Global Ecology & Biogeography (2003) **12**, 161–172

RESEARCH PAPER

Land use and cover with intensity of agriculture for Canada from satellite and census data

JEREMY T. KERR* and JOSEF CIHLAR

Canada Centre for Remote Sensing, Natural Resources Canada, 588 Booth Street, Ottawa, Ontario, Canada K1A 0Y7

ABSTRACT

Aim To develop the first national databases on land use and agricultural land use intensity in Canada for a wide variety of environmental monitoring applications.

Location Canada.

Methods In this paper, we describe a new system for the construction of both land use and land use intensity (within agricultural regions) called LUCIA (land use and cover with intensity of agriculture). Our methodology combines the highly detailed Canadian Census of Agriculture and recent growing season composites derived from the SPOT4/VEGETATION sensor. Census data are of much coarser resolution than the remotely sensed data but, by removing non-agricultural pixels from each census sampling area, we were able to refine the census data sufficiently to allow their use as ground truth data in some areas. The 'refined' census data were then used in the final step of an unsupervised classification of the remotely sensed data.

Results and main conclusions The results of the land use classification are generally consistent with the input census data, indicating that the LUCIA output reflects actual land use trends as determined by national census information. Land use intensity, defined as the principal component of census variables that relate to agricultural inputs and outputs (e.g. chemical inputs, fertilizer inputs and manure outputs), is highest in the periphery of the great plains region of central Canada but is also very high in southern Ontario and Québec.

Key words Canada, environmental monitoring, land use, land use intensity, remote sensing, sustainable development.

INTRODUCTION

Land use data represent an important baseline for environmental monitoring and policy initiatives (Frolking *et al*., 1999; Hurtt *et al*., 2001). Among these, the Kyoto Protocol and Convention on Biological Diversity both require detailed information on contemporary land use. The by-products of some land uses cause significant environmental damage and directly influence human and ecosystem health (Nielsen, 1999). For example, agricultural runoff is high in phosphorus, a nutrient that causes eutrophication of aquatic ecosystems (Schindler, 1974). Various land uses may also lead to toxic chemical accumulation in the environment (Blais *et al*., 1998). Pollutants may affect aquatic species, such as molluscs, particularly severely but terrestrial vertebrate species are also at risk (e.g. peregrine falcon *Falco peregrinus*; Martin, 1978; Bromley, 1992). Environmental factors that relate directly to human health, such as water quality, are also subject to degradation when agricultural land use intensity is too high (Medema *et al*., 1997).

Broad-scale land use measurements cannot be made purely based on remote sensing data, at least if detailed land use data are desired. Land cover, on the other hand, may be derived from remote sensing data alone (e.g. Townshend *et al*., 1987; Cihlar *et al*., 2000). A single land cover type (e.g. low biomass agriculture) may have multiple uses (e.g. rangeland, hay and grain production; Cihlar & Jansen, 2001). The role of ancillary data in the development of land use information is primarily to constrain cases where there is a one-to-many mapping from land cover to land use. Spatially extensive land uses are often agricultural in nature, so agricultural census data may help in the derivation of land use information (Frolking *et al*., 1999; Baban & Luke, 2000; Hurtt *et al*., 2001).

Remote sensing plays a key role in the development of land use data. Anderson *et al*. (1976) developed an early U.S. land use/cover analysis from aerial photography and Landsat 1 data based on a highly detailed, hierarchical framework, which was implemented largely through intensive manual

* Correspondence: Jeremy T. Kerr, Department of Biology, University of Ottawa, Ottawa, Ontario, Canada, K1N 6N5
E-mail: jkerr@uottawa.ca

methods. Frolking *et al.* (1999) developed land use predictions for agricultural areas in China using AVHRR LAC coverage (1.1 km nadir resolution) and agricultural census data resolved at county scale. These authors found that remote sensing measurements of total cropland area made consistent predictions ($R^2 = 0.80$) of census-based crop extents, but were generally 48–104% higher than the census estimates. This discrepancy is thought to arise in part from unreliable census data from China that under-report cropland extents (Ji *et al.*, 2001), while remotely sensed data probably overestimate the extent of agriculture in their study (Frolking *et al.*, 1999). Remote sensing vs. census estimates of particular agricultural land uses, however, tended to be rather poorly correlated in China (Frolking *et al.*, 1999). In the United States, recent land use measurements (Hurtt *et al.*, 2001) employ matrix transition approaches to establish a generalized (few classes), coarse resolution (0.5°) land use classification from AVHRR land cover data and agricultural census statistics. More recently, CORINE has been developing European land cover/land use data from high resolution SPOT HRV and Landsat Thematic Mapper (TM) sources (Mucher *et al.*, 2000), augmented by national statistical data sources (CEC, 1993).

While land use classification is of considerable interest and utility for environmental monitoring, aspects of land use relating to pollution are usually ignored in land use classifications. One approach to measuring the potential for pollution in agricultural areas, defined as 'land use intensity' in this study, is to measure agricultural inputs, typically fertilizers and chemical (pesticide) additions, and by-product outputs (e.g. manure production). Inputs and by-products include pesticides, fertilizers and manure. These substances are prime sources of non-point-source pollution, which frequently constitutes the main cause of aquatic pollution. Pesticides, which are typically relatively harmless to humans in trace quantities, may be biomagnified through trophic interactions (e.g. Kelly & Gobas, 2001) or accumulate in certain animal tissues, potentially causing direct or indirect long-term health effects for humans and other animals. Some common pesticides may impair human health (Safe, 2000) through endocrine disruption (e.g. vinclozolin; Kelce & Wilson, 1997; Sonnenschein & Soto, 1998). Integration of oft-overlooked land use intensity data from agricultural areas with land use classification procedures would expand the utility of land use monitoring initiatives considerably.

In this paper, we describe a hybrid procedure for generating land use and land use intensity maps for agricultural regions from remote sensing and census data sources and apply that process to Canada. The procedure, called Land Use and Cover with Intensity of Agriculture (LUCIA), relies on new satellite image classification techniques and data sources with ancillary data to derive national-scale land use predictions. It also enables estimation of land use intensity in agricultural regions, which are concentrated in the low relief, southern areas of the country, concomitant with human population density and areas of high species diversity (Kerr & Packer, 1997; Kerr *et al.*, 2001).

METHODS

The LUCIA process fuses spatially refined data from the Canadian Census of Agriculture with processed SPOT4/Vegetation data. The process is detailed below (and see flow chart in Fig. 1).

