Methods for Deriving Canada Wide Geo-Spatial Datasets in Support of Environmental Monitoring and Modelling

Goran Pavlic[1], Richard Fernandes[2], Wenjun Chen[2], Robert Fraser[2], Sylvain G. Leblanc[2]
[1]Intermap Technologies Corporation
588 Booth Street
Ottawa, Canada, K1A 0Y7
[2]Canada Centre for Remote Sensing
588 Booth Street
Ottawa, Canada, K1A 0Y7

ABSTRACT

Satellite remote sensing and numerical modelling are important tools to quantify trends in the environment across Canada. This paper describes methodology and initial results for the production of a new 1-km resolution map of waterbody fraction and 10-km resolution maps of forest cover type (hardwood/softwood) required as input parameters to these approaches. The waterbody fraction map is derived by gridding and aggregating existing national topographic database waterbody entities. The forest cover type map is produced by blending through decision rules a large number of datasets: 1-km resolution land cover based on satellite remote sensing, 1-km digital elevation model, Soil Landscapes of Canada information, and an ecoregion level descriptions of species occurrence. The quality of the water fraction map and forest cover type map is assessed. Applications of both products to national and regional scale mapping and modelling efforts are discussed. Strategies for improving both products are also outlined.
Keywords: Water, forest cover, national, modelling, raster

1. INTRODUCTION

Geo-spatial models of physical and socio-economic systems are now commonly used to address basic research questions as well as pressing policy issues at regional and national levels. These models require consistent and continuous national coverage of fundamental survey information such as census, land cover, and forest inventories, represented with mapping units on the order of 100 km^2. In this paper we discuss the production, quality, and application of two continuous data layers for Canada at much finer spatial scales:

- Waterbody Fraction Coverage of Canada (Waterbody Coverage) representing the fraction of liquid surface fresh water mapped within each 1-km grid cell across Canada.
- Forest Cover Type Fraction Coverage of Canada (Forest Cover Coverage) representing the fraction of each broadleaf and needleaf forested area in each 10-km grid cell.

2. METHODOLOGY

Both national coverages are a synthesis of survey data acquired over a large time frame (1949 to 2001) and at varying levels of quality. These surveys were conducted by mandated agencies with considerable resources. Therefore, it was deemed unrealistic to perform additional direct survey within the current study. Rather, we adopt the criteria that the national coverage consist of exactly the same information provided by the national surveys but aggregated to a useful reporting scale and supplemented by additional estimates where persistent biases in the national surveys are evident. An emphasis is placed on consistency and clear lineage.

2.1. Waterbody Fraction Methodology

Vector coverages of water body polygons were acquired from the Canadian National Topographic Database Version 3.1 (NTDB, Anonymous 1996) produced by Natural Resources Canada. These coverages correspond to 7209 1:50,000 scale map sheets and 1169 1:250,000 scale map sheets tiling Canada as shown in Figure 1. These maps are in the process of continuous revision using a number of technologies including fine to medium scale air photos, orthophotos and satellite imagery. Waterbodies are defined in the NTDB as one of:

- Generic/Unknown - a region covered with water at time of survey;
- Intermittent/slough - a waterbody normally dry at some time of the year;
- Other - other regions designated as waterbodies;
- Irrigation canal - a manmade waterbody;
- Flooded area - permanently flooded region.

The only other NTDB categories that may contain water at the surface are 'tundra ponds' and 'wells'. We have not included these categories in the water fraction product. 'Tundra ponds' are classified as wetlands and will be included in future wetland product. Figure 1 provides a flowchart of the processing steps used to derive the waterbody coverage. More details are given in the metadata provided with the product.

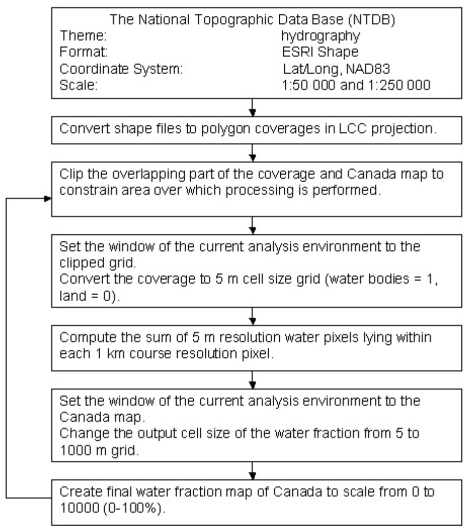

Figure 1: Flowchart of operations performed to generate Canada-wide water body coverage.

