A Preliminary Discussion of

A CFI Approach to Integrated Data Collection in British Columbia

Submitted to:

Jon Vivian, RPF
Terrestrial Information Branch
Ministry of Sustainable Resources Management
Victoria, BC

Project: MSI-002

October 19, 2001

J.S. Thrower & Associates Ltd. *Consulting Foresters*
Vancouver – Kamloops BC

Acknowledgements

This report was completed with the support of the BC Ministry of Sustainable Resources Management (MSRM) Terrestrial Information Branch (TIB). We thank Jon Vivian, *RPF* (MSRM TIB) for supervising the project. We also thank the following for providing information: Del Meidinger, *P.Ag*. (Research Branch, BC Ministry of Forests [MOF]); Stephen Titus, *PhD*, Univ. of Alberta, Edmonton; John Hennigman, *RPF* (Forest Practices Branch, MOF); Lawrence Bowdige, *RPF* (TIB); Keith Tudor, *RPF* (TIB); Joe Braz, *RPF* (TIB); and Gordon Nigh, *PhD*, *RPF* (Research Branch, MOF). Stephen Titus also reviewed an earlier version of this paper.

Table of Contents

List of Tables

List of Figures

1. INTRODUCTION

1.1 BACKGROUND

Many of the key forest management processes in BC rely on up-to-date inventories, growth and yield (G&Y) information, and G&Y models.[1] Data to develop each of these information sources have traditionally been collected in separate programs – which has often resulted in duplication of effort, complex systems, and high cost to the forest industry. The Terrestrial Information Branch (TIB) of the Ministry of Sustainable Resources Management (MSRM)[2] current data collection programs include the Vegetation Resources Inventory (VRI) Phase II (timber emphasis), Change Monitoring Inventory (CMI), and permanent sample plots (PSPs) for model development (Appendix II).[3] In response to the industry's need to streamline systems and reduce costs, the TIB has initiated a process to examine the potential to integrate some of these data collection systems.

One method proposed by the TIB is to integrate data collection and reduce overall program cost by using a continuous forest inventory (CFI)[4] approach to data collection in management units (Timber Supply Areas [TSAs] and Tree Farm Licenses [TFLs]). The proposed approach uses fixed-area PSPs installed in a management unit using a common plot design which are measured repeatedly over time for a variety of forest resource values (e.g., timber, ecology, coarse woody debris, soils, etc.). A similar CFI approach is used in the 400 ha Blue Mountain Woodlot east of Maple Ridge (Appendix III). This CFI was established in 1987 and includes 92 prism plots installed on a systematic square grid that are re-measured for tree attributes every five years.

1.2 PROJECT GOAL & OBJECTIVES

The primary goal of this project is to identify and discuss the pros and cons of using a single standard fixed-area plot configuration within a CFI design to meet the data collection needs for inventory, monitoring, and model development. The key questions asked by TIB are:

1. Can a single standard fixed-area plot design be used for all data collection needs?
2. Will this CFI approach meet the combined needs of inventory, monitoring, and model development?
3. Can monitoring data also be used for model development?

Additionally, the project Schedule A (Appendix IV) specifies that the following CFI design elements be considered:

1. A minimum set of variables to be collected on all plots.
2. A streamlined approach to data collection to allow sampling a plot in one day.

[1] A glossary of terms is given in Appendix I.

[2] Formerly the Resources Inventory Branch of the Ministry of Forests.

[3] Model development also requires data from experiments that are used to determine silviculture treatment responses. The Ministry of Forests experimental program (EP) addresses this to a large extent. The need for experimental data is not discussed further here, but there is a need to develop a separate comprehensive strategy for gathering model development data.

3. The implications on sample size and variance from using a single point instead of a plot cluster.
4. Using a grid as opposed to the current VRI PPSWR (probability proportional to size with replacement) to locate sample plots.
5. The need for pre- or post-stratification.
6. Selection of sub-samples and remeasurement schedules for monitoring or growth models development.
7. Statistical estimators and variance estimators.
8. Cost of plot installation and remeasurement.

1.3 TERMS OF REFERENCE

This paper is a preliminary discussion of the major issues surrounding a single standard plot configuration and the CFI approach envisioned by TIB. The development, testing, and possible future implementation of an integrated approach to data collection in BC will require more discussion and input from a wider range of experts and consideration of the higher-level needs of the provincial government. The intent is that this report will not be circulated to anyone outside of the MSRM or the expert panel. This paper was prepared for inventory experts familiar with the situation in BC and thus we have not defined many terms and acronyms.

This report was prepared by J.S. Thrower & Associates Ltd. under contract to Jon Vivian, *RPF* (TIB). Preparation of this paper was lead by A.Y. Omule, *PhD RPF* with contributions from Eleanor McWilliams, *MSc RPF*, Guillaume Thérien, *PhD*, Jim Thrower, *PhD RPF*, and Bill Warren, *PhD*.

