INT. J. REMOTE SENSING, 10 DECEMBER, 2004,
VOL. 25, NO. 23, 5337–5344

Taylor & Francis
Taylor & Francis Group

Sensitivity of hyperclustering and labelling land cover classes to Landsat image acquisition date

M. A. WULDER*†, S. E. FRANKLIN‡ and J. C. WHITE†

†Canadian Forest Service, 506 West Burnside Road, Victoria, BC, Canada V8Z 1M5
‡Department of Geography, University of Saskatchewan, Saskatoon, Saskatchewan, S7N 5A2, Canada

(*Received 10 March 2003; in final form 3 December 2003*)

Abstract. Seven Landsat images near Prince George, British Columbia, Canada, representing a range of within-year and between-year dates, were acquired to assess spectral variability and the concomitant impact upon hyperclustering classification results. Top-of-atmosphere (TOA) radiometric corrections, dark target subtractions and geometric corrections were applied to the imagery. Following application of an unsupervised hyperclustering procedure which employed the K-means classifier, post-classification comparisons examined the differences in spectral response patterns for several target classes, and area summaries were generated to compare the variability in the total area of classes as identified in each image. Finally, the kappa coefficient of agreement was used to quantify the degree of correspondence between the classified images. The results indicated that the spectral variability of the within-year image set exceeded the variability in the between-year image set and differences in class area were highly variable over the range of image acquisition dates. These findings suggest that off-year imagery (acquired on or near anniversary dates) may be preferred to off-season imagery when building large-area Landsat mosaics for land cover classification using the hyperclustering procedure.

1. Introduction

A hyperclustering and labelling procedure is a simple and relatively common approach to classify multiple scene Landsat mosaics (Talbot and Markon 1988, Bauer *et al.* 1994, Debinski *et al.* 1999, Cihlar 2000). Typically, this approach is recommended in areas where little is known of the class structure or where training (field) data are scarce or impractical to acquire (Franklin and Wulder 2002). The process is to generate many hyperclusters from the image data available by testing for within-cluster heterogeneity, merge the hyperclusters into a smaller number of more reasonable groupings which may resemble homogeneous classes, and then label the resulting classes as spatial features of interest according to a pre-determined map legend or class hierarchy. Hyperclustering and labelling requires relatively little manual effort, thereby limiting field and ancillary data needs (Homer

*e-mail: mwulder@nrcan.gc.ca

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online
http://www.tandf.co.uk/journals
DOI: 10.1080/0143116042000192385

et al. 1997). The approach appears to have the potential to be robust and repeatable and has been recommended as the basic method to be used in the production of a Landsat-based land cover map of the forested area of Canada representing year 2000 conditions (Wulder *et al.* 2003) using the Centre for Topographic Information (CTI) satellite image database (Wulder *et al*. 2002).

Typically, at the heart of this classification approach is a simple clustering algorithm; one of the most common such algorithms is known as K-means (Tou and Gonzalez 1974). The K-means algorithm is influenced by the number of cluster centres specified, the choice of initial cluster centres, the order in which the samples are taken and the geometrical properties of the data. This research addresses the question: what is the influence of the scene-dependent factors on the ability of the hyperclustering approach to generate the land cover classes required over large, spatially variable image mosaics? A test of the ability of the K-means algorithm to produce consistent classes from imagery acquired under different conditions was required; therefore, an experiment was devised involving seven different Landsat images acquired over one growing season and three years, in the region of Prince George, British Columbia, Canada.

2. Study area and methods

The study area was located near Prince George, British Columbia, Canada, and was centred at approximately 123° 23′ 32.50″ W, 54° 27′ 50.91″ N. Seven Landsat Thematic Mapper (TM) images (WRS Path 49, Row 22) were acquired from three different years and five different months (see table 1). Clouds were a problem in the majority of the images; only the 3 August 1999 image was reasonably cloud-free. This image was selected as the master image for the comparison tests in this paper. The master image was orthorectified using 15 Ground Control Points (GCPs) resulting in an RMS error of 1.33 pixels in the *x*-direction and 1.18 pixels in the *y*-direction. The remaining six images were registered to the master image using cubic convolution resampling. Both a top-of-atmosphere (TOA) radiometric correction and a dark-object subtraction procedure were applied to each of the seven images. Details of the methods used are given in Markham and Barker (1986) and Peddle *et al*. (2003). Since the analysis involved post-classification comparison between images, this degree of atmospheric correction was deemed appropriate (Song *et al*. 2001). Table 2 contains the values used in the dark-object subtraction for each image/band combination; for each image, these values were obtained in three deep,

Table 1. Seven TM images used in analysis of hyperclustering and labelling within-year and between-year variability.