Land use/cover

We used the most recent land cover classification for Canada (developed using SPOT4/Vegetation; see Cihlar *et al.*, 2001) as the starting point for the development of Canadian land use data. From this land cover map, we created a mask for agricultural, urban (within agricultural districts) and grassland pixels and used this mask to select the area for more detailed re-classification from initial remote sensing imagery. We isolated these parts of Canada for three reasons. Most importantly, the effects of human-dominated land uses are especially significant when natural habitats have been converted to agriculture or urban areas. Secondly, ancillary data needed to support the development of a land use data product are most readily available for agricultural areas, which includes grassland in agricultural censuses. Census of Agriculture data (Statistics Canada, 1996) were reported for each

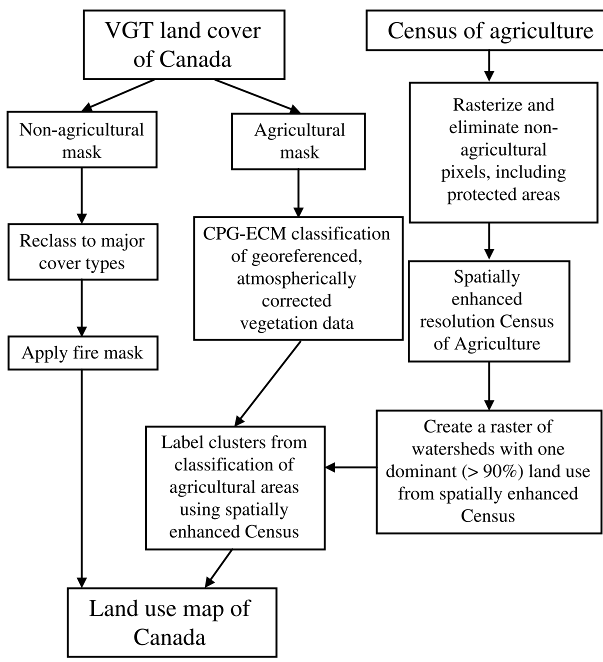

Fig. 1 This flowchart depicts the process that combines SPOT4/Vegetation data with spatially refined Census of Agriculture data to generate a 1-km resolution land use and cover database.

watershed that includes any agricultural activity throughout Canada. Thirdly, the types of land uses that may be detected using satellite data differ markedly across the boundaries of agricultural and nonagricultural areas.

The georeferenced 10-day composites of Canada's land surface used to create the land cover classification within agricultural regions consisted of atmospherically corrected, 1 km resolution, surface reflectance data for each of three image channels (red, NIR and SWIR) from the SPOT4/VEGETATION (VGT) sensor. We normalized all pixels to a 45° solar zenith angle and nadir view angle by adjusting for bidirectional reflectance effects with refinements for hotspot description (Cihlar *et al.*, 2003). Cloud and haze contamination were removed using CECANT (Cloud Elimination from Composites using Albedo and NDVI Trends; explained fully in Cihlar *et al.*, 1997). CECANT detects subpixel cloud contamination within each pixel by interpolating the seasonal trend in NDVI, determining whether a given pixel drops transiently below predicted levels (indicative of cloud contamination), and replaces that pixel with an interpolated value. This method has been addressed in detail elsewhere (see Cihlar *et al.*, 1997, 2001) and creates cloud-free image composites successfully for the entirety of Canada, which is a cloudy region. All imagery was projected to the Lambert Conformal Conic (LCC) projection through latitudinal parallels at 49° and 77° and with a central meridian at 95°W. We processed satellite data in PCI (PCI Geomatics, 2000), carried out all statistical analysis in Systat version 10 (SPSS Software, 2000) and all GIS modelling and analysis in Arc/Info GRID 8.02 (Environmental Systems Research Institute, 2000).

We classified the corrected VGT data for agricultural regions using the CPG-ECM unsupervised classification system (Classification by Progressive Generalization — Enhancement-Classification Method; Cihlar *et al.*, 2000; Beaubien *et al.*, 1999). Initially, three input image channels (NIR, SWIR, RED = RGB), consisting of fully corrected surface reflectance data, were contrast-stretched and used as input for K-Means cluster analysis. The initial cluster procedure created a large number of clusters (about 170) that retained nearly all detail from the contrast-enhanced data. CPG then identified spatially and spectrally dominant clusters and iteratively eliminated the 'least important' (small or dispersed) clusters through cluster merging. The procedure refined the classified data to 60–70 clusters with little or no visually noticeable loss of detail (Fig. 2). Clusters representing visually distinctive

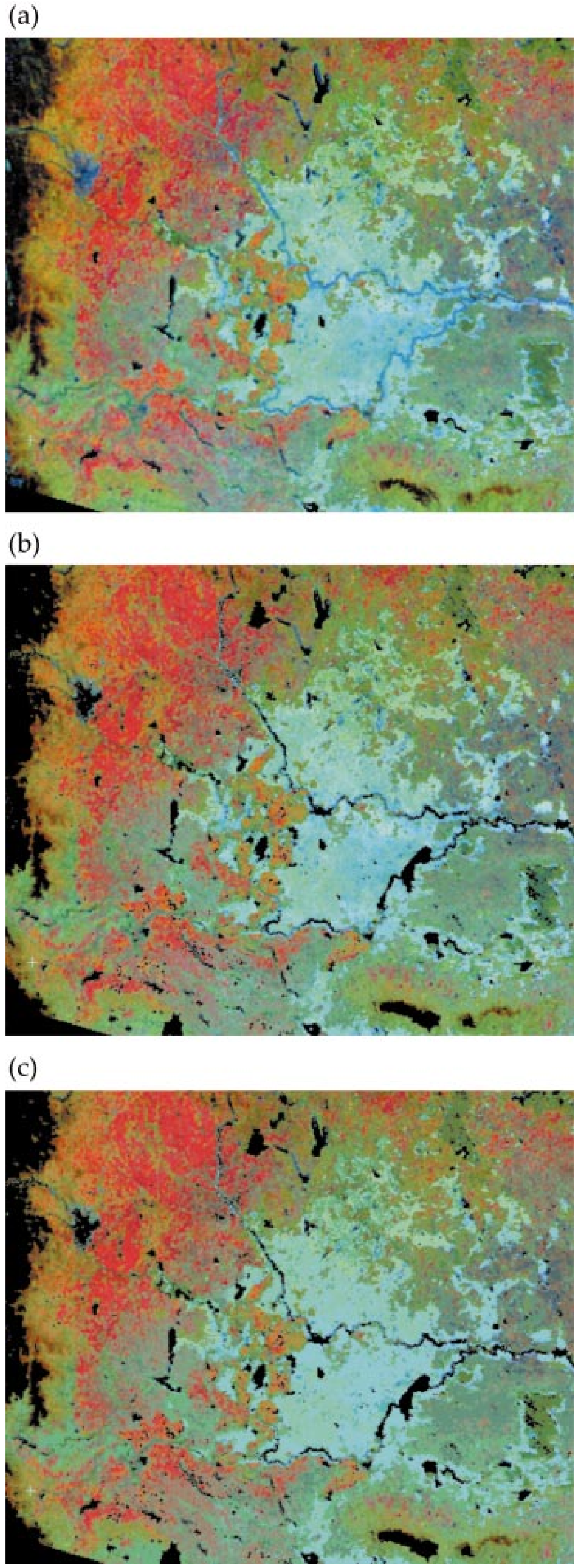

Fig. 2 The classification method used in this study retains visible detail while refining input imagery to progressively smaller numbers of spectral clusters. The contrast-stretched image (a) includes both agricultural and non-agricultural areas. The agricultural areas of this image were used to generate a 150-cluster image (b). This was then refined to about 70 clusters (c). While retaining visual detail, the exclusion of non-agricultural areas spatially refines the census database for labelling process for the final, reduced-cluster image.

varieties of high biomass agriculture were incorrectly merged into a single cluster between eastern (corn/soybean) and western Canada (oilseeds). These were separated manually. Subsequent labelling (assigning clusters to land cover categories within a legend) relied extensively on known or expected spectral characteristics of given vegetation types and required spatially refined Census of Agriculture data (see below). Urban land use cannot be identified reliably using coarse resolution remote sensing data in Canada and these areas were identified and delimited separately based on standard, urban area vector data from Statistics Canada.

