | | 50 | 67.3 | 0.397 |
| | | 60 | 75.2 | 0.505 |
| | | 80 | 89.1 | 0.777 |
| BOREAS NSA | ECM | 0 | 48.7 | 0.110 |
| | | 50 | 57.5 | 0.189 |
| | | 60 | 64.0 | 0.295 |
| | | 80 | 83.9 | 0.670 |
| | CPG | 0 | 48.8 | 0.114 |
| | | 50 | 57.6 | 0.190 |
| | | 60 | 64.0 | 0.295 |
| | | 80 | 84.5 | 0.687 |

Notes:
- SSA= Southern Study Area; NSA= Northern Study Area
- Pixel purity is the proportion of the 1km AVHRR pixel containing a single cover type as determined from the TM classification
- Diagonal accuracy is defined by Equation 2.

in class proportions in each study area and the AVHRR classification methods. The difference is dominated by the class distribution in the study areas used. For example, the highest error (74.9%) represents TM class 2 (dry conifers); according to the TM classification these occupy less in the SSA (0.8%, **Table 6**, part 6) than in the NSA (6.6%), thus biasing the overall estimates. This is consistent with the results of Penner (1995) who, in a study comparing AVHRR classification to the Canadian Forest Inventory, also found that the AVHRR classification tended to underestimate the extent of the less common classes and overestimate the dominant ones.

The impact of the area selected for high resolution coverage is also illustrated in the 4th part of **Table 6** which shows the difference between the total area per class estimated on the basis of SSA and NSA, respectively (e.g., |426487-424878|= 1609 km$^2$, 1609/425683 = 0.4% for wet conifers). The largest relative values occur in TM class 2, 9 and 11 which is again due to the difference in the class structure at each site; note that classes 9 and 11 are absent from the NSA (**Table 6**, part 6). Again, the differences in estimated area of individual classes are much smaller in % of the total area (range 0.2% - 5.1%, median 1.5%).

The estimated land cover composition over the entire area is shown in part 5 of **Table 6**, based on the mean area obtained from the combined ECM and CPG estimates for the two study areas; for the forested area (classified on the TM images) the corresponding area proportions are in part 3 of **Table 6**. Overall, the area is dominated by wet conifers, mixed forest, and water; these three types account for 51% of the total area (71.6% of forested area). The remaining classes are relatively small, occupying <5% each (median 2.2%) of the total area, and <7.4% (median 3.2%) of the forested area.

A comparison of the area estimates by AVHRR alone (**Table 5**) and by AVHRR/TM (**Table 6**) can be done only partially because of the differences in class definitions (**Table 1, 2**). For forest classes, the estimates are coniferous 48% (AVHRR alone) and 31.9 % (AVHRR/TM); deciduous 0.2% vs. 5.3%, and mixed 12.6% vs. 11.1%. The two area estimates thus differ substantially for coniferous and deciduous forest, probably due to the large AVHRR pixel sizes on the one hand, and land cover composition within the two TM scenes on the other. The coniferous AVHRR-only estimate is biased by the sub-pixel water bodies which are contained in the evergreen needleleaf classes. The remaining difference is caused by other cover types intermixed within the prevailing forest cover, and possibly by other factors. The presence of mixtures is also indicated by the confusion matrices (**Table 3**). For the same reason, the area of deciduous forest is underestimated by AVHRR alone. It should also be noted that the land cover composition of the TM scenes influences the estimated composition of the AVHRR pixels as evidenced by the area estimates based on SSA (**Table 6**, part 1) vs. that based on NSA (**Table 6**, part 2). The mixed forest estimates are fairly close.

Part 7 of **Table 6** shows the proportions of land cover obtained from a land cover map by Steyaert and Loveland (1995; labeled S&L below) which was prepared from 1992 monthly NDVI images. The area included in their study is smaller than that covered by the two AVHRR classifications (510517 km$^2$, compared to 1.44x10$^6$ km$^2$), so the comparison cannot be exact. Because of the differences in the area included, the comparison of forested area composition may be more relevant. The estimated distribution of the land cover types was fairly different, more than might be expected from the area covered. As for the ECM/CPG alone, classes occurring in small patches were absent in the S&L. The proportions of coniferous, deciduous and mixed forest were also different. This is true for

