

This is an author's preprint of an article that was published in *Environmental Monitoring and Assessment* 131:221-230. There may be minor editorial differences between this preprint and the version that was published in the journal.

The published version of this paper can be viewed at <http://dx.doi.org/10.1007/s10661-006-9470-8>

© Her Majesty the Queen in Right of Canada, 2006

Using spatial pattern to quantify relationship between samples, surroundings, and populations

Michael A. Wulder^{1*}, Trisalyn A. Nelson², and David Seemann¹

¹Canadian Forest Service (Pacific Forestry Centre), Natural Resources Canada, Victoria, British Columbia, Canada

²University of Victoria, Department of Geography, Victoria, British Columbia, Canada

* Corresponding Author:

Mike Wulder

Canadian Forest Service

506 West Burnside Road, Victoria, B.C., Canada V8Z 1M5

email: mwulder@nrcan.gc.ca, phone: 250 363-6090, fax: 250 363-0775

Submit: *Environmental Monitoring and Assessment*
Submitted: May 2006.

Abstract

The need for accurate carbon budgeting, climate change modelling, and sustainable resource management has lead to an increase in the number of large area forest monitoring programs. Large area forest monitoring programs often utilize field and remotely sensed data sources. Sampling, via field- or photo-plots, enables the collection of data with the desired level of categorical detail in a timely and efficient manner. When sampling, the aim is to collect representative detailed data enabling the statistical reporting upon the characteristics of larger areas. As a consequence, approaches for investigating how well sample data represent larger areas (*i.e.*, the sample neighbourhood and the population) are desired. Presented in this communication is a quantitative approach for assessing the nature of sampled areas in relation to surrounding areas and the overall population of interest. Classified Landsat data is converted to forest / non-forest categories to provide a consistent and uniform dataset over a 130,000 km² study region in central British Columbia, Canada. From this larger study area 322 2 x 2 km photo-plots on a 20 x 20 km systematic grid are populated with composition and configuration information for comparison to non-sampled areas. Results indicate that typically, within the study area, the spatial pattern of forest within a photo-plot is representative of the forest patterns found within primary and secondary neighbourhoods and over the entire population of the study. These methods have implications for understanding the nature of data used in monitoring programs worldwide. The ability to audit photo and field plot information promotes an increased understanding of the results developed from sampling and provides tools identifying locations of possible bias.

Key words: forest monitoring, land cover mapping, large area mapping, photo-plots, spatial pattern, Landsat, large area, forest inventory

1. Introduction

Large area land cover mapping and monitoring is important for assessing change in biological, chemical, and physical environmental processes (Roberts et al., 2003). Critical as input for climate modeling (Nonomura et al., 2003), carbon budgeting (Hoshizaki et al., 2004), and tracking environmental change and sustainability (Mayaux et al., 2005), the number of large area land cover monitoring programs is increasing worldwide (Franklin and Wulder, 2002). Multinational agreements, international marketplaces, and growing concerns over environmental degradation have increased pressure to account for forests and the sustainability of management practices (Kangis, 2006; Wulder et al., 2004).

Large area land cover and forest monitoring programs are increasingly utilizing a combination of general and detailed data sources (Tomppo, 1996; Huang et al., 2001). Plot based systems are typically required to provide the level of detail required for national and international reporting. Data with less spatial detail, such as satellite imagery enable stratification or allow general forest conditions to be characterized over large areas for a known point in time. Plot based systems, based upon field collected (Tokola, 2006; Smith, 2002) or photo interpreted plot data (Gillis, 2001; Kleinn et al., 2005), provide the basis for the generation of statistics for monitoring the characteristics and dynamics of larger, often country-wide, populations.

The usefulness of detailed forestry data is inherently tied to whether or not the information collected at sample locations represents more general forest conditions. As such, approaches for quantifying how well sample data reflect broader landscape characteristics are desirable. Determining the level of relatedness between data collected at sample locations and general conditions of larger regions, provides an indicator for the effectiveness of sampling programs and aids in understanding the information content of large area land cover maps. Such approaches also make it possible to quantitatively determine if there are biases in sampling scheme design. For instance, whether sample locations are systematically excluded from disturbance or subject to more disturbances than other regions in a local neighbourhood can be verified.

The goal of this research is to develop and demonstrate a method to quantify the representativeness of sample data of forest conditions in the surrounding area. The spatial pattern of forests is the quantitative measure used to assess the relatedness of the sample and population. Two objectives are used to meet this research goal.

1. A method is demonstrated for assessing how well the composition (or amount) of forest at a sample location represents the composition of forest within primary and secondary neighbourhoods, and over the study region population.
2. A method is demonstrated for assessing how well the configuration (or spatial arrangement) of forest at a sample location represents the configuration of forest within primary and secondary neighbourhoods, and over the study region population.

2. Background

2.1 Canada's Forest Monitoring Programs

While the methods developed and demonstrated through this work have broad application to a number of land cover and forest monitoring programs, the utility of this approach is demonstrated with reference to Canadian forest monitoring programs (Wulder et al., 2004). As home to approximately 10% of the world's forests, forest monitoring is a challenge, due to issues such as large area of interest, regional stewardship responsibilities, variable data vintages, differing standards and definitions. Canada is meeting these challenges primarily through two programs: the Canadian National Forest Inventory (NFI; Gillis, 2001) and Earth Observation for Sustainable Development of Forests (EOSD; Wulder et al., 2003).