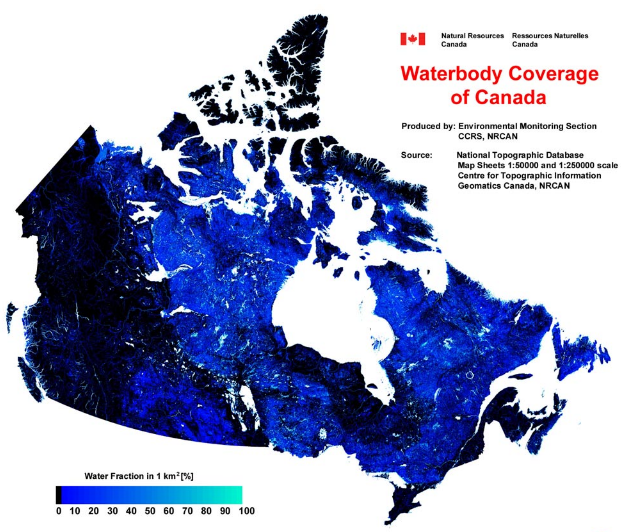

Figure 2: Water body fraction map of Canada together with inset showing detail. Note that only areas in white would be represented in raster maps with 1km grid cells using nominal presence absence information.

2.2 Forest Cover Type Methodology

The Canadian Forest Service compiles a National Forest Inventory nominally every five years. This inventory corresponds to a roll-up of provincial and territorial forest inventories. A publicly available spatial database of land type data, produced from Canada's Forest Inventory 1991, 1994 version (CanFI91.V94) by the Pacific Forestry Centre, Canadian Forest Service (Lowe et al., 1996) denoted 'CanFI' hereafter. Among other parameters, the database reports on the area of broadleaf and needleaf forest within mapping units ranging in size from 10 km to 100 km square (e.g. http://www.pfc.cfs.nrcan.gc.ca /monitoring/inventory/maps). The spatial distribution of the mapping units is reported in Gray and Power (1997). The vintage of the survey ranges from 1949 to 1994 with a mean date of 1979. Given concerns with vintage and the focus of CanFI on merchantable areas, additional input data were required to identify biases between CanFI estimates of forest cover and actual forest cover. A satellite based land cover map of Canada at 1-km resolution (Cihlar et al. 2001), adjusted by the waterbody fraction map, was used to estimate the actual amount of forest present for the 1998 growing season. Estimates of the fraction of needleleaf and broadleaf forest cover were also used to identify biases in CanFI sampling. However, these more detailed estimates were not used to adjust CanFI due to uncertainties in the mixed forest proportions (Cihlar et al. 2002). The CanFI forest cover type estimates were derived by a vector to raster conversion of the CanFI polygons into 10 km grid cells.

Independent estimates of forest cover type over forested areas were also derived based on narrative ecoregion descriptions and ancillary databases. The 217 ecoregions within the Ecological Framework of Canada (Ecological Stratification Working Group, 1996) was the primary means of constraining species estimates in the absence of CanFI data. These ecoregions contain, 1021 ecodistricts representing the smallest mapping unit within the Ecological Framework. Text attributes identify the co-variation of species with soil drainage class, elevation, and aspect. The Canadian Soil Information System (CANSIS, http://sis.agr.gc.ca/cansis/intro.html) was used to map soil drainage class according to the Soil Landscapes of Canada (Centre for Land and Biological Resources Research. 1996) mapping units. The soil drainage classes were aggregated into 3 levels to match the ecoregions descriptions (well drained, moderately drained and poorly drained). A 1-km resolution digital elevation model of Canada (Anonymous, 2000b) was used to derive elevation and aspect. Figure 3 provides a flowchart of processing steps applied to the input data to derive the ecoregion based estimates. The text attributes describing species occurrence were parsed resulting in a uniform set of rules designed to provide an estimate of forest cover type given the spatially geo-referenced input coverages.