Throughout non-agricultural areas, we combined land cover classes from the initial VGT land cover data to reflect distinctions in potential land use (as explained below) and updated these with ancillary data where possible. This procedure assumes that the VGT land cover classification successfully differentiates between agricultural and non-agricultural areas. We tested this assumption by examining the relationship between land cover-based estimates of agricultural area in each watershed vs. Census of Agriculture estimates, which are completely independent (Fig. 3). Small errors may arise because some areas that are classified as non-agricultural (e.g. Christmas tree farms or some orchards may appear to be forest land cover) may be used for agriculture. Correction of such errors, which are very small, is not possible with the data available for this study.

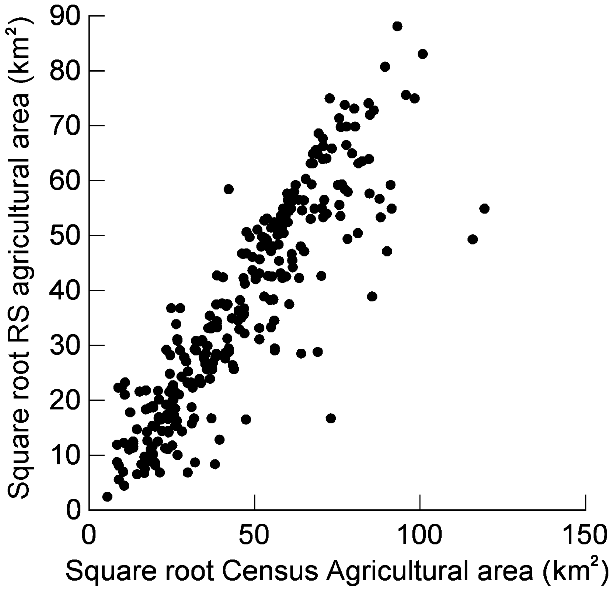

Fig. 3 The relationship between VGT land cover vs. Census of Agriculture estimates of agricultural area per watershed (km^2). There is strong agreement between these measurements ($r^2 = 0.78$, $P << 10^{-6}$, $n = 336$).

Ancillary data

We assembled ancillary data for the agricultural regions of Canada to develop a ground truth database for development of land use from remote sensing data. The primary data source was the 1996 Census of Agriculture. Because the high resolution census data are held as strictly confidential (to protect the privacy of census respondents), we could access only an aggregated form of the database that subdivides agricultural regions into 336 areas derived from watershed boundaries identified by Statistics Canada. This aggregated version provided little indication of the type of agricultural activities predominating in each region (e.g. the census may indicate that 30% of a watershed consists of pasture but not the distribution of that land use within the watershed). However, we refined the Census data spatially according to simple logical rules. Every watershed includes both agricultural and non-agricultural cover types, but the Census of Agriculture indicates land uses exclusively for agricultural areas. Using VGT-based land cover data, non-agricultural pixels could be removed from each watershed in a spatial refinement process. Protected areas (e.g. national parks) were also removed from the watersheds. For example, the Census of Agriculture may report that there are 3000 km^2 of pasture in a 6000 km^2 watershed. The VGT land cover product indicates that 2500 km^2 of the example watershed is coniferous or deciduous forest, so these forest pixels can be removed (they are not agricultural). The census data now indicate that 3000 of the remaining 3500 1 km^2 pixels remaining in the watershed consist of pasture. However, there is a single, large protected area that amounts to 350 km^2, which can also be eliminated because this area cannot be included in the census. The result is a spatially refined Census of Agriculture report for the watershed that indicates, once all non-censused pixels are removed, that each remaining pixel has a (3000 km^2/3150 km^2 × 100%) 95% chance of being pasture. By repeating this spatial refinement process for every partially agricultural watershed in Canada (336 of them), we transformed the extremely coarse Census of Agriculture data reports that are generated for each watershed into a strong indicator of particular land uses throughout Canada that we then related back to the results of the classification of the VGT imagery (see Fig. 1 and classification description, above). Overall, spatial refinement of watershed-level Census of Agriculture data reduced apparent watershed size, from 5905 km^2 before removing non-agricultural land to 4806 km^2 afterward (a 19% mean reduction because many watersheds were almost entirely agricultural while others were highly reduced).

Protected areas may have multiple simultaneous uses, such as conservation, recreation or resource extraction, but they are not subject to the same range of uses present in less regulated areas. Consequently, we maintain data on the location of protected areas in this analysis but we have not designated a single land use for them. The geographical extent of the protected areas in Canada was obtained from an unpublished source (World Wildlife Fund).

In forest ecosystems, both natural disturbances (e.g. fire and insect damage) and management operations (logging, road building) have important land use implications. We do not have access to reliable insect damage data so we have selected only fire for inclusion in LUCIA. Forest fires are a leading source of disturbance in the boreal forests. Forestry activities are affected by fire, albeit often by the need to conduct salvage harvests in extensive areas of burned forest: we include burned areas as a class in the land use map because of their impact on land use but acknowledge that specific land uses arising due to burns are not always predictable. Data on burned area extent and distribution are from 1994 to 2000. These factors are measured with intensively tested and highly accurate AVHRR fire detection and burned area-mapping algorithms (Fraser *et al*., 2000), while older burned areas are identified from the land cover map.

Intensity of agriculture

While reliable data for land use classes are desirable for a number of reasons, agricultural land use intensity data are also of considerable importance. 'Land use intensity' can be defined in a number of ways. In this study, we define it to be a combination of major agricultural pollution sources that are directly measured in the Census of Agriculture. These represent — or are surrogates for — the major material inputs and by-product outputs within agricultural areas. Specifically, we selected seven variables to characterize agricultural pollution sources:

1 Total fertilizer purchased (in $Cdn).
2 Total chemicals (insecticide, herbicide, fungicide) purchased ($Cdn).
3 Total fertilizer applied (tonnes).
4 Total area sprayed for insects (hectares).
5 Total area sprayed for weeds (hectares).
6 Total manure output (kg).
7 Number of cattle.