The Canadian NFI is a national program for collecting and reporting forests characteristics. The aim of the NFI is to provide consistent, relevant, and timely forestry information to assist management and monitoring (Gillis, 2001). Individual Canadian provinces and territories are responsible for the stewardship of natural resources with each jurisdiction controlling the definition and collection of information. To facilitate national integration in support of creating a consistent and systematic reckoning of Canada's forest land-base, the Canadian Forest Service (CFS) has drafted a series of guidelines, including standards and definitions to promote cross-jurisdiction consistency (Gillis, 2001). Further, NFI polygons are mapped with six hierarchical levels, enabling applicable mapped categories to be portrayed at differing levels of detail. With the first measurement of the new NFI approaching completion, guidelines for remeasurement have been developed, with provincial and territorial partners, to enable the systematic monitoring of the sample locations over time (Gillis et al., 2005).

The NFI is designed to produce national statistics through collection of a 1% sample of Canada's land mass. To do so, the country is partitioned with a 20 by 20 km grid, with 2 by 2 km photo-plots located at the grid vertices. Forest characteristics interpreted from the 2 by 2 km photo-plots are used to produce detailed information that represent the characteristics of a given 20 by 20 km grid. Over 20,000 photo-plots provide reliable area statistics and attribute estimation for monitoring Canada's forests (Gillis, 2001; Wulder et al., 2003). The NFI grid and photo-plot sampling design is one example of a large area forest monitoring program that integrates general forest characterization with detailed sample information. The approach described here, for assessing how well the sample reflects more general forest conditions, could be used with any plot based monitoring system.

In a second national program, the Canadian Forest Service and Canadian Space Agency are mapping the land cover of the forested ecozones Canada with satellite imagery (Wulder et al., 2003). The EOSD program is mapping land cover of the forested ecozones of Canada (Wulder et al. 2003) using 30 m spatial resolution orthorectified Landsat-7 ETM+ data produced for the entire Canadian landmass (Wulder et al., 2002). The EOSD land cover map is based on a unsupervised classification and hyperclustering followed by labeling and represents conditions centered on the year 2000 (Wulder et al., 2004). The Landsat based land cover mapping is focused upon the forested ecozones of Canada which encompasses approximately 6,472,601 km². To capture this area with Landsat imagery, over 475 images are required. Some overlap outside of the forested ecozones is also necessitated, resulting in a total mapped area of 8,229,109 km² when summing the area covered by the scenes being mapped. The EOSD classification links to

the NFI hierarchical classification, with the EOSD classification being a closed legend capturing the NFI level 4 cover type, with level 5 density classes (Wulder and Nelson, 2003).

The NFI and EOSD classification systems were designed to enable integration (Wulder and Nelson, 2003). By incorporating forest maps produced from two independently derived monitoring programs, a better understanding of the nature of information generated through each program is gained (Rommel et al., 2005). Using the 2 x 2 km photo-plot boundaries as context, the classified EOSD data may be used for characterization of forest composition and configuration, making it possible to assess how representative photo-plots are of larger areas and to provide an additional tool for understanding the nature of NFI data. The approach presented is generic and envisioned as portable to frameworks and datasets other than those used in this study.

2.2 Pattern Descriptors

There are two characteristics to spatial pattern when data are binary: composition and configuration. *Composition* relates to the aspatial element of pattern and *configuration* to the spatial element (Li and Reynolds, 1993, 1994, 1995; Gustafson, 1998; Boots, 2003). By characterizing the landscape as forest and non-forest, the composition of forest pattern can be quantified as the proportional amount of forested area. If forested pixels are coded as black (B) and non-forested coded as white (W), composition is determined by the proportion of forested pixels (p_b).

In this study, configuration refers to the spatial distribution of forest and can be quantified using a measure of spatial autocorrelation. When a pattern exhibits spatial autocorrelation there is clustering of either similar or dissimilar values on a map (Griffith, 1992). By enumerating the similarity of each pixel to its neighbours, it is possible to quantify the magnitude of spatial autocorrelation. Join-counts are often used for quantifying spatial autocorrelation in categorical data. Joins are defined by the common boundary between pixels and may be classified as BB, WW, or BW (Cliff and Ord, 1981). In this case a BB join would represent two forested pixels with a common boundary, and by counting the number of joins we can quantify the configuration of forested pixels. Typically, join-counts are used to determine if configuration varies from an expectation of complete spatial randomness. In such cases, the statistical hypothesis can be tested by converting join count measures to standardized Z-scores (Cliff and Ord, 1981). Here however, as will be detailed in the methods, we compare the number of BB joins in the photo-plot to a distribution of BB joins generated from the surrounding area.

3. Study Area and Data

This research focuses on 322 photo-plots that are centered near the city of Prince George (53°53'N, 122°40'W) in British Columbia, Canada (Figure 1). Covering approximately 14% of the province, the study area includes several land cover types and has variable topography. The main forest species in this area are hybrid white spruce (*Picea glauca* × *engelmannii*), sub-alpine fir (*Abies lasiocarpa*), lodgepole pine (*Pinus contorta*), and Douglas-fir (*Pseudotsuga menziesii*). Other less common species include black spruce (*Picea mariana*), trembling aspen (*Populus tremuloides*), balsam poplar (*Populus balsamifera*), and paper birch (*Betula papyrifera*) (Meidinger and Pojar, 1991). In this area, forestry is the dominant anthropogenic influence, interspersed with agricultural land use activities. Natural disturbances such as fire and insect infestations also have an impact on this landscape.

To prepare the data for analysis, the EOSD data are converted to binary grids, having a cell size of 25 m. Grid cells are coded as black (B) if more than 50% of the cell area is forest and white (W) otherwise. Where EOSD data are missing, due to cloud or shadow in the satellite imagery, pixels are categorized as non-forest. As may be expected, cloud and shadow may occasional reach amounts that necessitate exclusion of the impacted plots. The rationale and heuristics for plot exclusion are presented in the methods section. EOSD data are also overlain with photo-plot boundaries and the NFI 20 by 20 km grid.