Satellite	Date	Solar zenith (°)	Percentage cloud cover[1]
Landsat 5	31 July 1998	51	0 (1102)
Landsat 5	29 April 1999	47	0 (0010)
Landsat 5	18 July 1999	53	10 (1111)
Landsat 5	3 August 1999	50	0 (0000)
Landsat 7	12 September 1999	38.6	0 (1100)
Landsat 5	20 September 1999	35	10 (1111)
Landsat 7	26 June 2000	56.5	10 (1111)

[1]The first value is the overall percentage cloud cover. The second value in brackets is the cloud cover index of a scene quadrant in tens of percentage points. To obtain the percentage cloud cover of a quadrant, the corresponding index must be multiplied by 10.

Table 2. Dark-object subtraction values for each reflective band in each scene based on the average of three deep, dark lake radiances.

	TM bands					
Date	1	2	3	4	5	7
31 July 1998	59	19	18	9	4	2
29 April 1999	59	20	17	11	7	5
19 July 1999	58	17	16	10	5	2
3 August 1999	57	18	14	7	4	2
12 September 1999	45	25	17	12	9	9
20 September 1999	51	15	13	7	6	3
26 June 2000	55	30	22	10	9	8

dark lakes and then averaged to create the estimate for the subtraction process. The K-means algorithm was implemented using the five-step process outlined in table 3.

3. Results

Post-classification comparisons of the seven Landsat TM images indicated the existence of large differences in the outputs of the hyperclustering and labelling procedures. Visually, the different imagery contained large and variable amounts of cloud, haze, ice and snow, which affected the initial clustering sequence and results. The images displayed differences in general land cover condition that are consistent with seasonal changes in vegetation pattern (Guyot *et al*. 1989). For example, the visual inspection of normal colour composites (TM bands 3, 2 and 1 viewed through red, green and blue channels, respectively) indicated that there was more green coloured vegetation in the June, July and August images; conversely there was more brown vegetation in the April and September images. The presence of cloud also affected the classification results: in the 20 September 1999 image, clouds completely obscured the city of Prince George, and no urban/cultural/road class could subsequently be distinguished in the hyperclusters.

The analysis of the classification results involved several comparisons between the classified outputs generated from the K-means algorithm. First, the differences

Table 3. Implementation of the K-means algorithm.

Step	Action
1	Manual generation of a cloud mask for each image (interpreter delineated clouds manually with the aid of interactive reflectance thresholds).
2	K-means clustering after 12 iterations (the maximum allowable in the formulation available) and 50 clusters requested (this amount was thought to represent a reasonable compromise between the number of clusters that could be managed reasonably by a human interpreter in the following step and the maximum number – more than 240 – that can be generated in the available code). All six optical TM bands were used as inputs for the clustering.
3	Cluster merging based on Bhattacharyya distance separability measures (Bhattacharyya 1943) and visual cues in the imagery for known class features (utilizing air photos and local knowledge).
4	Recoding of clusters to a nine class scheme (unclassified/ice/snow/clouds, water, dense conifer, open conifer, deciduous/shrub, shrub/grass mix, grass, exposed soil, urban/cultural/roads).
5	Export of recoded class scheme to GIS map creation and display software.

in spectral response patterns for several target classes were compared across image dates. Second, the total class areas were compared between image dates and the maximum and minimum differences in class areas were summarized. Finally, a pixel-by-pixel comparison of the image classes was conducted by subtracting each image from the master and compiling the differences for the entire TM scene.