**Table 5.**

**Land cover area for the BOREAS area estimated using AVHRR by two classification methods***

| **Land cover category** | | | (ECM-CPG) | | | mean weighted diff./pixel (absol. value) | |
|---|---|---|---|---|---|---|---|
| | mean area | | absol.diff. | relat.diff. | IGBP L -2 | IGBP L -1 | IGBP L 0 |
| L 0, **L -2** | (km**2) | fraction | (unitless) | (unitless) | (unitless) | (unitless) | (unitless) |
| **1. Forest Land** | | | | | | | |
| High crown density (>60%) | 58770.5 | 0.04 | 0.007 | 0.174 | 0.007 | 0.007 | |
| Medium crown density (40-60%) | 29463 | 0.02 | 0.013 | 0.659 | 0.013 | 0.013 | |
| Low crown density (25-40%) | 392692.5 | 0.27 | -0.021 | -0.079 | | 0.021 | |
| Very low crown density (10-25%) | 205811.5 | 0.14 | 0.018 | 0.124 | | 0.018 | |
| **1.1 Evergreen needleleaf forest** | 686737.5 | 0.48 | 0.017 | 0.035 | | | 0.017 |
| Broadleaf forest | 2458.5 | 0 | 0 | -0.003 | 0 | 0 | |
| **1.2 Deciduous broadleaf forest (>80% broadleaf trees)** | 2458.5 | 0 | 0 | -0.003 | | | 0 |
| 60-80% needleleaf trees | **62526** | **0.04** | -0.005 | **-0.104** | | 0.005 | |
| 60-80% broadleaf trees | 35385.5 | 0.02 | -0.009 | -0.365 | | 0.009 | |
| 40-60% broadleaf or needleleaf trees | 3824.5 | 0 | 0.002 | 0.857 | | 0.002 | |
| **1.3 Mixed forest** | 101736 | 0.07 | -0.011 | -0.159 | | | 0.011 |
| 2. Open Land (tree crown density <10%) | | | | | | | |
| Closed shrubland (>50% shrubs), particularly wetlands | **19146.5** | **0.01** | -0.002 | **-0.139** | 0.002 | 0.002 | |
| Open shrubland (25-50% shrubs) | 55616.5 | 0.04 | -0.022 | -0.566 | 0.022 | -0.022 | |
| Shrubs (10-25%) and lichens (<50%) | 16536 | 0.01 | 0.001 | 0.059 | 0.001 | 0.001 | |
| Recent burns | 3965 | 0 | 0.006 | 2 | 0.006 | | |
| **2.1 Shrubland** | 95264 | 0.07 | -0.018 | -0.265 | | | 0.018 |
| 120 Grassland | 57430.5 | 0.04 | 0.012 | 0.304 | 0.012 | 0.012 | |
| **2.2 Grassland** | **57430.5** | **0.04** | 0.012 | **0.304** | | | 0.012 |
| **3. Developed and Mosaic Land** | | | | | | | |
| High biomass (well-watered or irrigated) | 31754 | 0.02 | -0.039 | -1.746 | 0.039 | | |
| Medium-High biomass | 32205 | 0.02 | -0.045 | -2 | 0.045 | | |
| Subhumid in the plains/Medium biomass | 99065.5 | 0.07 | 0.099 | 1.438 | 0.099 | | |
| Medium-Low biomass | 17560 | 0.01 | -0.024 | -2 | 0.024 | | |
| Low biomass (semiarid) | 31045.5 | 0.02 | 0.011 | 0.515 | 0.011 | | |
| **3.1 Cropland (>60% cropland)** | 211630 | 0.15 | 0.002 | 0.016 | | 0.002 | 0.002 |
| Built-up areas | 218 | 0 | 0 | 2 | 0 | | |
| Urban areas | 0 | 0 | 0 | | 0 | | |
| **3.2 Urban and built-up areas** | 218 | 0 | 0 | 2 | | 0 | 0 |
| Cropland-Woodland mosaic | 67733 | 0.05 | 0.035 | 0.738 | 0.035 | | |
| Woodland-Cropland mosaic | 30980.5 | 0.02 | -0.043 | -2 | 0.043 | | |
| **3.2 Cropland/Woodland Mosaic** | 98713.5 | 0.07 | -0.008 | -0.121 | | 0.008 | 0.008 |