4. Methods

The methods of this research are designed to test the null hypothesis that the composition and configuration of forest in a photo-plot are not samples of a reference distribution (population) generated from two surrounding neighbourhoods and the entire study region. The primary neighbourhood is the 20 by 20 km zone that each photo-plot is associated with in the NFI scheme. The secondary neighbourhood was comprised of the eight 20 by 20 km zones that are spatially contiguous to the primary neighbourhood (Figure 2). For comparison with the photo-plots, primary neighbourhood corresponds to 100 units 2 by 2 km in size, 96 of which do not intersect spatially with the photo-plots. Similarly, the secondary neighbourhood is composed of 800 units 2 by 2 km in size. The composition and configuration of photo-plots are also compared to a global reference distribution in which the entire study area was partitioned into 2 by 2 km units.

Of the 322 photo-plots in the study area, 209 could be completely contained within the secondary neighbourhoods and were used in the analysis. Results are not reported for photo-plots associated with EOSD data having a high proportion of missing information, due to cloud and shadow in the satellite imagery. A frequency distribution of the number of pixels with missing information in each of the 2 by 2 km plots had a natural break at 10%. Exploratory investigations indicated that where 10% of cells have missing data in five or more of the 96 plots (2 by 2 km in size) the spatial pattern of forest was different relative to the other plots. Therefore, these photo-plots were excluded from the analysis. A total of 191 photo-plots had missing data in less than 10% of cells. These were assessed in the global analysis. For analysis with the primary neighbourhoods 160 photo-plots were assessed, while for analysis on secondary neighbourhoods 150 photo-plots met the criteria. Thresholds used to identify locations having “too much” missing information can be determined on a study by study basis.

For each photo-plot, forest pattern composition is quantified as the proportion of black (B) pixels and forest configuration is quantified by counting the number of pixels with BB joins. Following a rook’s case definition of neighbourhood, joins were computed for one cell in each of the north, south, east, and west directions. To determine how well the photo-plot forest pattern represents forest pattern in neighbouring regions, photo-plot composition and configuration values are compared to distributions of similar measures generated for 2 by 2 km units of EOSD data that are within the primary or secondary neighborhoods. For analysis of the primary neighbourhood the reference distribution is generated using the 96 2 by 2 km units that do not intersect the photo-plot boundary. For analysis of the secondary neighbourhood a random sample of 96 2 by 2 km units was selected from the 800 possible units. The global analysis benefited from computations done for the primary and secondary neighbourhood analysis and all plots with less than 10% cloud were used to generate the global reference distribution (209 (96 + 96) – pixels with >10% cloud = 37249). The null hypothesis that the composition and configuration of forest in a photo-plot are not a sample from a reference distribution (population) generated from surrounding neighbourhoods was then accepted or rejected using a standard critical value of 0.10 and two tailed test.

For the primary and secondary neighbourhood analyses, results from statistical comparisons of the null hypotheses that forest composition and configuration for the NFI photo-plots are a sample of encompassing forest patterns are presented aspatially and spatially. In the aspatial

2 presentation, relative frequency distributions of p -values are presented and provide evidence of
the nature of the relationship between photo-plot forest pattern and the pattern in the surrounding
4 regions, as captured with measures of composition and configuration. The spatial presentation of
results characterizes spatial variation in how well photo-plot forest pattern represents forest
6 pattern on the landscape and is useful for locating areas that may benefit from more detailed
investigation. Global results are reported as single values.

5. Results

5.1 Composition

When the composition of photo-plot forest pattern is assessed relative to the forest pattern in the primary neighbourhood, the null hypothesis is rejected for 42 of 160 (26.3%) of the photo-plots. Testing of the same null hypothesis on the secondary neighbourhood shows similar results, with the null hypothesis being rejected for 40 of 150 (26.7%) photo-plots. Global results show even more similarity between photo-plots and the population with the null hypothesis only rejected in 17 of 191 (8.9%) cases. Regardless of the neighbourhood used for analysis, relative frequency distributions of p -values are skewed right, having an abundance of high p -values, indicating that the amount of forest within the photo-plot tends to be high (although not significantly different) in comparison with the surrounding regions (Figure 3).

Results of testing the null hypothesis associated with photo-plot forest composition for primary and secondary neighbourhoods are presented spatially in Figure 4. The two squares represent primary (inner square) and secondary (outer square) neighbourhoods. Square colour reflects whether the null hypothesis was accepted (yellow), rejected (red), or not considered due to an abundance of missing data (grey). The majority of plots missing substantial amounts of data are situated in the east and northeastern portions of the study area. This is likely the result of cloud caused by orographic lifting in the mountains east of the interior plateau.

Of the 140 plots that were not missing substantial amounts of data in either primary or secondary neighbourhoods, the null hypothesis associated with the composition of forests was accepted in both analyses for 95 photo-plots and rejected in both analyses in 28 photo-plots. For 21 photo-plots the null hypothesis was accepted for the primary neighbourhood and either rejected or excluded from the analysis in the secondary neighbourhood. For 15 photo-plots the null hypothesis was rejected for the primary neighbourhood and either accepted or excluded from analysis for the secondary neighbourhood.