We used these data as inputs for a principal components analysis (PCA), from which we saved the first two principal components. PCA reduces variables to a smaller number of factors that maximize the variation in common to all variables but that are completely uncorrelated with one another (are orthogonal in n-space). While PCA creates severe difficulties when used to describe the relationship between independent and dependent variables, it remains very effective for reducing the dimensionality of independent datasets (Kent & Coker, 1992). The two principal components (INTENSITY-I and INTENSITY-II) comprise integrated indices of land use intensity that were assigned to each watershed throughout the agricultural regions of Canada (Fig. 4). The reasons for using this statistical method were (i) to reduce the number of input/output variables into an 'index' of land use intensity and (ii) to mask the individual data values of each input variable,

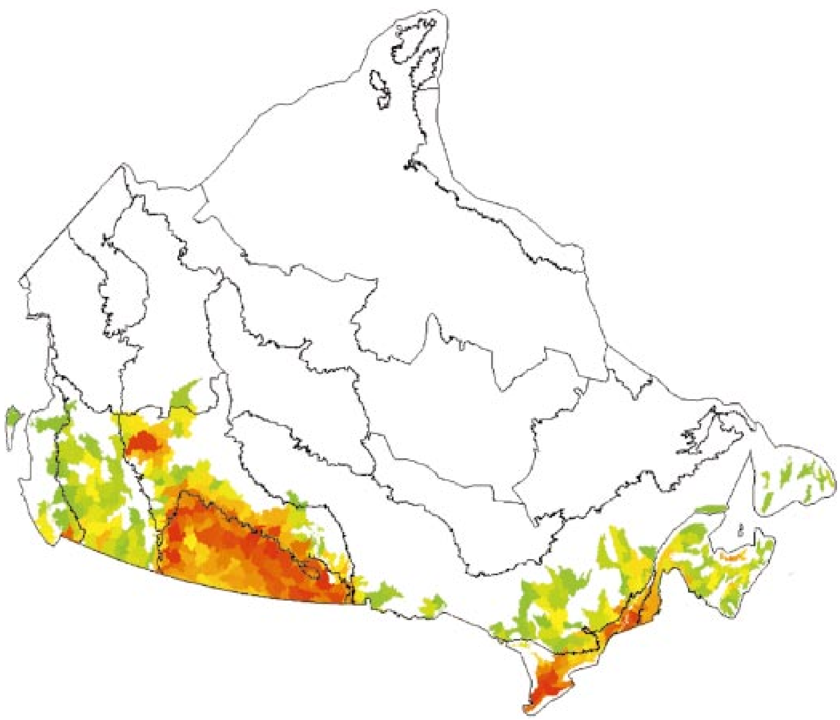

Fig. 4 The spatial distribution of the primary measurement of agricultural land use intensity in Canada (INTENSITY-I), derived from the first component of a PCA on Census variables that measure agricultural inputs and by-product outputs. Intensity-I extracts 71.7% of the variation in common to the variables used in the principal components analysis. The scale is relative: green indicates low land use intensity, yellow is intermediate and red indicates high agricultural inputs and/or by-product outputs.

enabling this procedure to be applied to confidential, higher-resolution census data. To determine which measurement of intensity related best to each of the input variables, we then examined the graphical relationships between both sets of variables (input and intensity) and tested the strengths of observed relationships using linear regression analysis.

Validation

There are no existing products describing contemporary land use for Canada that permit comprehensive product validation. This problem is common for land use products in general (e.g. Hurtt *et al.*, 2001). However, we were able to extract pasture statistics from the census data according to watershed and relate pasture extent to results from our land use classification. While the census data provide only summary statistics on a per watershed basis, they represent the most comprehensive and extensively verified land use summary available in Canada. It is crucial that any land use classification for Canada be broadly consistent with patterns of land use as reported by the census. Pasture is found throughout all agricultural regions and so can provide an index of the success of the land use classification throughout the agricultural area of Canada. We measured pasture area for each watershed from our land use classification (PA_{LUCIA}) and from the census (PA_{CENSUS}) along with watershed area ($AREA_{WS}$). We evaluated the relationship between PA_{LUCIA} and PA_{CENSUS}, including $AREA_{WS}$ as a covariate to reduce the likelihood of observing spuriously strong relationships because of area effects (large watersheds tend to have greater PA_{LUCIA} and PA_{CENSUS} values because of covariation with $AREA_{WS}$ rather than generally successful pasture detection).

RESULTS AND DISCUSSION

As in most countries, Canada's land use represents an array of agricultural, natural, and urban uses. Output from LUCIA (Figs 4 and 5, Table 1) refines these categories of use substantially without subdividing land use types to the point where we are likely to have exceeded the capacity of the census data to indicate specific land uses. Agricultural land uses are heavily concentrated in the prairie region, the Peace River area north-west of the prairies, southern British Columbia, southern Ontario and Quebec, and more sporadically in the Atlantic provinces. Overall, extensive and often intensive croplands and various types of pasture dominate Canadian agriculture. Agricultural land uses vary regionally with respect to which crops are most prevalent. Canadian agriculture from the prairie ecozone and west (including the Peace River valley, north-west of the central plains) is known for its grain (wheat, barley, and others) and oilseed (canola and sunflower) production. Many other crops are grown in the prairies and in the Peace River valley, such as sugar beets and pulses, but these could not be consistently discerned spectrally in the VGT composite data to allow labelling of these sorts of production. Corn, soybean, some grains and alfalfa are the most widespread crops grown in the agricultural land use regions of eastern Canada. Numerous other crops are also grown in some eastern regions, such as tobacco in southern Ontario, but these generally comprise small proportions of the total crop in each pixel. Wheat, canola, corn and soybean all have high biomass spectral signatures that are characterized by high NIR reflectance, slightly lower MIR and low red reflectance. Such spectral response is consistent with high leaf area index, high vegetation moisture content and high absorption of long wavelength visible (red) radiation by photosynthetic pigments. Agricultural areas with high alfalfa have lower NIR reflectance, higher MIR reflectance and comparable red absorption.

Pasture is widespread in Canada and an important constituent of agricultural land use. This is a compromise for some pasture areas (e.g. the Great Sand Hills in Saskatchewan) that retain substantially natural ecosystem processes, although contemporary grazing patterns are not comparable to pre-settlement conditions (Thorpe & Goodwin, 1997). However, in areas such as the Great Sand Hills in Saskatchewan large populations of native herbivores remain (as many as 10 000 mule deer in the Sand Hills; Thorpe & Goodwin, 1997) and continue to contribute to natural grazing cycles. Cattle now graze throughout these areas in place of bison but probably do not fill quite the same ecological niche. Alfalfa, included as pasture, is abundant in the east. Most farmers cut alfalfa crops two or three times a year and feed the hay to their own livestock (Desjardins, pers. comm.). Western pastures are somewhat more varied in their spectral signatures, particularly because of the presence of large ranches, expanses of rangeland, or because the pastures are found in areas of high topographical heterogeneity. Along the eastern foothills of the Rockies, for example, are extensive fescue grasslands that are used mainly for pasture. Similarly, the grassland area of the central prairies actually contains some nearly natural prairie ecosystem remnants but these are distributed disproportionately among small protected areas (such as Grasslands National Park; Government of Canada, 1997). Such grassland areas, including those that are protected, are also grazed to varying degrees.

