The Canada Vegetation and Land Cover: A Raster and Vector Data Set for GIS Applications - Uses in Agriculture

Stefan Palko and Louiselle St-Laurent
GIS Division
Geomatics Canada
Natural Resources Canada
615 Booth Street, Ottawa, K1A 0E9, Canada
Tel: 613-996-2812
Fax.: 613-952-0916
E-mail: palko@cc2smtp.emr.ca, Louiselle.St-Laurent@geocan.nrcan.gc.ca

Ted Huffman and Eric Unrau
Research Branch Agriculture and Agri-Food Canada
Central Experimental Farm
Ottawa, Ontario
Tel: 613-957-4347, x7837
Fax: 613-996-0646
E-mail: huffmant@ncccot.agr.ca

Abstract

This paper is a follow-up to the presentation entitled "Canada's New, Seamless Forest Cover Data Base" given at this conference in 1993. The final stages concerning the preparation of a complete vegetation and land cover data set for Canada are described, transformation of the raster product to vector format, integration into Geographic Information Systems (GIS) and limitations in GIS applications.

The data set was derived from the Advanced Very High Resolution Radiometer (AVHRR) sensor operating on board the United States National Oceanic and Atmospheric Administration (NOAA) satellites which have a resolution of 1.1 km. The vegetation and land cover data set includes twelve categories, all interpreted from the AVHRR imagery. Focus of this paper is on the vegetation and land cover classification, raster to vector conversion of the data to make the set more suitable for GIS integration, generalization for cartographic presentations and an example of its use in re-processing Census of Agriculture data to a biophysical (Soil Landscapes of Canada) framework for soil quality evaluation.

1.0 Introduction

As the use of geographic information systems is expanding, the availability of timely and up-to-date spatial data in digital format is an essential requirement for its success. For the user, it is a requirement which is expected to be easily fulfilled. Satellite imagery combined with the increased processing capabilities of current image analysis systems have made it possible to generate meaningful data sets which represent new knowledge not available with previous technologies. NOAA AVHRR data with 1.1 km resolution was used to produce such a data set, the Canadian Vegetation and Land Cover.

The beginning of this work was described at the 1993 GIS Conference (Palko et al, 1993). Since that time, the data set has been completed and published as the vegetation cover map included in the 5th edition of the National Atlas of Canada (EMR Canada, 1993). The raster data set has been transformed into a GIS-compatible vector format and has been used in several applications, particularly with GIS. Based on these recent applications, it has been possible to evaluate the data set, identify its advantages and limitations, as well as its desired characteristics in future updated versions. Evaluation in this paper employs applications of the data set in re-processing the Census of Agriculture data to a biophysical framework for soil quality evaluation (Soil Landscapes of Canada). The choice of this evaluation was made because it illustrates a typical example of using products derived from satellite imagery for GIS applications.

2.0 Generation of the Vegetation and Land Cover

The vegetation AND land cover data set was generated from NOAA AVHRR imagery (Palko et al, 1993). Approximately 45 images spanning several years (summer coverage, 1988-1991) were selected to produce a cloud-free composite for the entire country suitable for classification and interpretation. More specifically, the composite was derived from a combination of the red and near-infrared channels. The image was classified using a combination of supervised, automated, "maximum likelihood", and manual classifications.

Classified images were then subjected to filtering in order to reduce visual noise. In general, a minimum area of 4 contiguous pixels was retained. (Exceptions were isolated patches where single pixels of a class were retained if they were located more than 7 pixels inside the boundary of a larger polygon of a different class). Manual editing was used to supplement automated techniques and to incorporate changes from the review and verification process. The final data set was assembled digitally by merging the interpreted image data in raster format with selected National Atlas base map components in vector format.

2.1 Vegetation and Land Cover Classes

The cover classes included in this data set were selected after careful consideration of the image analysis and classification challenges inherent in mapping such a diverse and spatially extensive area. It was not feasible to differentiate wetland areas. In addition, vegetation zonation in the Arctic is poorly represented, with the result that areas dominated by low erect shrubs, by dwarfed and prostrate shrubs, and by herbs appear in the same class (Edlund 1990). Moreover, reflectance values were not sufficiently distinct to differentiate totally unvegetated terrain (barren land) from a sparse vegetation cover of woody plants, herbs and non vascular plants (lichens and mosses). Even some nearly continuously vegetated terrain dominated by lichen communities was classified in the same group. Although the approach was successful in delineating principal areas of agricultural lands, many dispersed areas were not identified. It is expected that further refinement of the classification techniques used will result in improved interpretation accuracy. The horizontal positional accuracy of digital data achieved is 1-4 km. Figure 1 illustrates the distribution of vegetation and land cover categories over the Canadian land mass.