5.2 Configuration

The null hypothesis that forest configuration within the photo-plot is a sample from a reference distribution generated from the surrounding area is rejected for 44 of 160 (27.5%) photo-plots when analyses are undertaken using the primary neighbourhood and for 42 of the 150 (28.0%) photo-plots when the secondary neighbourhood is analyzed. Again the global analysis lead to even fewer rejected null hypotheses, only 19 of the 191 (9.9%) photo-plots. Relative frequency histograms for configuration p -values are similar to those shown for the composition analysis, being skewed right with a high proportion of high p -values, indicating that the amount of forest within the photo-plot tends to be more clustered (although not significantly different) in comparison with the surrounding regions (Figure 5).

Results of the configuration analyses are presented spatially in Figure 6, using the same colour scheme described for Figure 4, and are similar to the results presented on composition. Of the 140 photo-plots that do not have a substantial amount of missing data in either the primary or secondary neighbourhoods, the null hypothesis associated with the configuration of forest in photo-plots is accepted for both scales of analysis for 95 photo-plots and rejected at both levels for 29 photo-plots. Twenty-one photo-plots have the null hypothesis associated with composition accepted for the primary neighbourhood and rejected or excluded from the analysis in the

secondary neighbourhood. Only 15 photo-plots are associated with rejected composition null hypotheses in the primary neighbourhood and acceptance or exclusion for the secondary neighbourhood.

When forest pattern in the photo-plot is compared to the primary neighbourhood 41 of the 160 (25.6%) photo-plots null hypothesis associated with both composition and configuration are rejected, and for only 5 (3.1%) photo-plots is one test rejected whilst the other accepted.

For analysis of the secondary neighbourhood, for 38 of the 150 (25.3%) photo-plots null hypotheses associated with both composition and configuration were rejected and for only 6 of 150 (4.0%) photo-plots was one null hypothesis rejected whilst the other accepted. For the global analysis only 16 of 191 (8.4%) photo-plots have both null hypotheses rejected. One photo-plot has the composition null hypothesis rejected, but the configuration photo-plot accepted; three photo-plots have the configuration hypothesis rejected while the composition hypothesis is accepted.

6. Discussion and Conclusions

With the increasing need for both large area mapping and detailed information, mapping programs that incorporate coarse and fine spatial resolution data have become more common. Ideally, detailed data gathered at sample locations are useful for statistically valid capture and subsequent reporting on the forest characteristics of larger areas. However, as with any sampling program, the sample data are most beneficial if they represent the population. Using spatial pattern, it is possible to quantify how well the forest conditions at sample locations reflect the larger population. The method outlined can be applied to any monitoring program based upon field or photo-plot sampling frameworks, where comprehensive data (i.e., satellite image classifications) representing a larger area, or population, are available. While methods in this paper are applied to photo-plots, a similar approach could be used when sampling is based on field plots. In a similar fashion, image data could be clipped to a consistent shape and size guided by field plot locations. The clipped image data would be indicative of forest conditions at the sample site and may in turn be compared to the local or global neighbourhoods. If applicable, the clipped data around the sample locations may also represent specific attributes not only forest / non-forest conditions.

6.1 Composition

Results from the composition analysis indicate that for both primary and secondary neighbourhoods, approximately 75% of the photo-plots, the forest composition is not significantly different from forest composition in the surrounding region. When considering the results of the global analysis, most photo-plots (91.1%) adequately represent the proportion of forest in the study region, and results provide confidence that the sampling scheme used in the NFI provides an adequate representation of the amount of forest at most locations. It must be recalled that any given plot in a systematic sample does not need to exactly characterize its neighbourhood; the important factor is to be representative over the entire population. Knowledge of more local relationships (e.g., the primary and secondary neighbourhoods in this study) provide a means for vetting locations where issues may exist, not necessarily to undermine the confidence in the samples to adequately characterize the population.

Although the proportion of photo-plots that have the composition null hypothesis accepted is similar for primary and secondary neighbourhood analysis, it is more common for primary neighbourhoods to be accepted and secondary rejected or excluded than vice versa. It is notable when forest composition of any photo-plot is more related to secondary than to primary neighbourhoods, and photo-plots having such patterns could be flagged for more detailed investigation in the future. Similarly, photo-plots where the composition null hypothesis is rejected at both spatial scales could be identified for investigation.

Photo-plots where the composition null hypothesis is accepted for analysis of both neighbourhoods dominate the landscape. In instances where the null hypothesis is rejected, inspection indicated that the rejection was typically explained by topography, land cover, land use, or harvesting activities (or combinations thereof). For instance, the presence of a photo-plot in a river valley will result in less, or an absence, of harvesting, which can result in a significant difference in configuration to surrounding areas. As forest harvesting activities are the dominant disturbance on the landscape, areas where harvesting is not undertaken (i.e. parks, alpine areas,

wetlands) are typically where significant differences in composition and surrounding characteristics are found.

6.2 Configuration

Results of the configuration analysis are consistent with composition results. For analysis on both primary and secondary neighbourhoods approximately 75% of photo-plots have forest composition that is not significantly different from surrounding areas. In the global analysis an even larger proportion (90.1%) of photo-plots adequately reflect population conditions. Results from configuration analysis assure users that within the study region, Canada's NFI photo-plots are a representative sample of the spatial distribution of nearby forests. Configuration analysis also produces similar results to composition analysis in terms of photo-plots having a combination of significant and insignificant results at the different analysis spatial scales.

As would be expected from the geographic theory of spatial autocorrelation which suggests that near objects are more related than far (Griffith 1992), there is evidence that photo-plots better represent the spatial configuration forest in nearby locations or primary neighbourhoods than in farther away secondary neighbourhoods. Photo-plots tend to provide a better representation of forest configuration for primary neighbourhoods than for secondary neighbourhoods. Geographic differences (e.g., topography, rivers, lakes) that occur in a portion of the secondary neighbourhood will result in differing forest conditions.