Several of the land use classes are mixed, typically combining local high biomass crops (e.g. corn and soybean) and pasture. These are found often in transition zones between regions dominated by cropland and pasture, respectively. The problem of mixed pixels is, unsurprisingly, ubiquitous in satellite data: explicit recognition of mixed land uses reflects land use transitions far better than arbitrary assignment of spectrally mixed pixels to a single unmixed land use class.

LUCIA considers protected areas separately, reflecting their multiple use, multiple stakeholder milieu. Instead of reporting

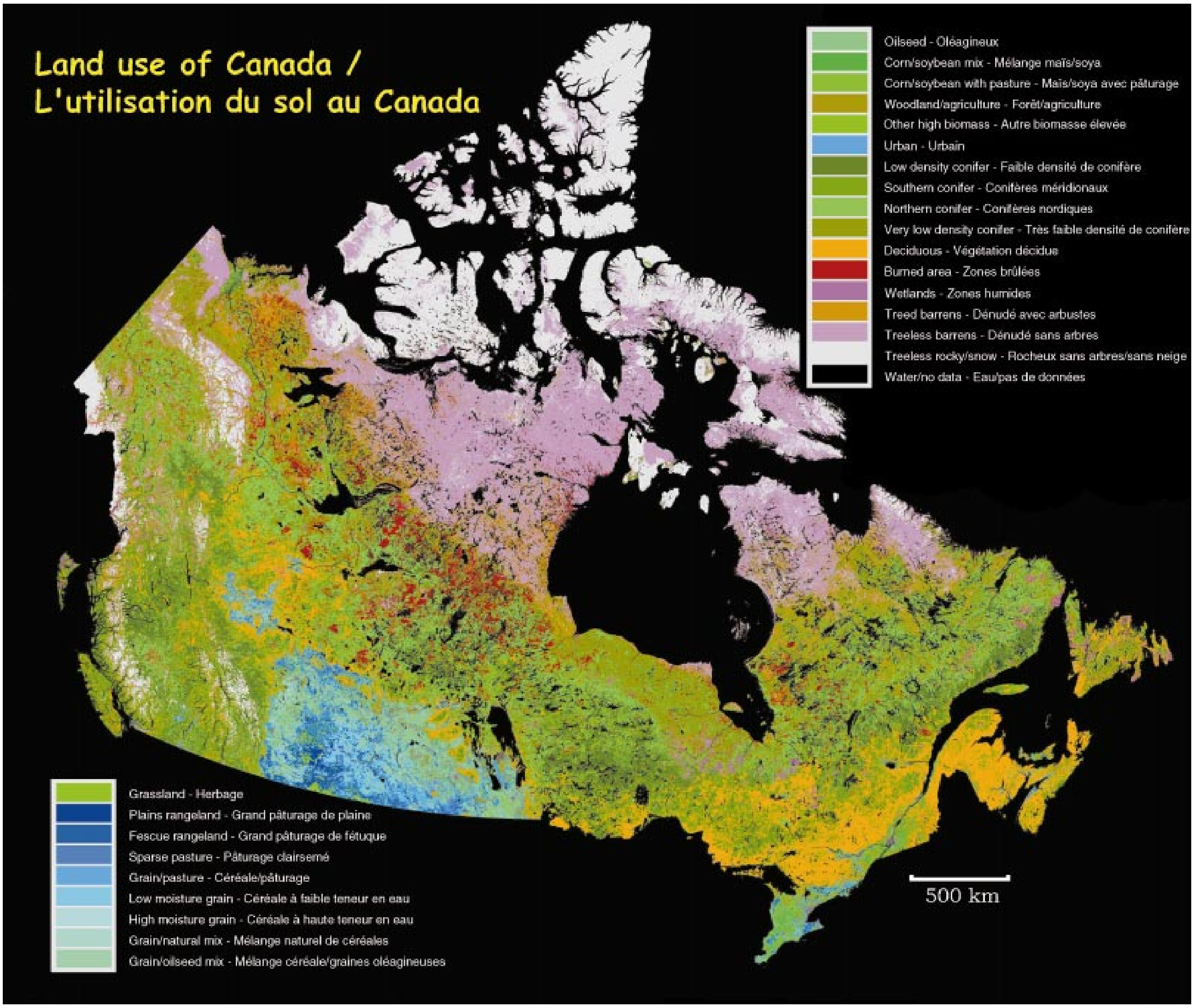

Fig. 5 Land use classification for Canada derived from the LUCIA process. VGT data, providing 1-km spatial resolution, are used for the classification of land use, which also relies on Census of Agriculture data. Protected areas, which may have multiple uses, are not included on this figure so as to retain the clarity of the land use classification.

protected areas as a particular land use, we provide the boundaries of protected areas and show their constituent land cover, as derived from the LUCIA land use/cover classification procedure. The justification for this approach is that land cover in protected areas strongly influences their respective uses. Protected areas take many forms in Canada, including community pasture areas managed by provincial agencies, national and provincial parks, occasional private nature reserves, areas managed cooperatively with First Nations communities, private resource extraction interests, and those concerned with biodiversity conservation. In addition, multiple use PAs are fairly common in the Canadian parks system (e.g. Yellowstone to Yukon wildlife corridor; Soulé & Terborgh, 1999). Algonquin Park, a key constituent of the Ontario provincial protected areas network, is subject to logging throughout most of its extent, and is bisected by a highway and affected by a neighbouring rail corridor (Government of Ontario, 1999). Most parks are also associated with recreational land uses and all serve conservation purposes.