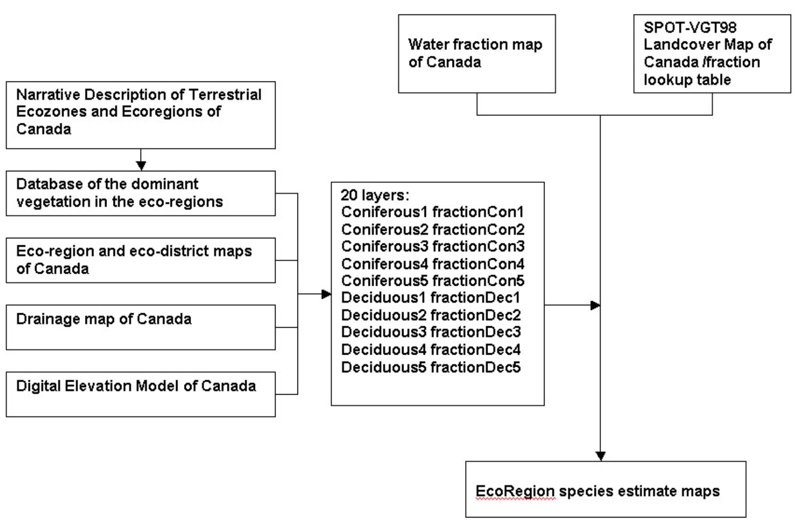

Figure 3: Flowchart of processing steps applied to derive national species cover type coverage from ecoregion level information.

CanFI cover estimates were adjusted in areas where they underestimated forest cover by comparison to the satellite land cover maps. A satellite based land cover map of Canada at 1-km resolution (Cihlar et al. 2001), adjusted by the waterbody fraction map, was used to estimate the actual amount of forest present for the 1998 growing season. Satellite-based estimates of the fraction of needleleaf and broadleaf forest cover were also used to identify biases in CanFI sampling. However, these more detailed estimates were not used to adjust CanFI due to uncertainties in the mixed forest proportions (Cihlar et al. 2002). Rather, adjustment was accomplished using species cover information from co-located ecoregions species maps. No correction for overestimates was performed both to avoid having to remove forests surveyed by CanFI and because the degree of overestimation was less then the uncertainty in relating the CanFI and the forest cover maps. Blending of CanFI and ecoregion maps were performed over 10-km grid cells as follows:

1. The CanFI forest cover underestimate if any relative to the land cover map was computed.
2. The ratio, R1, of underestimated forest cover to total ecoregions forest cover was computed.
3. The histogram of ecoregions forest cover type for the cell was scaled by R1. If R1 was infinity, then the CanFI estimate was completely replaced by the ecoregions estimate.
4. The CanFI cover was subtracted from the cover of the corresponding cover type represented in the scaled histogram. Negative cover values were set to zero to produce an adjusted histogram of forest cover type.
5. The adjusted histogram of ecoregions species cover was once again scaled so that the total cover matched the land cover based forest cover underestimate.
6. Forest cover type estimates from the scaled and adjusted histograms were added to the corresponding CanFI forest cover type layer.

This blending method forces agreement between CanFI and land cover estimates over the areas identified as having large biases. Also, the output will match CanFI if there is no overestimate and exactly match the ecoregions cover in cases where there is total underestimation (i.e. CanFI indicates no forest and land cover identifies a forest). It is likely that the CanFI underestimate does not occur over inventoried forests but over areas not considered merchantable that are described in an ecological framework of the region. Simply adding the ecoregions species distribution scaled to match the extent of CanFI underestimate will not account for this sampling bias. Step 4 reduces the likelihood of double counting of inventoried forest cover type.