Similarity in the relative frequency histograms of p -values and the spatial distribution of significant results, associated with tests of forest composition and configuration, reflect the similarity of these spatial pattern measures. While the two measures quantify different elements of forest spatial pattern, the amount of forest will impact the likelihood of forest occurring in nearby pixels. Used in combination, these measures provide a valuable tool for identifying locations where composition and configuration results are substantially different and may alert the analyst to photo-plots possibly requiring further investigation. It is interesting that regardless of comparisons with primary neighbourhoods, secondary neighbourhoods, or the entire study area, using configuration tests the null hypothesis that the photo-plot forest characteristics are representative of the population is rejected more often.

The method demonstrated in this paper provides an approach to using the spatial pattern of forests to gain a better understanding of how well data generated from sampled locations represent the population or surrounding areas. Using these approaches it is possible to quantify how well photo-plots reflect both the proportion (composition) and spatial arrangement (configuration) of forests, and provide a means for determining if photo-plots are biased. Presenting results both aspatially and spatially gives insight into both the strength and spatial variation in sampling design efficacy, and enables locations in need of further investigation to be identified.

Acknowledgements

2 This research is enabled through funding of the Government Related Initiatives Program of the
Canadian Space Plan of the Canadian Space Agency and Natural Resources Canada – Canadian
4 Forest Service. Joanne White of the Canadian Forest Service is thanked for valuable discussions
and insights.

6

References

- Boots, B. (2003). Developing local measure of spatial association for categorical data. *Journal of Geographical Systems*, 5, 139-160
- Cliff, A. & Ord J. (1981). *Spatial Processes Models and Applications*. (London: Pion Limited)
- Franklin, S.E. & Wulder, M.A. (2002). Remote sensing methods in medium spatial resolution satellite data land cover classification of large areas. *Progress in Physical Geography*, 26, 173-205
- Gillis, M. (2001). Canada's National Forest Inventory. *Environmental Monitoring and Assessment*, 67, 121-129
- Gillis, M., Omule, A.Y. & Brierley, T. (2005). Monitoring Canada's forests: The National Forest Inventory. *The Forest Chronicle*, 81, 214-221
- Griffith, D. (1992). What is spatial autocorrelation? Reflections on the past 25 years of spatial statistics. *L'Espace géographique*, 3, 265-280
- Gustafson, E. (1998). Quantifying landscape spatial pattern: what is the state of the art? *Ecosystems*, 1, 143-156
- Hoshizaki, K., Niiyama, K., Kimura, K., Yamashita, T., Bekku, Y., Okuda, T., Quah, E.S. & Noor, N.S.M. (2004). Temporal and spatial variation of forest biomass in relation to stand dynamics in mature, lowland tropical rainforest, Malaysia. *Ecological Research*, 19, 357-363
- Huang, C., Yang, L., Homer, C., Coan, M., Rykhus, R., Zhang, Z., Wylie, B., Hegge, K., Zhu, Z., Lister, A., Hoppus, M., Tymcio, R., DeBlander, L., Cooke, W., McRoberts, R., Wendt, D. & Weyermann, D. (2001). Synergistic Use of FIA Plot Data and Landsat 7 ETM+ Images for Large Area Forest Mapping. (In Proceedings of the Third Annual Forest Inventory and Analysis Symposium, Traverse City, Michigan, October 17-19, 2001. U.S. Department of Agriculture, Forest Service GTR NC-230. pp 50-55)
- Kangis, A. (2006). Assessing the world's forests. (In A. Kangis & M. Maltamo (Eds.), *Forest Inventory: Methodology and Applications* (pp.279-366). Dordrecht, The Netherlands:Springer) 362p.
- Kleinn, C., Ramirez, C., Holmgren, P., Valverde, S.L. & Chavez, G. (2005). A national forest resources assessment for Costa Rica based on low intensity sampling. *Forest Ecology and Management*, 210, 9-23
- Li, H. & Reynolds, J.F. (1993). A new contagion index to quantify spatial patterns of landscapes, *Landscape Ecology*, 8, 155-162
- Li, H. & Reynolds, J.F. (1994). A simulation experiment to quantify spatial heterogeneity in categorical maps. *Ecology*, 75, 2446-2455