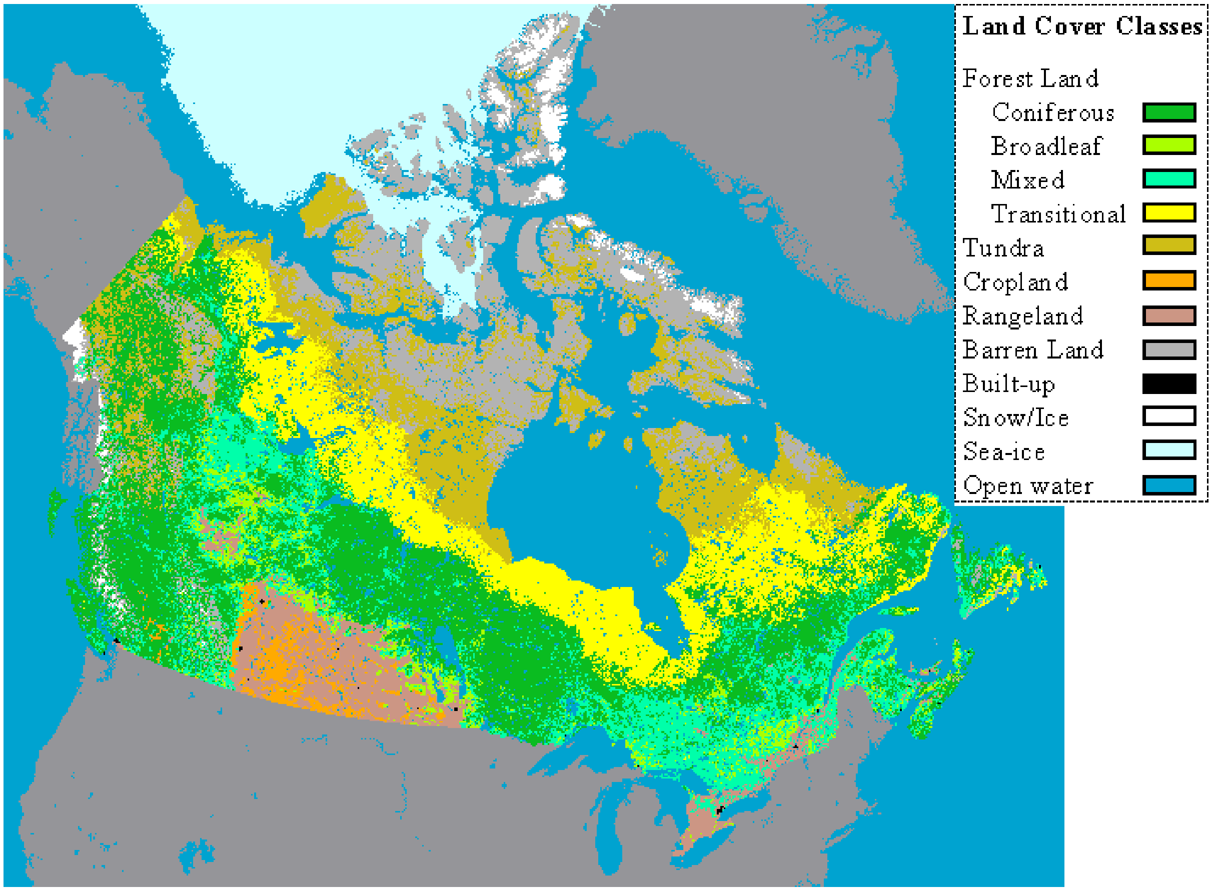


Figure 1. Distribution of vegetation and land cover categories over Canada.

The following list (Table 1) gives a brief description, classification and definition for each category.