Land use intensity

Land use intensity varies substantially within agricultural regions. INTENSITY-I accounts for most of the variability

Table 1 Land use classification legend with descriptions of each land use class. Classes 1–15 constitute agricultural land uses and were newly derived from processed VGT data. Other classes are derived directly from the most recent national land cover data for Canada, also derived from the SPOT4/Vegetation sensor

Land use/cover type	Description
1. Grassland	Limited to southern-central prairies, these areas are rarely grazed by native bison but rather at low intensity by other herbivores, mostly cattle
2. Rangeland: central prairies	Very low vegetation cover with little moisture. Moderately low intensity grazing
3. Pasture	Found along the shoulder of the Rockies predominantly, and also in eastern Canada
4. Very low vegetation cover	These areas (e.g. Great Sand Hills) are often used as community pastures but may also be more open rangelands
5. Grain/pasture mixed use	Often found in areas of transition between pasture and grain production
6. Grain: low moisture	Found most commonly in prairies but also in very well-drained soil areas of southern Ontario
7. Grain: improved moisture/soil	This land use/cover is concentrated in the prairies and is often found in irrigated areas, leading to higher biomass crop production than in drylands
8. Grain or natural vegetation mixed with water	A mixed land use/cover class often found around the periphery of water bodies in high-intensity agricultural areas
9. Grain and oilseed	A mixed land use primarily observed in the prairies, the major site for oilseed production, which blends various oilseeds (e.g. canola and sunflower) with grain production
10. Oilseeds	Croplands used for oilseed production (e.g. canola and sunflower) are most frequently found in the prairies
11. Corn/soybean	Southern Quebec and Ontario have relatively extensive areas dedicated to corn, often used for livestock feed, and soybean
12. Corn/soybean with pasture	A mixed land use class often observed in eastern Canada that frequently abuts land use class 3
13. Woodland: agricultural use	May vary from majority cropland to majority woodland
14. Other high biomass crops	These may include pulses (peas, beans, etc.), tobacco, sunflower, sugar beets, or other crops
15. Urban/bare	Urban areas not detected in original agricultural mask but found outside of major urban areas Some confusion is possible between urban areas identified using standard urban area masks (from Statistics Canada) and bare ground detected with 1-km VGT data
16. Conifers in boreal region	High–medium density coniferous forests
17. Southern conifers	Southern coniferous forests often mixed with broadleaf
18. Northern conifer	Low-density coniferous cover with variable undergrowth
19. Low-density conifer/barren	Various understories are possible in this transitional, predominantly northern cover type
20. Deciduous forests	May be mixed with some conifer and have variable crown density
21. Post-burn	Derived from initial VGT land cover classification and burned area detection algorithms for 1994–2000
22. Wetlands	Wetlands are not easily mapped in Canada and may include a variety of distinctive ecosystems, such as bogs, fens, swamps, and seasonally inundated areas
23. Treed barren lands	Predominantly a northern (or very high elevation) class of taiga regions
24. Treeless barrens	These ecosystems are found in tundra regions (or very high elevations)
25. Treeless rocky/snow cover	Little vegetation with predominantly northern distributions

Boundaries of protected areas are depicted on the land use/cover layer. These may have multiple simultaneous uses so their constituent covers are depicted rather than assigning all protected areas to a single category.

among the variables used to derive the two land use intensity principal components ($R^2 = 0.717$; see Fig. 6a–c; Table 2). INTENSITY-II makes a significant, secondary contribution to the measurement of land use intensity ($R^2 = 0.161$) but relates more strongly to manure output per watershed than does INTENSITY-I. Its relationship with other input variables is considerably weaker than that between input variables and INTENSITY-I. Interestingly, manure output correlates best with INTENSITY-II but total cattle per watershed was best described by INTENSITY-I. This may be attributable to the inclusion in manure output estimates of contributions from other livestock (e.g. horses and pigs) apart from cattle. We did not have access to numbers of horses, for example, nor chickens and pigs, all of which contribute substantially to manure production in some areas.

Agricultural land use intensity (INTENSITY-I and INTENSITY-II) is highest in southern Ontario, at the periphery of the prairie ecozone, and in south-western British Columbia near the city of Vancouver. Intensity declined in the central and southern areas of the prairies, reflecting the lower density

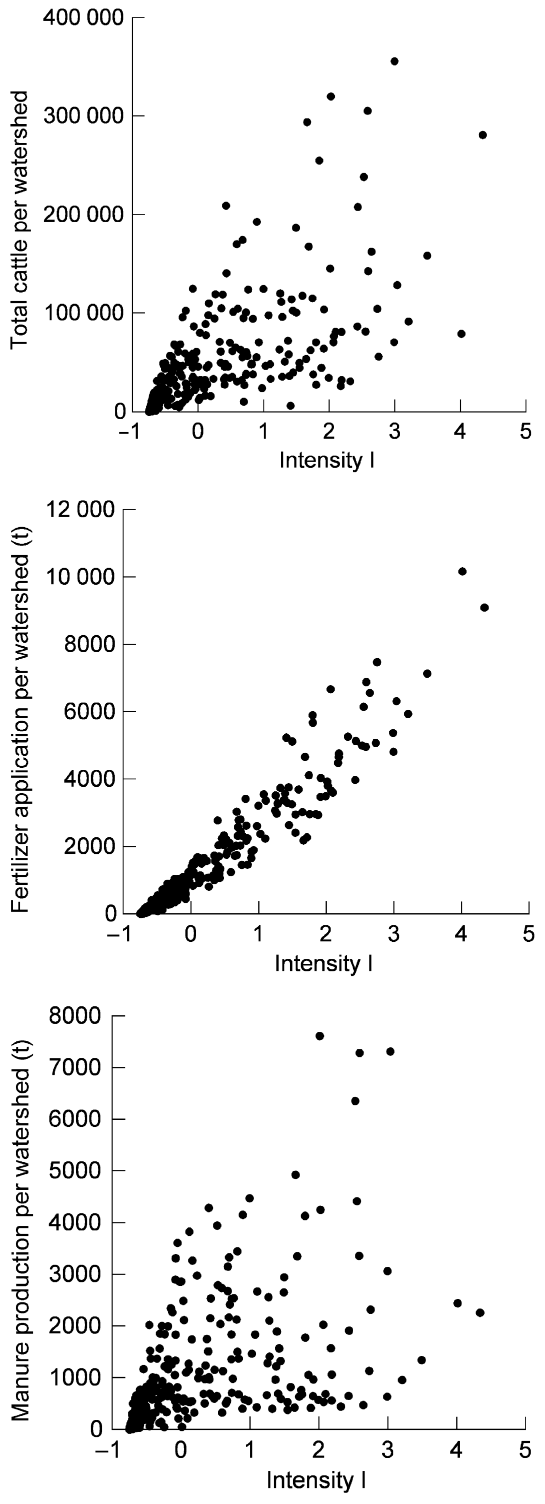

Fig. 6 The relationships between land use intensity (INTENSITY-I) and three of the variables used to generate intensity by PCA. All measurements here are measured per sampling area (watershed) from the Census of Agriculture.

Table 2 The component loadings of the Census of Agriculture measurements of agricultural inputs and by-product outputs on agricultural land use intensity indices from PCA

Agricultural input/by-product output	INTENSITY-I	INTENSITY-II
Total cattle per watershed	0.714	–0.517
Total fertilizer purchases ($)	0.972	< 0.001
Total fertilizer applied (tonnes)	0.966	–0.019
Total chemical purchases ($)	0.964	0.189
Area sprayed for insects (ha)	0.764	0.478
Area sprayed for weeds (ha)	0.880	0.296
Total manure production (kg)	0.590	–0.713

grazing lands found in this region. This central prairie area has high middle IR reflectance relative to the NIR and red bands, respectively, in the July–August composite period. Low MIR reflectance is characteristic of xeric conditions (de Boer, 1993; Fraser *et al*., 2000) and low moisture content in vegetation, factors that lead to low intensity agricultural practices relative to moister areas nearby.