[4] The term CFI is not used consistently in the literature. In this report, CFI is broadly defined as methods to sample an area repeatedly over time and includes overall sample design, plot design, and estimation methods.

2. BUSINESS NEEDS

2.1 PRIMARY NEEDS

The primary data collection business needs of TSAs and TFLs in BC directly considered in this paper are: 1) inventory; 2) monitoring of growth, yield, and other forest attributes; and 3) growth models for regular inventory updates and application in timber supply analysis.

1. *Inventory.*
   All TSAs and TFLs in the province require a forest inventory. The provincial standard is the VRI program which is a collection of tools to complete forest inventories that includes: a) Phase I photo estimation of tree and stand attributes; b) Phase II ground samples (generally a star cluster of five prism plots) used to statistically adjust the Phase I estimates; c) stem analysis to develop net volume adjustment factors (NVAF) to adjust the cruiser-called net volumes taken in the Phase II samples; d) sampling to estimate the within polygon variation (WPV) in a management unit to refine the precision estimate of the Phase II adjustment; and e) the recently developed CMI procedures to monitor growth, yield, and other attributes in the management unit. The Phase II plots can be relocated, but are considered temporary by the TIB to provide a single point-in-time measurement of the inventory.

2. *Monitoring.*
   Provincial legislation does not yet require monitoring of growth, yield, and other forest attributes in TSAs or TFLs; however, there has been considerable discussion recently on how monitoring can be used to address needs for third party market certification (e.g., Forest Stewardship Council) as well as some inventory and G&Y information needs. To help address this emerging need, the TIB has designed the CMI process (mentioned above). This program includes a series of PSPs located randomly or systematically across a management unit.[5] The intent is that these plots are re-measured periodically over time to provide information to track timber and non-timber attributes against predictions or expectations of those attributes. A primary use of CMI data is to check G&Y model projections as applied in the management unit.

3. *Developing growth models.*
   Growth models are used to update managed and natural stand inventories (short-term yield projection) and for application in timber supply analysis (long-term yield projection). Growth models are also used (although less widely) for strategic silviculture, harvesting, and corporate planning. The Province has a long history of establishing and maintaining PSPs to develop and maintain G&Y models for these purposes. These PSPs were mostly subjectively located in natural and treated stands or as part of designed experiments. Developing growth models includes deriving basic relationships that comprise the model, validation, and calibration of models.

2.2 SECONDARY NEEDS

The secondary business needs for TSAs and TFLs that require data collection and are indirectly considered in this paper are:

1. *Site Index Adjustment (SIA).*
   Many TFLs and most Innovative Forestry Practices Agreement (IFPA) areas in the province have

---

[5] The provincial level CMI for the National Forest Inventory (NFI) is a separate program that should be kept independent of the management unit sampling programs.

chosen a statistical approach to adjust preliminary estimates of potential site index for yield projection in post-harvest regenerated (PHR) stands and application in timber supply analysis. The data for this application are collected from randomly selected plots across the management unit where potential site index is estimated in PHR stands and used to adjust preliminary estimates attached to each polygon of ecological maps for the entire management unit. This approach has not been advocated by the MOF, but has been widely used by industry and is accepted for generating managed stand yield tables for application in timber supply analysis.

2. *Site Index-Biogeoclimatic Ecosystem Classification (SIBEC)*.
   The MOF is developing estimates of average potential site index for the major tree species on all site series in the Biogeoclimatic Ecosystem Classification (BEC) system in BC. The data for these estimates are from detailed plots subjectively located where potential site index is estimated and the BEC site series is classified. The intent is that these estimates can be combined with ecological maps to generate managed stand yield tables similar to the SIA approach. These estimates can also be used to estimate potential site index and for yield projections for any area with an estimate of BEC site series. These estimates are not well developed for many areas of the province because of the enormous data requirements and the approach has not been widely applied in timber supply analysis in BC.

3. *Forest health monitoring.*
   Forest health issues are increasingly important in BC and there is a desire among practitioners to develop a program to monitor the level of various forest health agents, particularly in managed stands. Possible monitoring objectives include comparing actual yields with model predictions, demonstrating the efficacy of treatments, and checking existing pest thresholds. Required data includes stand descriptors and the incidence and severity of selected damage agents (40-50 province wide). Currently there is no program in place to provide this information. There are several established forest health surveys and the Operational Planning Regulation requires relevant forest health issues be evaluated in areas under forest development plans; but there is no system to coordinate and integrate data collection activities. A forest health monitoring program could provide data to compare model predictions with observed yield of healthy and unhealthy managed stands to demonstrate efficacy of silviculture treatments against damage agents, and to check existing pest incidence and severity thresholds.

4. *Linking VRI with Predictive Ecosystem Mapping (PEM)*.
   The recent use of computer-based predictive systems to generate ecological maps has created the need to provide some form of accuracy assessment for these maps. Some practitioners have suggested using the ecological classification of VRI Phase II plots as point estimates to compare with PEM maps so as to develop some form of accuracy assessment. The benefits of this suggestion is that the VRI plots are randomly located within management units, however; most VRI Phase II samples in the province are based on timber emphasis plots (TEPs) and thus do not include ecological classification.