Predominant Plant Cover	Class	Definition
	Forest land	Land on which trees are the dominant vegetative cover with tree crown cover of 10% or more. Includes land where trees are stunted owing to site limitations, undetectable owing disturbance, or temporarily absent.
	Continuous Forest	Land cover type where forest land occupies more than 50% of the area.
Tree	• Coniferous Forest	Continuous forest in which 76-100% of the canopy is composed of coniferous trees.
	• Broadleaf Forest	Continuous forest in which 76-100% of the canopy is composed of broadleaf trees.
	• Mixed Forest	Continuous forest in which 26-75% of the canopy is composed of coniferous or broadleaf trees.
Tree/Shrub/Herb/ Nonvascular*	• Transitional Forest	A mixture of land cover classes where tree cover is discernible but forest land occupies less than 50% of the area.
Shrub/Herb/ Nonvascular*	• Tundra	Low arctic or alpine vegetation with discernible cover. Although generally located beyond the tree line, low woody plants (ericaceous shrubs, willows, etc.) and patches of stunted trees may occur.
	Agricultural Land	
Crops	• Cropland	Cultivated land with crops, fallow, feedlots, orchards, vineyards nurseries, shelter belts, and hedgerows.
Non-Cultivated (Shrub/Herb)	• Rangeland and Pasture	Land supporting native vegetation, shrubs, grass and other herbaceous cover with less than 10% tree cover. Includes improved land dedicated to the production of forage, and upland and lowland meadows.
	Non-vegetated Land	
	• Perennial Snow or Ice	Perennial snow fields and glaciers.
	• Barren Land	Land without discernible vegetation cover. May include sand, rock, bare soil and open pit mines.
Non-vegetated	• Built-up Area	Cities and towns of sufficient size to be depicted at the scale of mapping.
	Water/Sea Ice	
	• Open water	
	• Sea-ice Minimum Cover	

*Plants lacking an internal vascular system (e.g. moss, lichen)

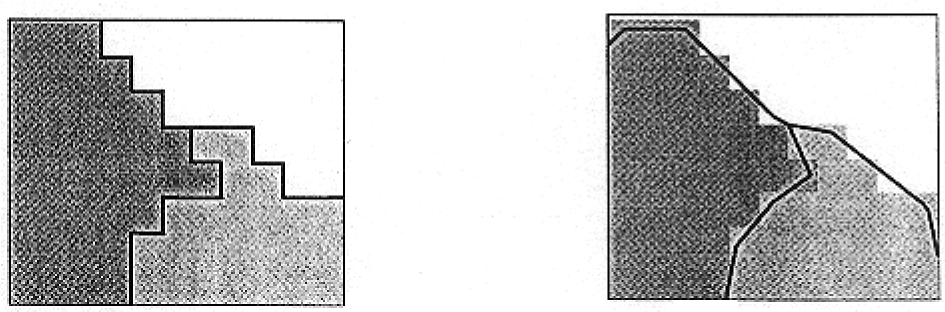
Table 1. Description of the vegetation and land cover classes.

3.0 Raster to Vector Conversion and Generalization

The necessity of the conversion of the Canadian vegetation cover into a vector product was justified for full integration of remote sensing into GIS. The capabilities of GIS software to use remote sensing products have progressed in the recent years, from simple display of a remote sensing image in the background to application of computations between the remote sensing product and the GIS data base information. Computations between the images and the GIS data base are based on the capabilities of the GIS software to handle both the raster and the vector domains. Currently, existing GIS software does not allow direct computations between vector and raster structure and the majority of GIS software acquired by users are restricted to analysis in one of the domains only. Commonly, raster-based GIS are preferred for applications using remote sensing as the main source of variable data. However, for vector-based software, full integration becomes possible only when remote sensing products are vectorized. To make the vegetation cover data set available to vector-based system, the raster data set was converted into a vector product.

3.1 Raster-to-Vector Conversion

Raster-vector-conversion is a classical problem in graphic technologies from and many programs, including semi-automated vectorization software packages, have been developed as a result. For area maps or thematic images, the conversion is usually done by generating a line between adjacent pixels of different values which represent the boundary between two raster areas. The coordinates of this line can be generated either by taking the position of the corner of each boundary pixel (Figure 2a), which results in a line with horizontal and vertical segments (the staircase problem) or they can be approximated to a position which represents the tendency taken by surrounding boundary pixels. The latter approach results in a smoother representation (Figure 2b).



(a) (b)

Figure 2. Raster-vector conversion.
a) Staircase effect.
b) Smoothing with two-pixel look-ahead vectorization

The generated polylines are then structured into polygon topology and coded with the thematic values of the raster. From the raster-to-vector conversion, the data set can increase in size with increased complexity of the thematic data. The algorithm has been implemented on some commercial GIS software packages with limitations as to the capabilities of the software to handle the large data sets. The methodology used to transform the vegetation raster data set uses a two-pixel look ahead algorithm where the resulting polyline passes on the "correct" side of the pixel centre (EPS Technologies Inc.). In the conversion of the raster vegetation data set, the full power of available high-end software was exploited, The conversion aimed at reproducing a vector data set which would keep the integral spatial distribution of the original raster classes. The procedure consisted in the following steps:

A. Automatic transformation of the raster into smooth vectors.
B. Interactive editing to correct long thin polygons.
C. Polygon topology building with creation of point in polygons.
D. Overlays of vector-polygons over the original raster cover for attribute coding the polygons with the vegetation class values.
E. Verification of the coded polygons by transformation of the vector set back to the raster representation, followed by a difference calculation between the vegetation class codes of the generated and the original raster version.