Product validation

The area of agriculture per watershed determined from the VGT land cover classification was related strongly to census estimates of the extent of agriculture ($R^2 = 0.78$, $P << 10^{-6}$, $n = 336$; Fig. 3). However, the slope of the relationship is 0.741, indicating a consistent tendency for the remotely sensed measurement of agricultural area to underestimate the census reported value. Refinements to this work might alleviate this problem through reliance on finer resolution imagery (e.g. Landsat 7 ETM +). Higher-resolution imagery would facilitate detection of small patches of agricultural land that are not discerned easily using coarse resolution data sources. Alternatively, census data may overestimate actual agricultural area in each sampling district. Neither census nor remotely sensed data provide a completely accurate assessment of agricultural area, but their strong correlation is promising.

Further validation of the final land use classification is difficult because alternative data sources that are often marshalled for this purpose (such as TM scenes) provide data on cover rather than use. As with previous studies (Baban & Luke, 2000; Hurtt *et al*., 2001), we estimate the accuracy of our classification by assuming that the census data depict actual land use and by comparing LUCIA results to the census data. In general, land use derived from LUCIA relates quite well with results expected based on the Census of Agriculture. Estimates of pasture extents at the watershed level made by the LUCIA use/cover product and the census data correlate well ($R^2 = 0.716$, $P << 10^{-6}$, $n = 336$; Fig. 7). We controlled for the possibility that area might create spuriously strong

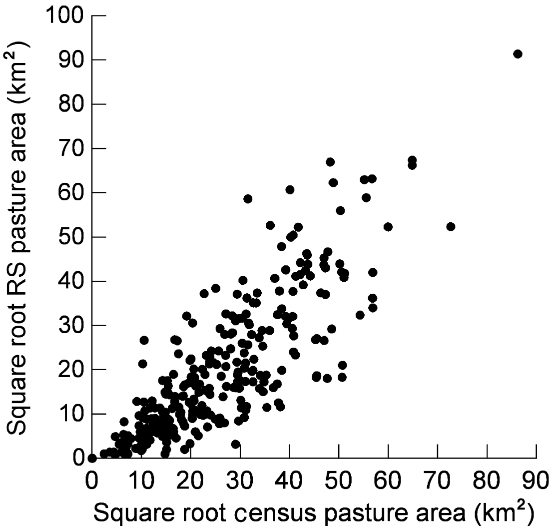

Fig. 7 The relationship between pasture area per watershed as estimated by Census of Agriculture and the land use data derived from the LUCIA process (in km^2). The strong correlation between these measurements ($r^2 = 0.716$, $P << 10^{-6}$, $n = 336$) indicates that the land use output of LUCIA is reasonably consistent with Census of Agriculture data.

correlations among these data by including area as a covariate in a second regression model (for example, watersheds with large agricultural areas will tend to have larger pasture areas). Agricultural area per watershed improved the pasture regression model ($R^2 = 0.817$, $P << 10^{-6}$, $n = 336$). In this model, both independent variables retain a high degree of significance. The partial coefficient (the slope of the relationship after adjusting for watershed agricultural area) of the remote sensing estimate of pasture extent is 0.484: the classification underestimates pasture extent. There may be considerable pasture area within mixed agriculture-woodland classes or other classes that are dominated by different land uses.

More detailed, field-based validation work would improve our estimates of land use in Canada. There are few alternative sources of land use data that could be used to generate an error matrix or index of agreement. Current land cover products rarely include classes that may be translated directly into land use. For example, agriculture is usually classified according to biomass estimates rather than crop types or other land use-related cover. This problem is serious and renders most land cover databases inappropriate for validation purposes except a general assessment of the agreement between estimates of agricultural extent.

Future directions

Data from sensors with greater spatial resolution, such as the Moderate Resolution Imaging Spectroradiometer (MODIS), may be useful in the future for land use development. A major benefit of working with such imagery, particularly at its 250-m resolution, would be the reduction of the number of mixed pixel classes: MODIS makes measurements that are closer to the scale of actual land use. Landsat 7 ETM + data would, of course, be better yet, but a national-scale TM mosaic for Canada will not be available for several years and is subject to serious technical challenges (e.g. mosaicking cloud-free TM scenes for Canada requires sampling from different parts of the growing season). A serious obstacle that remains for land use research at fine resolutions and very broad geographical extents is the scale/resolution mismatch between remote sensing and ancillary data sources. While the Canadian Census of Agriculture provides useful guidance in land use research, it is three to four orders of magnitude coarser in resolution that the spectral data available even from the 1-km resolution VGT sensor (and around eight order of magnitude less resolved than TM imagery). Comprehensive validation procedures will remain elusive in the absence of detailed local data on land use. The development of such data should be a high priority for future land use research in Canada.

The land use and land use intensity data that we have developed using the LUCIA process should be useful in a range of applications, such as endangered species research (Kerr & Cihlar, in press), environmental observation for sustainable development (Cihlar *et al.*, 2003) and carbon flux modelling. These data also provide baseline measurements of land use and land use intensity in Canada for land use (required under Framework Convention on Climate Change (FCCC) rules) and land use intensity change detection, respectively. The LUCIA process itself forms a basis for further land use research using data from VEGETATION, VEGETATION2 or higher resolution sensors (e.g. MODIS), provided sufficiently detailed ancillary data are available.

ACKNOWLEDGMENTS

Claire Gosson (Natural Resources Canada), Vince Zelazny (New Brunswick Ministry of Natural Resources) and Hussein Alidina (WWF) provided data for protected areas in Canada. Statistics Canada's Census of Agriculture data are courtesy of Nancy Hofmann. Robert Fraser provided valuable comments that improved this manuscript while Rasim Latifovic provided software and scientific tools to develop the SPOT4/Vegetation image composites for all of Canada.

REFERENCES

Anderson, J.R., Hardy, E.E., Roach, J.T. & Witmer, R.E. (1976) *A land use and land cover classification for use with remote sensor data.* U.S. Geological Survey Professional Paper 964. U.S. Printing Office, Washington, D.C.

Baban, S.M.J. & Luke, C. (2000) Mapping agricultural land use using retrospective ground referenced data, satellite sensor imagery and GIS. *International Journal of Remote Sensing*, **21**, 1757–1762.

Beaubien, J., Cihlar, J., Simard, G. & Latifovic, R. (1999) Land cover from multiple thematic mapper scenes using a new enhancement-classification methodology. *Journal of Geophysical Research*, **104**, 909–920.

Blais, J.M., Schindler, D.W., Muir, D.C.G., Donald, D.B. & Rosenberg, B. (1998) Accumulation of persistent organochlorine compounds in mountains of western Canada. *Nature*, **395**, 585–588.