a comparison with the ECM/CPG AVHRR classifications (**Table 5**) and to a lesser extent also for the combined AVHRR/TM results (**Table 6**, parts 5 and 7). The AVHRR/TM results contained substantially larger proportions of coniferous and deciduous classes than S&L; significantly smaller mixed forest, medium-age regeneration, and burn classes; and it also contained smaller classes which were not resolved by the AVHRR alone. Some of the differences can be explained by the inclusion of the TM classifications which increase area estimates for the small classes. Beyond that, the input data and the labeling procedure may play some role. For example, the NDVI-sensitive classes containing significant proportions of the deciduous species occupy higher proportions of the S&L classification than for ECM/CPG, and converse is true for the coniferous species with low NDVI; this would be expected since the S&L classification is more sensitive to NDVI differences. While the S&L values for coniferous forest appear to be underestimated, the available data make the resolution of differences difficult. The TM classifications (**Table 6**, part 6) indicate much higher proportions of the coniferous forest and

**Table 5. (continued)**

| Land cover category | | | (ECM-CPG) | | | mean weighted diff./pixel (absol. value) | |
|---|---|---|---|---|---|---|---|
| 4. Non-Vegetated Land | | | | | | | |
| Patches of stunted black spruce | 8246 | 0.01 | 0.006 | 1.082 | 0.006 | | |
| Lichens and shrubs | 15333 | 0.01 | -0.01 | -0.94 | 0.01 | | |
| Low vegetation cover (< 25%) | 6912 | 0 | 0.009 | 1.773 | 0.009 | | |
| Very low vegetation cover (<10%) | 5094.5 | 0 | 0.001 | 0.145 | 0.001 | | |
| Low/very low cover | 35585.5 | 0.02 | 0.005 | 0.211 | | 0.005 | |
| Barren | 2373 | 0 | -0.002 | -1.064 | | 0.002 | |
| **4.1 Barren** | 37958.5 | 0.03 | 0.003 | 0.131 | | | 0.003 |
| | | 0 | | | | | |
| Water, snow, land ice | 147853.5 | 0.1 | 0.002 | 0.017 | 0.002 | 0.002 | |
| **4.2 Other** | 147853.5 | 0.1 | 0.002 | 0.017 | | | 0.002 |
| Total | 1440000 | 1 | | | | | |
| Mean | | | | | 0.014 | 0.005 | 0.007 |
| Stdev | | | | | 0.02 | 0.009 | 0.007 |

*** Notes:**
Amean= (AECM+ ACPG )/2
ADA= (AECM+ ACPG)/1440000
ADR= (AECM+ ACPG)/Amean
MWDP= |ADA/Amean|;
ACPG = area estimated by the CPG classification
AECM= area estimated by the ECM classification
Amean= mean area (km2)
ADA= area difference (absolute)
ADR= area difference (relative)
MWDP= mean weighted difference per pixel.

less regeneration than S&L but these represent only a portion of the whole area and were selected in relatively well-forested areas. Furthermore, some of the TM classes have low accuracies (**Table 2**). It should be expected that the area proportions would also change if S&L results are combined with TM data.

## Comments

Although combined AVHRR/TM cover type area estimates in Table 6 are considered fairly representative of the conditions in the BOREAS Region it should be possible to refine these further through several steps.

First, the TM classifications used in the calibration process can be improved. The TM classifications used were based on older data (**Table 2**), thus introducing uncertainty into the classification process, especially because of the disturbances and regeneration that occur between the TM and AVHRR data acquisitions. The TM classification methodology also needs more attention to maximize the classification accuracy of these data. In addition, better sampling of the cover type mixtures with TM data is needed to minimize the bias introduced into the calibration (**Table 6**). This can be accomplished by adding TM scenes, selected along the transect joining NSA and SSA or sampling using the ecoregion framework (Ecological Working Group, 1989). However, it will be important to ensure that the scenes are representative of the proportions of cover types in the ecoclimatic region; deviations from this requirement will lead to errors in estimating the areas of individual cover types. Another important step is the reconciliation of the two classification legends. Although the IGBP classification has the advantage of linking the results to the national and global estimates, the TM legend has been specifically designed to facilitate biogeochemical modeling. Ideally, the two should be combined, to allow estimates of all the classes at the regional level. It should be readily possible to map all the classes at the TM level.