The spectral response patterns of several target classes for TM bands 1, 2 and 3 are shown in figure 1. Several features changed reflectance dramatically as a consequence of the formation and melting of snow and ice, classes that were most prevalent in the 29 April 1999 image. In this image, the recently cut areas appeared to contain significant variability related to the presence of snow ('snow-effect') and this greater reflectance is common in all three visible bands; however, a reasonably stable spectral response pattern existed for these features in the visible bands in the other images. There was also a marked increase in the reflectance of recent cuts (and all other target classes) in the 12 September 1999 imagery, specifically in TM band 2. This may be caused by a relatively larger amount of haze in this scene or by the senescence of deciduous species found in these recent cuts.

In contrast to the spectral response seen in new cuts, older cuts appeared to have less snow-effect in the 29 April 1999 image, and less variability from one image date to the next. Also in the older cuts, in the 26 June 2000 scene, the reflectance in TM band 2 was greater than either bands 1 or 3; this was interpreted as indicating a higher shrub content in these older cut blocks and early emergence of deciduous plants (leaf-on). The other target classes of interest (coniferous, deciduous, wetland) followed a similar pattern in spectral response to older cuts: the snow-effect in the April image resulted in increased reflectance for TM bands 1, 2 and 3; reflectance is relatively stable for July and August images; and reflectance increases in the 12 September 1999 image in bands 2 and 3, most likely as a result of a relative increase in vegetation wetness. Extraction of the TM band 4 (infrared) reflectance in the dense conifer class provided for a detailed consideration of within- and between-year spectral characteristics. The histograms presented in figure 2 illustrate that the reflectance values are less stable in the within-year comparison than in the between-year comparison.

To determine the impact of the spectral variability on classification results, the area differences between the classes were compared. In table 4 the minimum and maximum area differences for each of the images, relative to the master image (3 August 1999), are presented. In six of nine classes, the maximum difference (with clouds masked) occurred between the master image and the April image, suggesting that the use of images with large amounts of ice and snow strongly impairs generation of the desired target classes. As proof of this, four of the target classes were not present in the April image (open conifer, deciduous/shrub, shrub/grass mix and grass). The next greatest number of differences was found between the master and the 31 July 1998 image, probably caused by the presence of a large cloud bank in this image. Finally, the two September images contained the maximum difference from the master image in three classes: the shrub/grass mix, grass and urban/cultural/roads. The latter difference was almost certainly caused by the large cloud positioned over the city of Prince George in the 20 September 1999 image. Thus, although seasonal differences may strongly influence classification results, image anomalies such as clouds and haze may also have an impact on hyperclustering and labelling of classes.

The final step in the analysis was to complete a pixel-by-pixel post-classification comparison of each of the images with the master image in order to quantify the