- Li, H. & Reynolds, J.F. (1995). On definition and quantification of heterogeneity. *Okios*, 73, 280-284
- Mayaux, P., Homgren, P., Achard, F., Eva, H., Stibig H. & Branthomme, A. (2005). Tropical forest cover change in the 1990s and options for future monitoring. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 360, 373-384
- Meidinger, D. & Pojar, J. (Eds.) (1991). Ecosystems of British Columbia. *British Columbia Ministry of Forests Special Report*, Series No. 6, 330p.
- Nonomura, A., Songa-Ngoie, K. & Fukuyama, K. (2003). Devising a new digital vegetation model for eco-climatic analysis in Africa using GIS and NOAA AVHRR data. *International Journal of Remote Sensing*, 24, 3611-3633
- Rommel, T.K., Csillag, F., Mitchell, S. & Wulder, M.A. (2005). Integration of forest inventory and satellite imagery: A Canadian status assessment and research issues. *Forest Ecology and Management*, 207, 405-428
- Roberts, D.A., Keller, M. & Soares, J.V. (2003). Studies of land-cover, land-use, and biophysical properties of vegetation in the Large Scale Biosphere Atmosphere experiment in Amazonia. *Remote Sensing of Environment*, 84, 377-388
- Smith, W.B. (2002). Forest inventory and analysis: A national inventory and monitoring program. *Environmental Pollution*, 116, S233-S242
- Tokola, T. (2006). Europe. (In A. Kangis & M. Maltamo (Eds.), *Forest Inventory: Methodology and Applications* (pp.295-366). Dordrecht, The Netherlands: Springer) 362p.
- Tomppo, E. (1996). Multi-source national forest inventory of Finland. (In Proceedings of the Subject Group S4.02-00 *Forest Resource Inventory and Monitoring* and Subject Group S4.12-00 *Remote Sensing Technology*; Tampere, Finland. IUFRO World Congress, pp.27-49)
- Wulder, M.A., Dechka, J.A., Gillis, M., Luther, J., Hall, R.J., Beaudoin J. & Franklin, S.E. (2003). Operational mapping of the land cover of the forested area of Canada with Landsat data: EOSD land cover program. *Forestry Chronicle*, 79, 1075-1083
- Wulder, M.A. & Nelson, T. (2003). EOSD legend: characteristics, suitability, and compatibility. *Technical Report*, Canadian Forest Service, Pacific Forestry Centre, Victoria, British Columbia. URL: http://www.pfc.forestry.ca/eosd/cover/eosd_report_e.html
- Wulder, M., Loubier, E. & Richardson, D. (2002). Landsat-7 ETM+ orthoimage coverage of Canada. *Canadian Journal Remote Sensing*, 28, 667-671
- Wulder, M.A., Kurz, W. & Gillis, M. (2004). National level forest monitoring and modeling in Canada. *Progress in Planning*, 61, 365-381

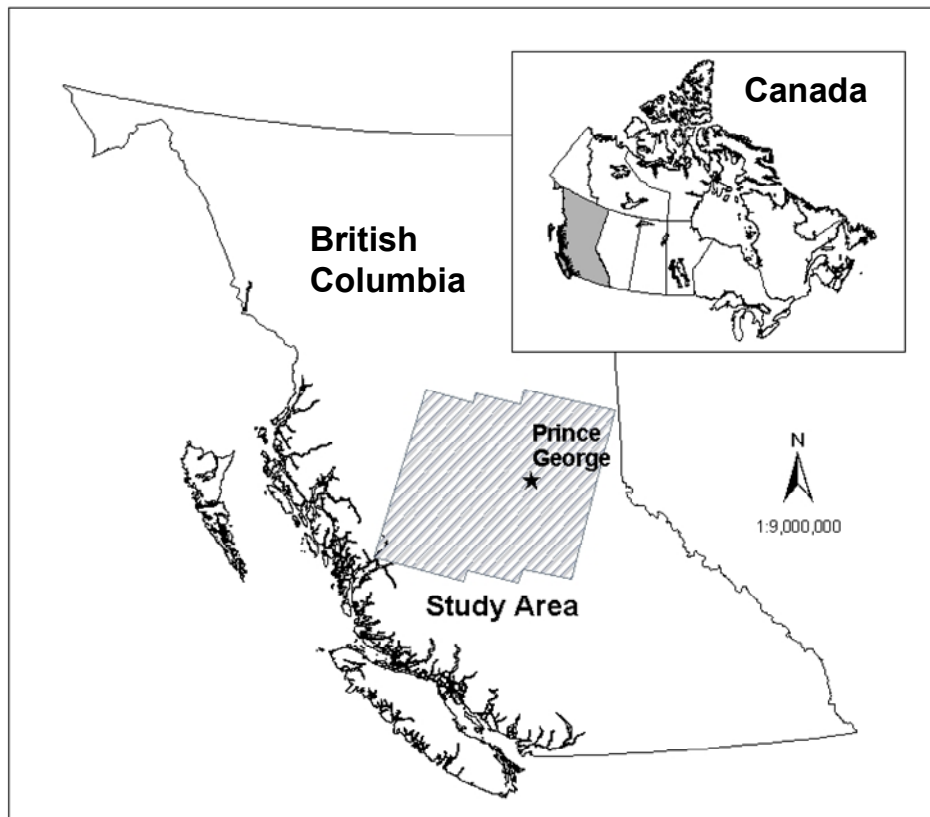


Figure 1. The study area is centered on Prince George, British Columbia, Canada (53°53'N, 122°40'W) and includes 322 2 by 2 km photo-plots located on the 20 by 20 km NFI grid.