3. QUALITY CONTROL

3.1 Waterbody Coverage

Two potential sources of error in the final map are: the differences between the NTDB vectors and reality and processing artefacts caused by converting the NTDB to a 1-km raster grid. The first error source is quantified in the NTDB documentation. Errors in producing the raster include:

- Numerical rounding errors: This source is negligible as double precision floating point computations are performed.
- Vector to raster conversion error: We performed a 1% sampling of both 1:50,000 and 1:250,000 map sheets focussed on areas with the most convoluted water body shapes (e.g. Boreal Shield, river deltas). The relative root mean square difference between vector polygon and raster gridded waterbody proportions for 1 km pixels in sampled NTDB map sheets was always below 3% and usually below 1% for 1:250,000 map sheets due to their coarser initial scale.
- Raster aggregation: The total water fraction error between 1km gridded waterbody estimates and the polygon-based area of the map sheet was quantified for all 1:250,000 map sheets. The mean bias for 1:250,000 map sheets was 0.1287% of the area over the map sheet (standard deviation of 0.1539%). A similar bias was found over the 1% sample of 1:50,000 map sheets with a slightly higher standard deviation of 0.183% due to the increased level of initial generalisation when converting to 5 m grids.

3.2 Forest Cover Type Coverage

Two quality control exercises were possible given the available data sets. Firstly, the ecoregions attribute parsing exercise was examined for consistency (i.e. would the text attributes be parsed into the same %cover distributions by different interpreters?). Two interpreters parsed the same random sample of 10 ecoregions. The root mean square difference was under 20% for any ecoregion and the bias was under 10% over the sample. Given that crown cover estimates from forest survey are typically generalized to 20% or coarser intervals, the parsing exercise is likely no less precise, thematically, then the CanFI databases. This is not to say that the ecoregion estimates are more precise than CanFI when mapping actual forest cover since they represent typical rather than actual forest cover.

The second quality control exercise was the comparison of CanFI needle leaf and broadleaf cover with the satellite based land cover estimates. The CanFI estimates of total broadleaf and needleleaf cover were compared to estimates derived by combining the land cover and water fraction maps. Histograms, shown in Figure 4, indicate no bias between CanFI and land cover estimates of broadleaf species and a rapid exponential decay in residuals. This is encouraging on two counts. First, broadleaf species are typically well inventoried as they are found in southern Canada and in upland areas where management is possible. Therefore, the good agreement is a marginal validation of the land cover product. Second, the comparison suggests that errors due to vintage and imputation of CanFI cells are small over mixed and pure broadleaf zones. The comparison of needleleaf species shows a similar behavior for areas where CanFI overestimates needleleaf cover. This suggests that the low errors due to imputation may apply to all surveyed regions. However, as expected, the land cover map frequently maps more needleleaf cover than CanFI. This is most likely due to the surveys being biased towards merchantable stands – especially since the broadleaf estimates match relatively closely.

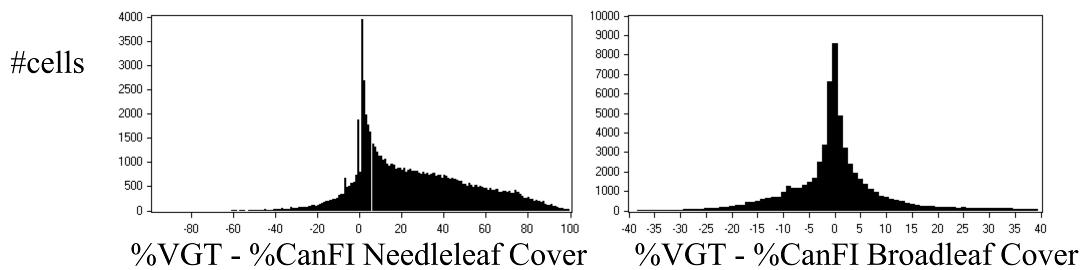

Figure 4: Histograms of differences between land cover and CanFI estimates of needleleaf and broadleaf forest cover.

Both the forest cover and water fraction quality control exercises focus on consistency of source data and derived products. No direct comparison was made between the final maps and independently surveyed forest or water cover sources. Rather, the emphasis of this study was to present the best available survey data provided by mandated agencies in a systematic, national coverage.

4. APPLICATIONS AND CONCLUSIONS

The waterbody fraction coverage has been used by Statistics Canada to modify per capita land use statistics and to revise the estimate of surface water for Canada. At a research level, the new coverage is being used to generate revised carbon budgets and evapotranspiration estimates of Canada and has been applied to correct leaf area index and burned area estimates from remote sensing. The water coverage could be further improved by incorporating recent NTDB surveys and by using high resolution remote sensing (radar or optical) in areas where water cover is known to have changed (e.g. dams).