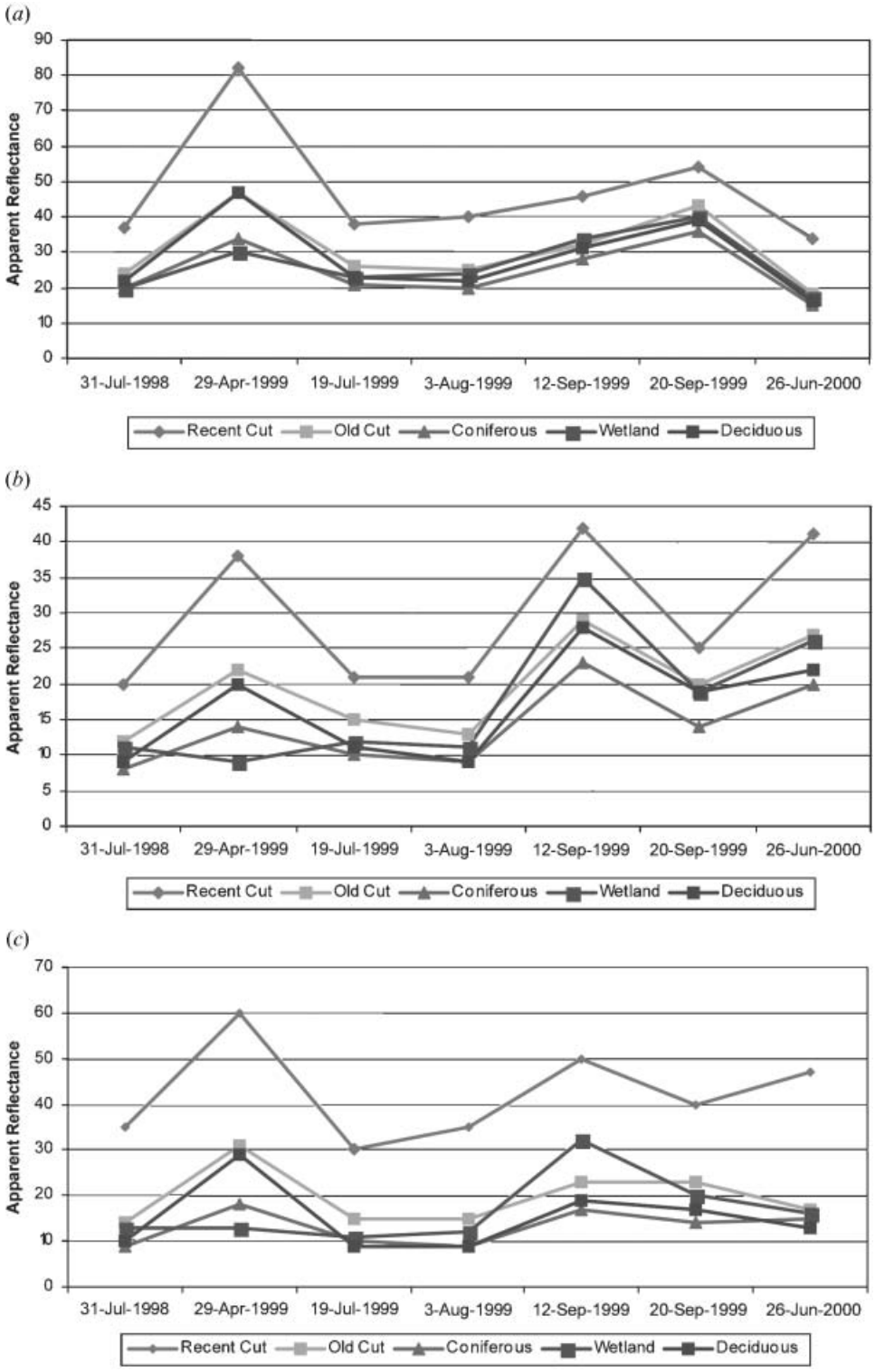

Figure 1. Spectral response patterns for certain target classes. (*a*) Landsat TM band 1; (*b*) Landsat TM band 2; (*c*) Landsat TM band 3.