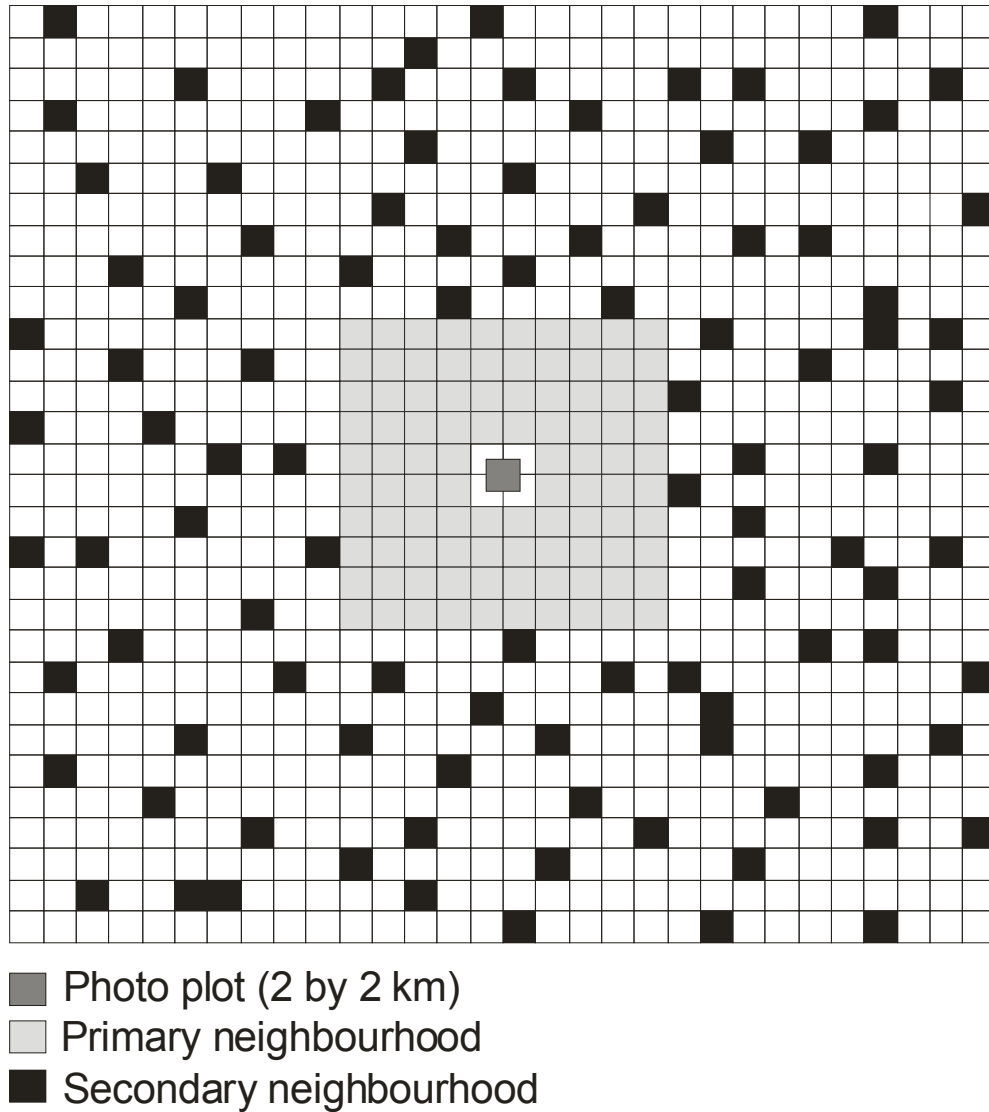


Figure 2. Primary (20 by 20 km) and secondary (60 by 60 km) neighbourhoods for each photo plot. Reference distributions for the primary neighbourhood are generated from the 96 2 by 2 km plots that do not intersect the photo-plot. Reference distributions for the secondary neighbourhood are generated from 96 randomly selected 2 by 2 km plots.

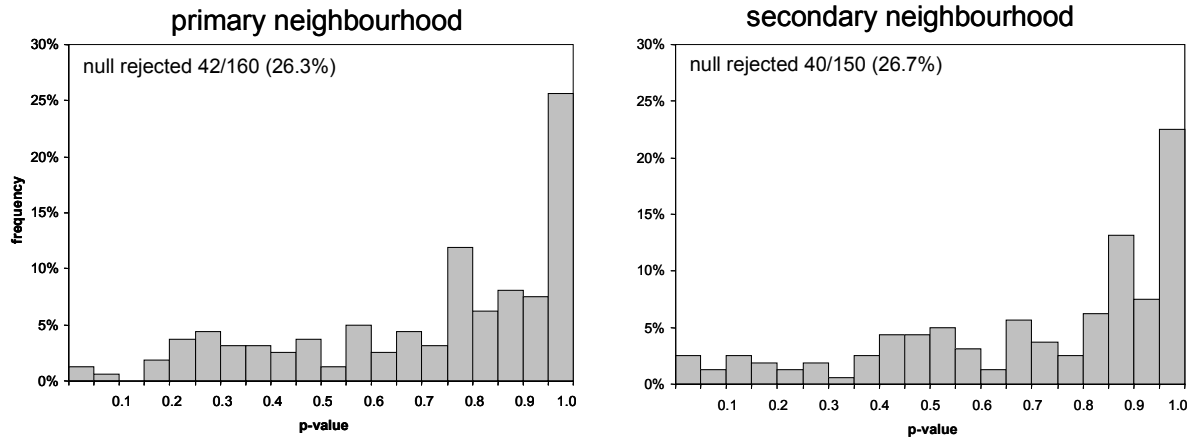


Figure 3. P values from testing the null hypothesis that forest composition is similar for NFI photo plots and EOSD data in primary and secondary neighbourhoods ($\alpha = 0.10$, test = two tailed)