In the automatic transformation, step A, features represented by one pixel-wide areas would result in discontinuous polygons where these pixels would be arranged in a diagonal fashion and touched only at the corners. In the cases of features such as rivers and coastal inlets, these polygons were connected and validated by comparison with existing digital hydrography. The verification process (step E) demonstrated the integrity of the vector product in representing the original raster cover.

The result of the raster-to-vector conversion increased the size of the vegetation data set by a factor of four, to approximately 100 megabytes, which is typically too large for most users to handle. However, in the process, the data set was broken down into portions based on a tiling schema using the political boundaries of the country. The 1:2 000 000 scale (political boundaries and hydrography) NAIS digital data was used for the break down. Each portion of the set was edge-matched at these boundaries to permit the regeneration of a seamless cover.

Through the transformation process, it was possible to derive an evaluation of the vegetation cover product based on the existing hydrography and coastline at the 1:2 000 000 scale. Comparison of the two sets showed an overall relative accuracy of 1 km. However on the edge of the set, on the eastern and western coasts, differences up to 4 km are observed, the far edges being on the far reach of the initial NOAA imagery.

3.2 Generalization

The large size of the generated vector product limits the use of the data set to portions for analytical GIS applications on a continental scale. For cartographic representations at a very small scale (1:7 500 000), the quantity of data is too large for existing software to handle as a whole. Generalization becomes a necessity. Using existing GIS software package functionalities, the vector product was generalized to a smaller scale through simple elimination of polygons on the basis of polygon area size. For a scale of 1:7 500 000, eliminating all areas smaller than 50 square kilometers reduces the volume of the data set both in size and in the number of polygons by a factor of 10. This simple approach still respects the relative distribution of the thematic categories for cartographic representation and results in a set which can be handled as one coverage for the country.

4.0 National Applications in Agriculture

The Canadian vegetation vector product has proven useful to demonstrate the integration of remote sensing in GIS and to extend the application domain of the cover. In one particular application, the vector map was used by the Centre for Land and Biological Resources Research (CLBRR) of Agriculture and Agri-food Canada as a 'filter' during the integration of Statistics Canada's "Census of Agriculture" data (Statistics Canada,1992) with Soil Landscapes of Canada (SLC) (Shields et al,1991) polygons.

CLBRR is currently involved in developing and reporting on 'agri-environmental indicators' pertaining specifically to resource (soil) quality (McRae et al, 1994). Several of these indicators, such as soil cover and land management, find considerably useful data in the Census, but the smallest reporting unit is the Enumeration Area (EA). An EA typically encompasses between 20 and 100 farms, and due to confidentiality provisions, data is not generally segregated below that level. Also, EA boundaries typically conform to those of cultural units such as municipalities and have no relationship to soil or vegetation types.

In order to 'reorganize' the Census to the spatial stratification of SLC's, we intersected EA polygons with SLC polygons and apportioned data according to area. However, to apportion on the basis of total area, without regard for the spatial distribution of agricultural land within and between SLC's was considered inappropriate. The vectorized Vegetation and Land Cover map was used to determine the portions of EA's that intersect with agricultural portions of SLC's.

5.0 Limitations of the Vegetation Cover Map

The vectorized cover map was particularly helpful in apportioning EA data along the boundary between the 'agricultural' and the 'forested' portions of the country (i.e. the agricultural 'fringe'). In these areas, SLC's tend to reflect the 'margin' effect by having a clear separation into agricultural and non-agricultural portions. The farm data from an EA that straddles this area was appropriately apportioned to SLC's on the basis of agricultural area. Although this was the case in all areas of the fringe, it was especially important in districts such as the Peace River of Alberta and British Columbia, the interlakes of Manitoba and the St. John River valley in New Brunswick.