Further examination of the ECM and CPG procedures is needed to better understand the reasons for the differences in area estimates in some cases, and possible ways to resolving these. This includes the possible effects of residual noise in the AVHRR input data. Although the 1993 data are not likely to be improved significantly by further processing, it should be possible to use data from another year, alone or in combination with those from 1993, to refine the cover type estimates for the region. However, it appears that the use of various classification methodologies (and also data from adjacent years) provides an additional quality control mechanism for ensuring the accuracy and consistency of land cover estimates from remote sensed data. Another potential area for improvement is in the integration

**Table 6.**

**Cover type area estimates for the BOREAS Region.**

| **Area estimate based on:** | TM 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Based on SSA | | | | | | | | | | | | |
| SSA-CPG | 427951 | 8360 | 144773 | 103361 | 59327 | 29533 | 161978 | 39334 | 39989 | 65531 | 5987 | 1086124 |
| SSA-ECM | 425023 | 8624 | 122979 | 66446 | 12249 | 29177 | 170682 | 37930 | 17121 | 60261 | 7661 | 958153 |
| Mean | 426487 | 8492 | 133876 | 84903 | 35788 | 29355 | 166330 | 38632 | 28555 | 62896 | 6824 | 1022139 |
| % area | 42 | 1 | 13 | 8 | 4 | 3 | 16 | 4 | 3 | 6 | 1 | 100 |
| difference (% of mean) | 0 | 2 | 8 | 22 | 66 | 1 | 3 | 2 | 40 | 4 | 12 | 6 |
| 2. Based on NSA | | | | | | | | | | | | |
| NSA-CPG | 419949 | 57709 | 188371 | 71077 | 55491 | 33705 | 135049 | 21418 | | 47276 | | 1030045 |
| NSA-ECM | 429808 | 57121 | 183871 | 63434 | 43201 | 33299 | 118671 | 24362 | | 47820 | | 1001587 |
| Mean | 424878 | 57415 | 186121 | 67256 | 49346 | 33502 | 126860 | 22890 | | 47548 | | 1015816 |
| % area | 42 | 6 | 18 | 7 | 5 | 3 | 12 | 2 | 0 | 5 | 0 | 99 |
| Difference (% of mean) | 1 | 1 | 1 | 6 | 12 | 1 | 6 | 6 | | 1 | | 1 |
| 3. Based on SSA+NSA | | | | | | | | | | | | |
| Mean | 425683 | 32954 | 159998 | 76080 | 42567 | 31428 | 146595 | 30761 | 28555 | 55222 | 6824 | 1018977 |
| % area | 41.6 | 3.2 | 15.7 | 7.4 | 4.2 | 3.1 | 14.3 | 3 | 2.8 | 5.4 | 0.7 | 99.7 |
| stdev(%mean) | 1 | 86 | 20 | 24 | 50 | 8 | 16 | 30 | 57 | 17 | 17 | 5 |
| difference (% mean) | 1.2 | 74.9 | 20.4 | 26.2 | 55.3 | 7.2 | 17.7 | 29.1 | 40 | 16.5 | 12.3 | 6.3 |
| 4. Differences due to the TM classified area | | | | | | | | | | | | |
| Area | 1609 | 48923 | 52246 | 17648 | 13558 | 4147 | 39470 | 15742 | 28555 | 15347 | 6824 | 6323 |
| % of area | 0.4 | 148.5 | 32.7 | 23.2 | 31.9 | 13.2 | 26.9 | 51.2 | 100 | 27.8 | 100 | 0.6 |
| % of total area | 0.2 | 4.8 | 5.1 | 1.7 | 1.3 | 0.4 | 3.9 | 1.5 | 2.8 | 1.5 | 0.7 | 0.6 |
| 5. Fraction of total area | | | | | | | | | | | | |
| % of total area | 29.6 | 2.3 | 11.1 | 5.3 | 3 | 2.2 | 10.2 | 2.1 | 2 | 3.8 | 0.5 | 70.8 |
| % of forested area | 41.8 | 3.2 | 15.7 | 7.5 | 4.2 | 3.1 | 14.4 | 3 | 2.8 | 5.4 | 0.7 | 100 |
| 6. Fraction of area based on TM only | | | | | | | | | | | | |
| SSA (%) | 46.9 | 0.8 | 17 | 9.5 | 1.1 | 3.7 | 9.5 | 2.8 | 1.3 | 6.3 | 0.4 | 99.3 |
| NSA (%) | 45.3 | 6.6 | 20.1 | 4.2 | 4 | 2.9 | 10.3 | 2.4 | 0 | 3.7 | 0 | 99.5 |
| 7. Fraction of total area (Steayert and Loveland, 1995) | | | | | | | | | | | | |
| % of total area | 24.9 | 5.1 | 20.8 | 0 | 6 | 0 | 12.2 | 14.8 | 0 | 0 | 4.1 | 87.9 |
| % of forested area | 28.3 | 5.8 | 23.6 | 0 | 6.8 | 0 | 13.9 | 16.9 | 0 | 0 | 4.6 | 100 |