The forest cover type maps are being used to estimate biomass at a provincial and eventually national extent. Ideally, the methodology could be extended to mapping species cover. A revision of CanFI (CanFI2000) is soon to be released with increased coverage and careful consideration of between province differences in reporting. This database could improve the existing forest cover type map and could perhaps be used to extend the result to a species cover map of Canada.

The choice of a coarse spatial scale for the national coverages is, in part, a conscious decision designed to reflect the level of spatial and thematic accuracy in the source data and ancillary sources used to address biases. Recent efforts to improve the input waterbody and forest inventory surveys will be propagated into these products and may eventually supercede these maps. Until that point, users of these two national coverages should be aware of their limitations. These are primarily related to vintage and thematic errors in the survey data. For example, the NTDB maps are surveyed mid-summer and infrequently updated unless large flooding or droughts have occurred in areas effecting populations, transportation networks, or the energy sector. Similarly, users of the forest cover type coverages should factor in disturbance effects (e.g. fire) and the synoptic nature of the ecoregion text attributes. Users of the both coverages should use recent, local survey information to assess inter annual changes and should factor in seasonal changes due to snowmelt to summer conditions.

Nevertheless, these consistent and clearly documented coverages fill large data gaps at a national scale. The emphasis on the preservation of survey information from mandated agencies guarantees that these maps are at least as informative as the source data aggregated to the reporting level selected.

Acknowledgment: This project was supported by Program for Energy Research and Development (PERD), Climate Change Action Plan 2000 Initiatives, and Long-term Space Plan (LTSP). We thank David Price, Josef Cihlar, and Quanfa Zhang for constructive discussions and assistances.

REFERENCES

Anonymous (1996) Standards and specifications of the National Topographic Database, Geomatics Canada Catalogue# M52-70/1996E. Minister of Supply and Services Canada, Canada pp 25

Anonymous (1997) Concepts, sources and methods of the Canadian System of Environmental and Resource Accounts (Econnections: linking the environment and the economy), Statistics Canada Technical Report # 16-505-GPE. Statistics Canada, Canada pp 68

Anonymous (2000a) Geospatial Database - Standards and Specifications, Centre for Topographic Information Unpublished Technical Report. Centre for Topographic Information, Geomatics Canada, Sherbrooke pp 42

Anonymous (2000b) Canadian Digital Elevation Data, Standards and Specifications, Centre for Topographic Information Unpublished Technical Report. Centre for Topographic Information, Geomatics Canada, Sherbrooke pp 13

Anonymous (2001) A Plot-based National Forest Inventory Design For Canada, accessed on March 10, 2002 at http://www.pfc.cfs.nrcan.gc.ca/monitoring/inventory/canfi/docs/design2_e.pdf. An interagency partnership project, Canadian Forest Service, Canada pp 59

Centre for Land and Biological Resources Research (1996) Soil Landscapes of Canada, v.2.2. Research Branch, Agriculture and Agri-Food Canada, Ottawa

Chen W, Chen J, Cihlar J (2000) An integrated terrestrial ecosystem carbon-budget model based on changes in disturbance, climate, and atmospheric chemistry. Ecological Modelling 135 55-79

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

Cihlar J, Latifovic R, Beaubien J, Guindon B, Palmer M (2002) TM-based accuracy assessment of a land cover product for Canada derived from SPOT VEGETATION data. Submitted to Canadian Journal of Remote Sensing.

Ecological Stratification Working Group (1996) A National Ecological Framework for Canada. Agriculture and Agri-Food Canada, Research Branch, Centre for Land and Biological Resources Research and Environment Canada, State of Environment Directorate, Ottawa/Hull pp 125.

Gray SL, Power K (1997) Canada's forest inventory 1991: the 1994 version - technical supplement. Canadian Forest Service, Information Report BC-X-362E pp 31.

Lowe JJ, Power K, Gray SL (1996) Canada's Forest Inventory 1991: the 1994 version. An addendum to Canada's Forest Inventory 1991. Canadian Forest Service, Information Report BC-X-362E.