Bromley, M. (1992) *Updated status report on the tundra peregrine falcon,* Falco peregrinus tundrius*, in Canada.* Committee on the Status of Endangered Wildlife in Canada, Ottawa.

CEC (1993) *CORINE land cover technical guide.* European Union, Directorate-General Environment, Nuclear Safety and Civil Protection, Office for Official Publications of the European Communities, Luxembourg, ISSN 1018-5593, ISBN 92-826-2578-8.

Cihlar, J., Beaubien, J. & Latifovic, R. (2001) *Land cover of Canada 1998. Digital data set documentation.* Natural Resources Canada, Ottawa.

Cihlar, J., Guindon, B., Beaubien, J., Latifovic, R., Peddle, D., Wulder, M., Fernandes, R. & Kerr, J. (2003) From need to product: a methodology for completing a land cover map of Canada with Landsat data. *Canadian Journal of Remote Sensing*, in press.

Cihlar, J. & Jansen, L.J.M. (2001) From land cover to land use: a methodology for efficient land use mapping over large areas. *Professional Geographer*, **53**, 275–289.

Cihlar, J., Latifovic, R. & Beaubien, J. (2000) A comparison of clustering strategies for unsupervised classification. *Canadian Journal of Remote Sensing*, **26**, 446–454.

Cihlar, J., Latifovic, R., Chen, J., Du Trshchenko, A.Y., Fedosejevs, G. & Guindon, B. (2003) Systematic corrections of AVHRR image composites for temporal studies. *Remote Sensing of Environment*, in press.

Cihlar, J., Ly, H., Li, Z., Chen, J., Pokrant, H. & Huang. F. (1997) Multitemporal, multichannel AVHRR data sets for land biosphere studies: artifacts and corrections. *Remote Sensing of Environment*, **60**, 35–57.

De Boer, T.A. (1993) Botanical characteristics of vegetation and their influence on remote sensing. *Land observation by remote sensing* (ed. by H.J. Buiten and J.G.P.W. Clevers), pp. 89–103. Gordon and Breach Science Publishers, Langhorne.

Environmental Systems Research Institute (2000) *Arc/Info, v. 8.02*. Environmental Systems Research Institute, Redlands.

Fraser, R.H., Li, Z. & Cihlar, J. (2000) Hotspot and NDVI differencing synergy (HANDS): a new technique for burned area mapping over boreal forest. *Remote Sensing of Environment*, **74**, 362–376.

Frolking, S., Xiao, X., Zhuang, Y., Salas, W. & Li, C. (1999) Agricultural land-use in China: a comparison of area estimates from ground-based census and satellite-borne remote sensing. *Global Ecology and Biogeography*, **8**, 407–416.

Government of Canada (1997) *State of the parks report.* Government of Canada, Ottawa.

Government of Ontario (1999) *Algonquin provincial park management plan.* Government of Ontario, Toronto.

Hurtt, G.C., Rosentrater, L., Frolking, S. & Moore, B. (2001) Linking remote-sensing estimates of land cover and census statistics on land use to produce maps of land use of the conterminous United States. *Global Biogeochemical Cycles*, **15**, 673–685.

Ji, C.Y., Liu, Q., Sun, D., Wang, S., Lin, P. & Li, X. (2001) Monitoring urban expansion with remote sensing in China. *International Journal of Remote Sensing*, **22**, 1441–1455.

Kelce, W.R. & Wilson, E.M. (1997) Environmental anti-androgens: Developmental effects, molecular mechanisms, and clinical implications. *Journal of Molecular Medicine*, **75**, 198–207.

Kelly, B.C. & Gobas, F.A.P.C. (2001) Bioaccumulation of persistent organic pollutants in lichen–caribou–wolf food chains of Canada's Central and Western Arctic. *Environmental Science and Technology*, **35**, 325–334.

Kent, M. & Coker, P. (1992) *Vegetation description and analysis*. CRC Press, Ann Arbor.

Kerr, J.T. & Cihlar, J. (in press) Land use and biodiversity: patterns and causes of species endangerment in Canada, Ecological Applications.

Kerr, J.T. & Packer, L. (1997) Habitat heterogeneity as a determinant of mammal species richness in high energy regions. *Nature*, **385**, 252–254.

Kerr, J.T., Southwood, T.R.E. & Cihlar, J. (2001) Remotely sensed habitat diversity predicts butterfly species richness and community similarity in Canada. *Proceedings of the National Academy of Sciences*, **98**, 11365–11370.

Martin, M. (1978) *Status report on the peregrine falcon,* Falco peregrinus anatum*, in Canada*. Committee on the Status of Endangered Wildlife in Canada, Ottawa.

Medema, G.J., Van Asperen, I.A. & Havelaar, A.H. (1997) Assessment of the exposure of swimmers to microbiological contaminants in fresh waters. *Water Science and Technology*, **35**, 157–163.

Mucher, C.A., Steinnocher, K.T., Kressler, F.P. & Heunks, C. (2000) Land cover characterization and change detection for environmental monitoring of pan-Europe. *International Journal of Remote Sensing*, **6**, 7, 1159–1181.

Nielsen, N.O. (1999) The meaning of health. *Ecosystem Health*, **5**, 65–66.

PCI Geomatics (2000) *PCI Imageworks and EASI*. PCI Geomatics Ltd, Richmond Hill.

Safe, S.H. (2000) Endocrine disruptors and human health: is there a problem? An update. *Environmental Health Perspectives*, **108**, 487–493.

Schindler, D.W. (1974) Eutrophication and recovery in experimental lakes: implications for lake management. *Science*, **184**, 897–899.

Sonnenschein, C. & Soto, A.M. (1998) An updated review of environmental estrogen and adrogen mimics and antagonists. *Journal of Steroid Biochemistry and Molecular Biology*, **65**, 143–150.

Soulé, M.E. & Terborgh, J. (1999) Conserving nature at regional and continental scales: a scientific program for North America. *Bioscience*, **49**, 809–815.

SPSS Software Ltd (1999) *Systat, v. 10.04.* SPSS Inc., Chicago.

State of the Parks Report (1997) Government of Canada, Ottawa.

Statistics Canada (1996) *Census of agriculture.* Government of Canada, Ottawa.

Thorpe, J. & Goodwin, B. (1997) Range use by cattle and deer in the Great Sand Hills. *Grazing Gazette* 6.

Townshend, J.R.G., Justice, C.O. & Kalb, V.T. (1987) Characterization and classification of South American land cover types using satellite data. *International Journal of Remote Sensing*, **8**, 1189–1207.

BIOSKETCHES

Jeremy Kerr is an assistant professor of biology at the University of Ottawa and is interested in the relationships among human activities, environmental gradients and aspects of both biological diversity and community structure. He uses remote sensing, GIS tools and field data to address research issues at different spatial scales.

Josef Cihlar is a senior research scientist at the Canada Centre for Remote Sensing. His primary interests are in the use of earth observations for improved understanding of environmental processes and change, particularly in relation to land cover characteristics and the carbon cycle.