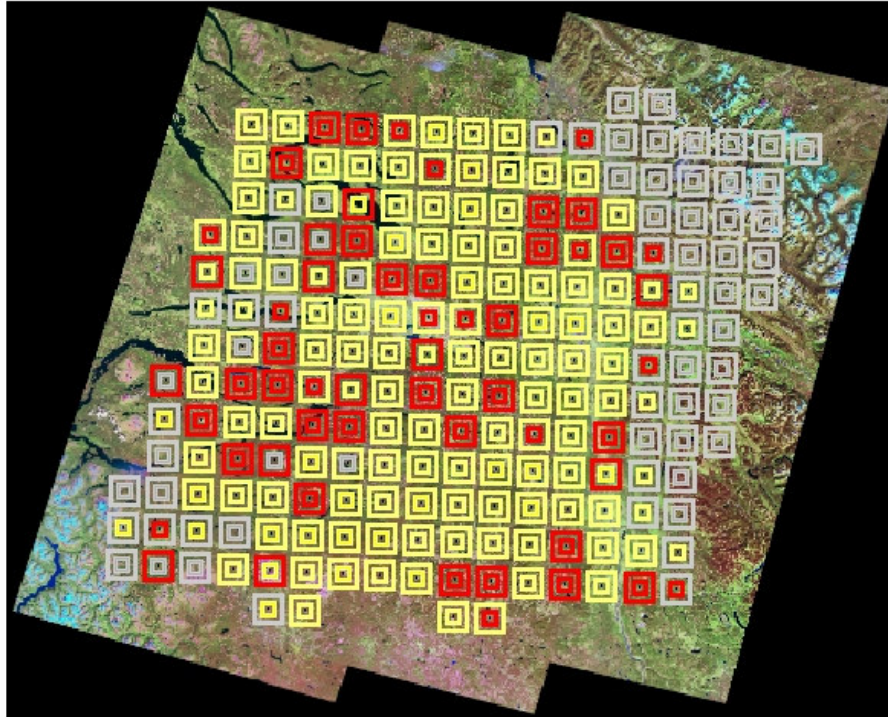


Figure 4. Map of results of testing the null hypothesis that forest composition is similar for NFI photo-plots and EOSD data in primary (inner square) and secondary neighbourhoods (outer square). Red = null hypothesis rejected and yellow = accepted ($\alpha = 0.05$, test= two tailed). Grey represents locations excluded from the analysis due to missing image data (e.g. cloud, shadow).

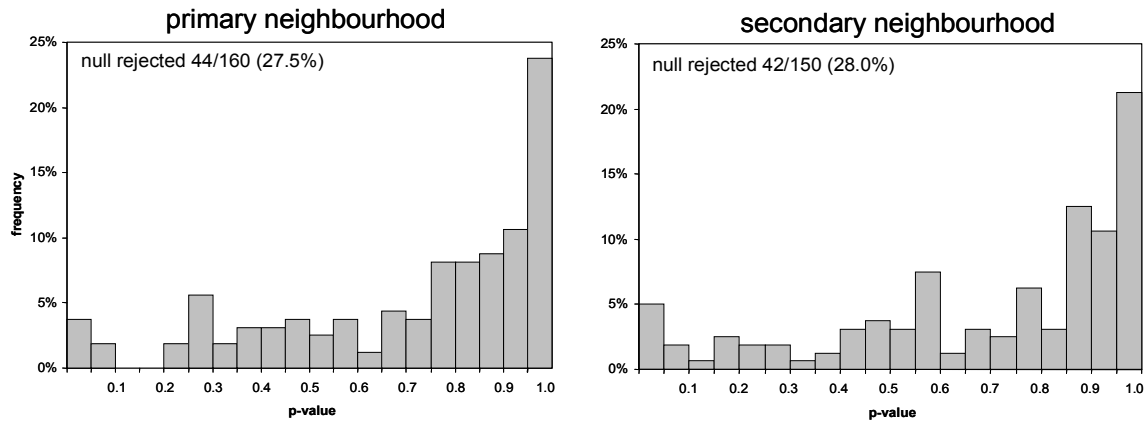


Figure 5. p -values from testing the null hypothesis that forest configuration is similar for NFI photo-plots and EOSD data in primary and secondary neighbourhoods ($\alpha = 0.10$, test= two tailed).

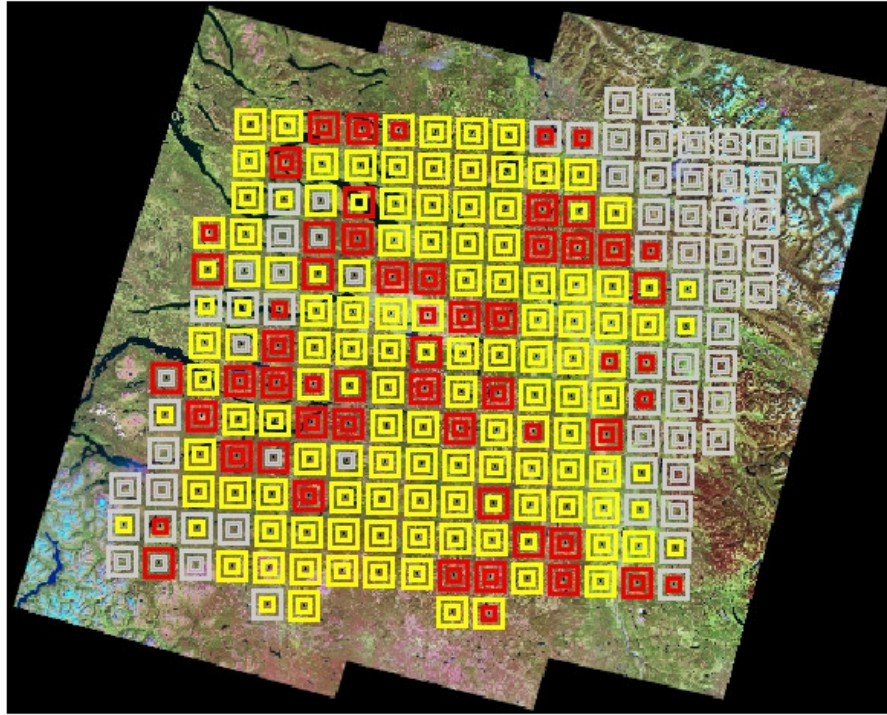


Figure 6. Map of results of testing the null hypothesis that forest configuration is similar for NFI photo-plots and EOSD data in primary (inner square) and secondary neighbourhoods (outer square). Red = null hypothesis rejected and yellow = accepted ($\alpha = 0.05$, test= two tailed). Grey represents locations excluded from the analysis due to missing image data (e.g. cloud, shadow).