of TM and AVHRR classifications results, e.g. by the use of inverse calibration methods (Walsh and Burk, 1993). Work in these areas is in progress.

## Summary and Conclusions

This study was undertaken to characterize land cover composition in the BOREAS Region. A combination of AVHRR and TM satellite data was used, and two AVHRR classification methods were employed. The results can be summarized as follows:

1. The forested areas in the BOREAS Region contains approximately 30% wet conifers, 11% mixed, and 10% water; the remaining 8 categories occupy <5% each. It is expected that these estimates can be further refined by using time-coincident data sets, more extensive high spatial resolution coverage, and single classification legend.
2. Combining higher resolution classifications over limited areas with coarser resolution ones over the entire region of interest is a cost-effective approach to land cover mapping over large areas. However, the area estimates are sensitive to the representativeness of the cover type mixtures in the area mapped at high spatial resolution.

3. The two AVHRR classification methods (ECM and CPG) yielded consistent and similar results. Using these different classification methods may be a useful strategy for increasing the robustness of the cover type area estimates.

## ACKNOWLEDGMENTS

The 1993 AVHRR data were processed by Ms. Pat Hurlburt at the Manitoba Remote Sensing Centre. The AVHRR post-processing was conducted by Ms. Fengting Huang of Intermap Technologies. M. Guy Simard of the Canadian Forest Service (CFS) made strong contribution to AVHRR cluster labeling. Dr. Steen Magnussen from CFS provided constructive comments on a draft of this paper. The assistance of these scientists is gratefully acknowledged.

## REFERENCES

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

Beaubien, J., and G. Simard. 1993. "Methodologie de classification des donnees AVHRR pour la surveillance du couvert vegetal." *Proceedings of the 16th Canadian Remote Sensing Symposium*, Sherbrooke, Quebec: 597-603.

Beaubien, J., J. Cihlar, G. Simard, and Q. Xiao. 1997. "Land cover from Thematic Mapper using new enhancement – classification methodology." *Canadian Journal for Remote Sensing* (this issue).

Belward, A. (Ed.) 1996. "The IGBP-DIS global 1 km land cover data set DISCover." Proposal and implementation plans. IGBP Working Paper # 13, IGBP Data and Information System Office. 61p.

Brown, J.F., T.R. Loveland, J.W. Merchant, B.C. Reed, and D.O. Ohlen. 1993. "Using multisource data in global land-cover characterization: concepts, requirements, and methods." *Photogrammetric Engineering and Remote Sensing* 59: 977-987.

Cihlar, J. 1996. "Identification of contaminated pixels in AVHRR composite images for studies of land biosphere." *Remote Sensing of Environment* 56:149-163.

Cihlar, J., and P.M. Teillet. 1995. "Forward piecewise linear model for quasi-real time processing of AVHRR data." *Canadian Journal for Remote Sensing* 21: 22-27.

Cihlar, J., H. Ly, and Q. Xiao. 1996. "Land cover classification with AVHRR multichannel composites in northern environments." *Remote Sensing of Environment* 58:36-51.

Cihlar, J., J. Beaubien, Q. Xiao, J. Chen, and Z. Li. 1997a. "Land cover of Canada in 1993: methodology and preliminary results". In preparation.