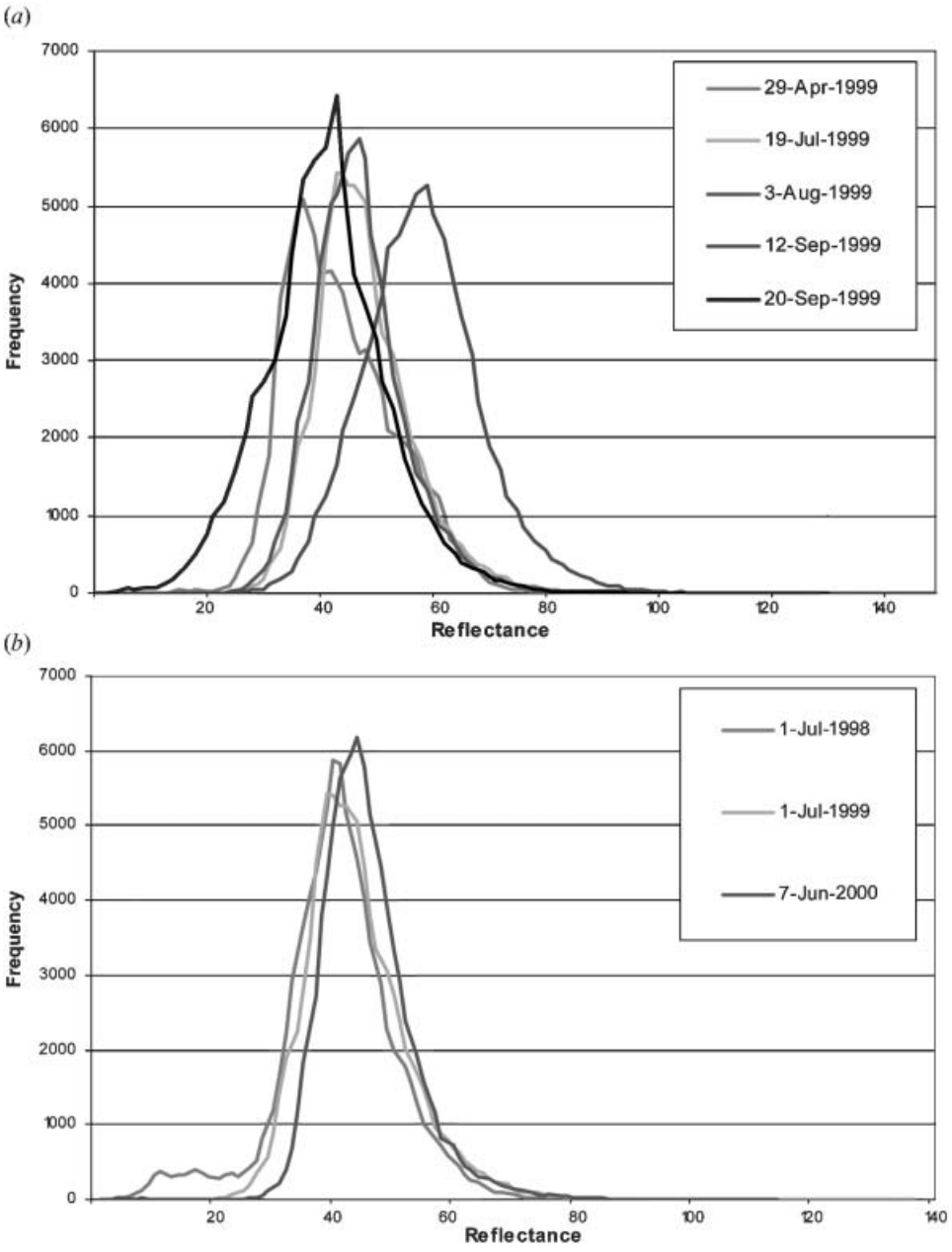

Figure 2. Histograms of Landsat TM channel 4 (near-infrared) reflectance for the dense conifer class in: (*a*) five within-year TM images of Prince George, British Columbia, Canada and (*b*) three near-anniversary, between-year images.

level of agreement between them. The kappa coefficient of agreement suggested that the correspondence was at best moderate between the master and the 18 July 1999 and 26 June 2000 classifications. The other four classifications only weakly corresponded to the master (table 5). These results suggested a low persistence and stability of classifications representing the differing dates.

Table 4. Class areas and maximum difference in class areas using the 3 August 1999 image as the master (areas calculated with clouds masked).

Class[1]	Master (ha)	Minimum difference (%)	Date	Maximum difference (%)	Date
0	545 186	−1.73	12 Sep 1999	3.27* (2.30)	31 July 1998
1	30 268	3.60	18 July 1999	262.93* (46.59)	26 June 2000
2	786 783	−2.32	20 Sep 1999	−37.12	31 July 1998
3	98 232	−8.58	26 June 2000	276.82	31 July 1998
4	140 906	−3.89	12 Sep 1999	−100.0* (56.87)	18 July 1999
5	80 148	−7.44	26 June 2000	171.51* (−49.50)	20 Sep 1999
6	49 508	−18.39	26 June 2000	−100.0* (−47.21)	12 Sep 1999
7	75 250	−4.92	26 June 2000	−100.0* (89.08)	31 July 1998
8	10 268	12.49	12 Sep 1999	756.47	20 Sep 1999

*The largest difference was with the snow-covered 29 April 1999 imagery; second largest difference shown in brackets.

[1]Classes: 0 = unclassified/ice/snow/clouds; 1 = water; 2 = dense conifer; 3 = open conifer; 4 = deciduous/shrub; 5 = shrub/grass mix; 6 = grass; 7 = exposed soil; 8 = urban/cultural/roads.

Table 5. Kappa coefficient of agreement over all classes using the 3 August 1999 image as the master.