Limitations in the use of the vegetation cover map became apparent, however, within the large areas designated 'forest' or 'agricultural'. In the case of the forest area, our EA database showed the existence of considerable numbers of farms and agricultural land in areas where the cover map indicated no agricultural cover. This was particularly noticeable in the northern interior of British Columbia, the clay belt area of Ontario and Quebec, the Eastern Townships and St. Lawrence Valley in Quebec and the Fredericton area in New Brunswick (Figure 3).

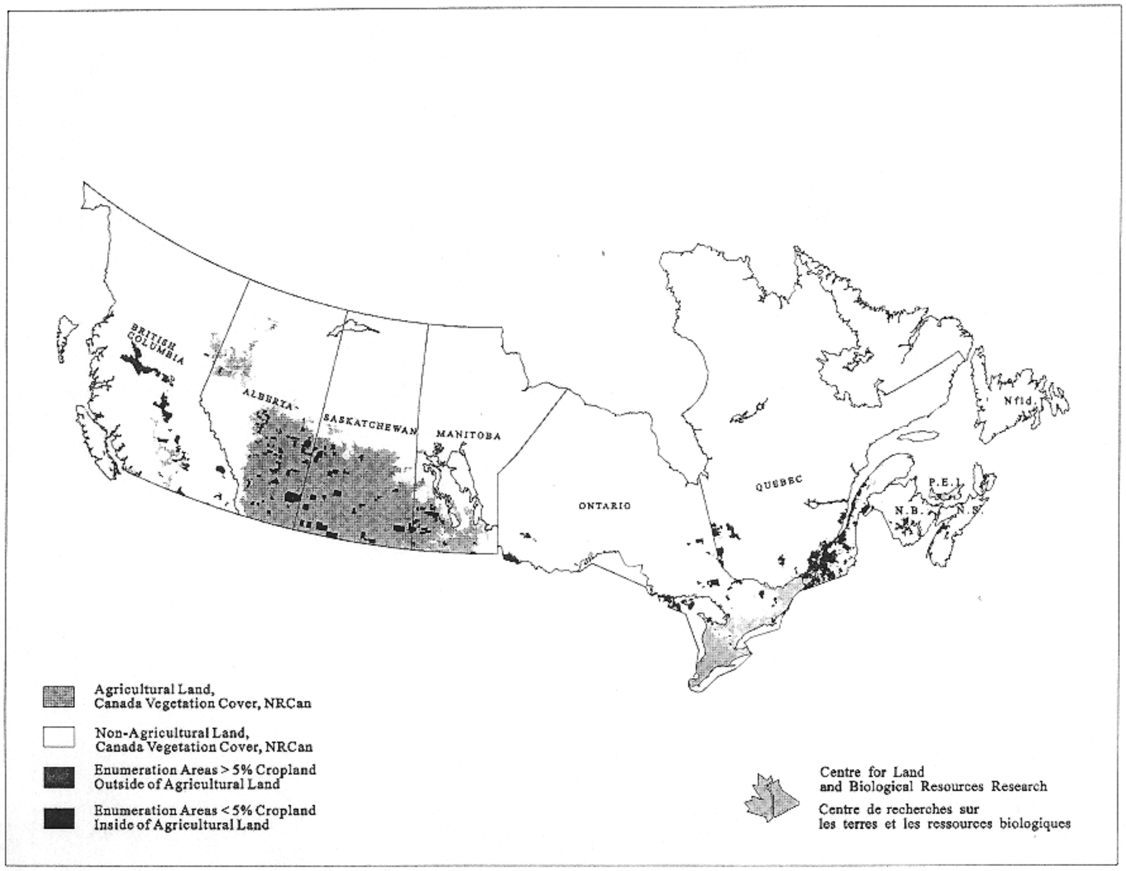


Fig. 3 Distribution of Agricultural Lands in Canada

Similarly, a number of cases of EA's with low proportions of agricultural land were identified within large areas of homogeneous agricultural cover (Figure 3). These occured primarily in the prairies near Milk River and Brooks in Alberta and in the 'Badlands' of Southern Saskatchewan.

The implication of these limitations is twofold; we attempted to use the vegetation cover map at a scale beyond its resolution capabilities or we found classification inconsistencies. We suggest that both are in play; that ommission (from an agricultural perspective) errors are mainly the result of farmland existing in small patches within a predominantly forested landscape, and commission errors are the result of short-grass prairie that is not farmland being classified as 'rangeland'. While the former is a technical issue for which we make recommendations in the following section, the latter is a case of semantics in which the term 'rangeland and pasture' is presented as a land cover term when it more properly denotes a land use.