Cihlar, J., H. Ly, Z. Li, J. Chen, H. Pokrant, and F. Huang. 1997b. "Multitemporal, multichannel AVHRR data sets for land biosphere studies: artifacts and corrections." *Remote Sensing of Environment.* (in print)

Cihlar, J., J. Chen, and Z. Li. 1997c. "Seasonal AVHRR multichannel data sets and products for scaling up biospheric processes." *Journal of Geophysical Research - Atmospheres.* Accepted.

Cihlar, J., Q. Xiao, J. Chen, and J. Beaubien. 1997d. "Classification by Progressive Generalization: a new automated image analysis methodology for remote sensing data". In preparation.

Cihlar, J., J. Beaubien, Q. Xiao, J. Chen, and Z. Li. 1997e. "Land cover of Canada in 1993: methodology and preliminary results". In preparation.

Ecological Working Group. 1989. "Ecoclimatic Regions of Canada." Ecological Land Classification Series No. 23, Environment Canada. 118p.

Eidenshink, J.C., and J.L. Faundeen. 1994. "The 1 km AVHRR global land data set: first stages in implementation." *International Journal for Remote Sensing* 15: 3443-3462.

Hall, F.G., and D. Knapp. 1994a. "Landsat TM forest cover classification image of BOREAS Southern Study Area." The BOREAS Information System, NASA Goddard Space Flight Center, Greenbelt, MD. (available as a digital file).

Hall, F.G., and D. Knapp. 1994b. "Landsat TM forest cover classification image of BOREAS Northern Study Area." The BOREAS Information System, NASA Goddard Space Flight Center, Greenbelt, MD. (available as digital file).

Li, Z., J. Cihlar, X. Zhang, L. Moreau, and H. Ly. 1996. "Detection and correction of the bidirectional effects in AVHRR measurements over northern regions." *IEEE Transactions on Geoscience and Remote Sensing* 34: 1308-1322.

Loveland, T.R., J.W. Merchant, D.O. Ohlen, and J.F. Brown. 1991. "Development of a land-cover characteristics database for the conterminous U.S." *Photogrammetric Engineering and Remote Sensing* 57: 1453-1463.

Penner, M. 1995. "Comparing AVHRR and CANFI 91 land classes." *Canadian Journal for Remote Sensing* 21: 10-15.

Robertson, B., A. Erickson, J. Friedel, B. Guindon, T. Fisher, R. Brown, P. Teillet, M. D'Iorio, J. Cihlar, and A. Sancz. 1992. "GEOCOMP, a NOAA AVHRR geocoding and compositing system." *Proceedings of the ISPRS Conference*, Commission 2, Washington, D.C.: 223-228.

Rowe, J.S. 1972. "Les regions forestieres du Canada." Publication No. 1300F, Service canadien des forets. 172p.

Sellers, P., F. Hall, H. Margolis, R. Kelly, D. Baldocchi, J. den Hartog, J. Cihlar, M. Ryan, B. Goodison, P. Crill, J. Ranson, D. Lettenmeier, and D. Wickland. 1995. "The Boreal Ecosystem-Atmosphere Study (BOREAS): an overview and early results from the 1994 field year." *Bulletin of the American Meteorological Society* 76: 1549-1577.

Sellers, P.J., C.J. Tucker, G.J. Collatz, S.O. Los, C.O. Justice, D.A. Dazlich, and D.A. Randall. 1994. "A global 1o by 1o NDVI data set for climate studies. Part 2: The generation of global fields of terrestrial biophysical parameters from the NDVI." *International Journal of Remote Sensing* 15: 3519-3545.

Steyaert, L., and T. Loveland. 1995. "1-km AVHRR seasonal land cover classification for the Boreal Ecosystem-Atmosphere Study (BOREAS)." The BOREAS Information System, NASA Goddard Space Flight Center, Greenbelt, MD. (Available as a digital file).

U.S. Department of Commerce, 1977. World Data Bank II. Volume 1 - North America and General User's Guide. National Technical Information Service, Springfield, VA.

Walsh, T.A., and T.E. Burk. 1993. "Calibration of satellite classifications of land area." *Remote Sensing of Environment* 46: 281-290.

Wu, A., Z. Li, and J. Cihlar, J. 1995. "Effects of land cover type and greenness on AVHRR bidirectional reflectances: analysis and removal." *Journal of Geophysical Research* 100:9179 9192.