Date	Kappa
31 July 1998	0.3593
29 April 1999	0.3266
18 July 1999	0.5762
12 Sep 1999	0.4485
20 Sep 1999	0.3794
26 June 2000	0.5234

4. Conclusions

Differences in class area were highly variable over the analysed range of image acquisition dates when a K-means hyperclustering and labelling procedure was used in Prince George, British Columbia. In addition, the spectral variability of the within-year image set was observed to exceed that in the between-year image set. This is an important finding as it suggests that off-year imagery (acquired on, or near anniversary dates) may be preferred to off-season imagery when building large-area Landsat mosaics for land cover classification using the hyperclustering procedure. When selecting imagery it is important to remain as close to peak photosynthetic activity as possible. However, because of anomalies such as clouds and haze, acquisition during these peak times may not be feasible and the results of this research indicate that users must consider the implications of the selection and inclusion of an off-season image with caution.

Acknowledgments

This research was funded by the Canadian Space Plan of the Canadian Space Agency support of Earth Observation for Sustainable Development of Forests program. Additional support was also provided by the Canadian Forest Service and the Natural Sciences and Engineering Research Council of Canada.

References

BAUER, M. E., BURK, T. E., EK, A. R., COPPIN, P. R., LIME, S. D., WALSH, T. A., WALTERS, D. K., BEFORT, W., and HEINZEN, D. F., 1994, Satellite inventory of Minnesota forest resources. *Photogrammetric Engineering and Remote Sensing*, **60**, 287–298.

BHATTACHARYYA, A., 1943, On a measure of divergence between two statistical populations defined by their probability distributions. *Bulletin of the Calcutta Mathematical Society*, **35**, 99–109.

CIHLAR, J., 2000, Land cover mapping of large areas from satellites: status and research priorities. *International Journal of Remote Sensing*, **21**, 1093–1114.

DEBINSKI, D.M., KINDSCHER, K., and JAKUBAUSKAS, M. E., 1999, A remote sensing and GIS-based model of habitats and biodiversity in the Greater Yellowstone ecosystem. *International Journal of Remote Sensing*, **20**, 3281–3291.

FRANKLIN, S., and WULDER, M., 2002, Remote sensing methods in high spatial resolution satellite data land cover classification of large areas. *Progress in Physical Geography*, **26**, 173–205.

GUYOT, G., GUYON, D., and RIOM, J., 1989, Factors affecting the spectral response of forest canopies: A review. *Geocarto International*, **3**, 3–18.

HOMER, C. G., RAMSEY, R. D., EDWARDS, T. C., and FALCONER, A., 1997, Landscape cover-type modeling using a multi-scene Thematic Mapper mosaic. *Photogrammetric Engineering and Remote Sensing*, **63**, 59–67.

MARKHAM, B., and BARKER, J., 1986, Landsat MSS and TM post calibration dynamic ranges, exoatmospheric reflectances and at satellite temperature. *EOSAT Landsat Technical Notes*, **1**, 3–7.

PEDDLE, D. R., TEILLET, P. M., and WULDER, M. A., 2003, Radiometric image processing. *In Remote Sensing of Forest Environments: Concepts and Case Studies*, edited by M. A. Wulder and S. E. Franklin (Boston: Kluwer Academic Publishers), pp. 181–208.

SONG, C., WOODCOCK, C. E., SETO, K. C., LENNEY, M. P., and MACOMBER, S. A., 2001, Classification and change detection using Landsat TM data: When and how to correct atmospheric effects? *Remote Sensing of Environment*, **75**, 230–244.

TALBOT, S. S., and MARKON, C. J., 1988, Intermediate-scale vegetation mapping of Innoko National Wildlife Refuge, Alaska, using Landsat MSS digital data. *Photogrammetric Engineering and Remote Sensing*, **54**, 377–383.

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

WULDER, M., DECHKA, J. A., GILLIS, M. A., LUTHER, J. E., HALL, A., BEAUDOIN, A., and FRANKLIN, S. E., 2003, Operational mapping of the land cover of the forested area of Canada with Landsat data: EOSD land cover program. *Forestry Chronicle*, **79**, 1075–1083.

WULDER, M., LOUBIER, E., and RICHARDSON, D., 2002, A Landsat-7 ETM+ orthoimage coverage of Canada. *Canadian Journal of Remote Sensing*, **28**, 667–671.