6.0 Considerations for the Data Set Use and Future Updates

We can make several recommendations with respect to the limitations we encountered. The first is that the map (especially the electronic version) be accompanied by an explicit note with respect to spatial resolution, rather than relying on an overview of the source and methodology to serve that purpose. Similarly, the inconsistency in nomenclature of grassland should be noted.

Our second comment relates to future updates and is a suggestion that a 'mixed pixel' category be reviewed in order to determine the possibility of identifying areas of complex mixes of agriculture and forest.

Thirdly, with respect to a 'national vegetation cover' map, we suggest that our use and analysis can serve to identify areas where a different source of imagery might be used. By identifying 'target' areas of spatial mixes, it becomes economically realistic to consider the use of imagery such as Landsat TM or SPOT to update the AVHRR derived product.

Finally, the multi-year coverage (i.e. 1988 to 1991 summer imagery) was a consequence of the availability of data at the time of production of the data set. When this project was initiated in 1988, Canada was covered only partially by the NOAA AVHRR imagery; much of the Yukon and the Northwest Territories was not covered. Consequently, while arrangements wer being made to include the missing northern coverage during the summer of 1990 and 1991, the project was proceeding to map the rest of Canada from the existing pre-1990 coverage. Experience gained during the project indicates that it is quite possible to generate similar data sets from one summer period. In addition, processing capabilities have improved to the extent that producing a new data set on a short cycle is a realistic possibility.

7.0 Summary and Conclusion

The Canadian vegetation and land cover mapping project provided a good opportunity to experiment with the application of leading technologies in remote sensing and GIS. In mapping the vegetation and land cover categories, the NOAA AVHRR imagery permitted a new generation of data of sub-continental scale to be developed directly from imagery acquired from space platforms. It proved to be a viable alternative to the previous data sources, such as large-scale aerial photography, LANDSAT imagery, or secondary sources based on ground observations. In merging the NOAA AVHRR product with GIS, the project demonstrated practical solutions for converting raster-based RS sources into vector-based inputs into GIS.

While the integration of remote sensing into GIS applies generally to both small scale and large scale imagery (i.e., NOAA AVHRR and LANDSAT), methodologies involved in delineation of cover categories are specific to the NOAA AVHRR imagery.

Advantages of the image compositting, analysis and classification methodologies are in the use of a consistent data set of large aerial extent and short duration, a consistent analytical procedure, and the relatively short time period needed to generate the data set. Existing applications, such as described in this paper, have clearly demonstrated the usefulness of the data set and have helped to identify areas for improvment in future updates. Potential applications are evident for global environmental monitoring, for resource management, and for the creation of various continental and sub-continental perspectives. Limitations of this approach include the level of accuracy of the data (maximum cellular resolution of 1 km), and the limited number of cover categories that were identified.

8.0 References

Canada, Energy, Mines and Resources Canada. 1993. Canada-Vegetation Cover. *The National Atlas of Canada*, 5th edition. Map Scale 1:7 500 000. Ottawa.

Canada, Environment Canada. 1990. *The Canadian Vegetation Classification System.* First Approximation. National Vegetation Working Group, Canada Committee on Ecological Land Classification. Edited by W. L. Strong, E. T. Oswald, and D. J. Downing. Ecological Land Classification series, No. 25. Ottawa.

EPS Technologies Inc, 1992, Conversion of Surface to Vector, *PAMAP GIS Software Manual,* Victoria.

McRae,T., T. Huffman, J. Culley, D. Trant, J. Dumanski, B. Harron, C. de Kimpe. 1994. Using indicators to assess the health of Canadian agroecosystems. *Poster, First International Conference on Agroecosystem Health*, Ottawa, Canada.

Palko, S., Lowe, J. J., Pokrant, H. T. 1993. Canada's New Seamless Forest Cover Data Base. *Proceedings, GIS'93 Symposium*, Vancouver, February 1993.

Pokrant, H. T.; Palko, S.; and Lowe, J. J. 1991. The Use of Remote Sensing in Producing the National Atlas of Canada. *Proceedings, Ministry of Natural Resources and Canadian Institute of Surveying and Mapping GIS Seminar,* Toronto, November, 1991.

Shields, J.A., C.Tarnocai, K.W.G.Valentine and K.B.MacDonald.1991. *Soil Landscape of Canada - Procedures Manual and User's Handbook*. LRRC Contribution No.88-29. Agriculture Canada. Ottawa.

Statistics Canada. 1992. *1991 Census of Agriculture* - Products and Service Catalogue 92-303. Agriculture Division, Statistics Canada. Ottawa.