3. SAMPLE DESIGN OPTIONS

3.1 Overview

This section describes the CFI-type approach suggested by TIB (Option A) and a simplified sampling with partial replacement (SPR) design that we propose as an alternative to promote discussion (Option B). Both design options have advantages and disadvantages that are discussed below in relation to some key components. Further details are given in Section 4. These options are presented to generate discussion of possible alternatives – not as the only two possibilities to meet the combined data collection needs.

3.2 Option A

A set of fixed-area PSPs similar to the current CMI plots[6] is the CFI design envisioned by the TIB. Re-measurements would be based on the traditional periodic inventory where the plots are measured in a given time period (e.g., every 5-10 years).

Advantages

The major advantages of this sampling design are:

1) *Good estimates of change.*
   Maintaining a fixed set of plots over time provides the best estimates of change and mortality (assuming there is persistence [similarity] in the population spatial pattern over time [Appendix V]). This in turn provides better data for monitoring and model development.
2) *Simplicity.*
   Using a single fixed-area plot design will simplify the production of field manuals, training of crews, data entry, compilation, storage, and analysis. This will result in some cost savings.

Disadvantages

The major disadvantages of this sampling design are:

1) *Potential high cost.*
   Fixed-area plots are less efficient than variable radius plots for many tree attributes, thus the cost of establishing the number of fixed-area plots (e.g., PSPs) required to estimate VRI Phase II yields with currently accepted precision targets would be very expensive. In addition, this approach is not sensitive to budgets fluctuating over time.
2) *Potential for a persistent, bad sample.*
   There is a risk that the fixed set of plots may not represent the area of interest very well. The consequence is that this unrepresentative sample will persist over time as the same plots are measured at each occasion.
3) *Reduced ability to address spatial and temporal issues.*
   The spatial pattern in the population will change over time and this cannot be addressed with a fixed set of plots. This is an important issue for estimates of yield (VRI Phase II) but not as important for estimates of change (monitoring and model development).

---

[6] B.C. Ministry of Forests. 2001. Change Monitoring Inventory Ground Sampling Procedures for the Provincial Change Monitoring and the National Forest Inventory, Version 1.1 March 30, 2001. MOF, Victoria, B.C. 203 pp.

3.3 Option B

Maintain a fixed set of fixed-area plots (e.g., CMI plots) as proposed in option A, but using fewer of these plots and augment them with independent samples of temporary plots (TSPs) as needed for VRI Phase II. This can be considered a simplified sampling with partial replacement (SPR) design.[7] Data from the PSPs would be used for monitoring and model development, and if all or a randomly chosen subset of the PSPs are re-measured when the TSPs are established, the two data sources can be combined to give a better estimate of current yield for VRI Phase II. Further, this option includes the provision to use varying plot and plot cluster designs for the TSPs to make this sampling as efficient as possible to reflect local conditions (including the use of variable radius plots).

Advantages

The major advantages of this approach are:

1) *Potentially lower overall cost.*
   For a given precision level, smaller samples are required to estimate growth (change) than yield. This option allows for maintaining a sample size of PSPs sufficient to estimate change, but not requiring that the set of PSPs also be sufficient for estimating yield at any point in time. In addition, variable radius plots are more efficient for estimating timber yield than fixed-area plots.
2) *Ability to address potential bad samples.*
   Reduces the implications of a persistent bad sample if the PSP sample locations provide an estimate of yield that does not well represent the area of interest.

Disadvantages

The major disadvantage of this approach is:

1) *Increased complexity.*
   This option includes several plot or plot cluster designs and increases the complexity of data capture, compilation, storage, and analysis. This will add cost to developing the procedures over option A.

[7] SPR estimators to estimate change do not have to be used. Change can be estimated using re-measured plots only and current values from all the plots (re-measured and independent).

4. SAMPLE DESIGN DETAILS

4.1 OVERVIEW

This section discusses the details of the two sample design options presented in Section 3. The design elements specified in the project objectives (Section 1.2) are addressed here. Unless specifically noted, the design elements discussed below could be addressed with either sample design option A or B.

4.2 OBJECTIVES

An integrated sampling approach must address the objectives of monitoring, model development, and VRI Phase II. Specific objectives may vary by management unit but in general will include providing data to:

1) Adjust VRI Phase I estimates to estimate overall totals, averages, and associated confidence intervals of timber and selected non-timber resources. (Data currently provided by VRI Phase II ground sampling).
2) Check G&Y estimates used in timber supply and silviculture decisions. Managers need a reasonable level-of-comfort that the G&Y models adequately reflect the actual G&Y achieved on the ground. (Data currently provided by new CMI program being pilot tested – not yet implemented).
3) Track non-tree vegetation to detect changes in presence or proportion of plant species as an early warning indicator of climatic or other changes. (Data currently provided by new CMI program being pilot tested – not yet implemented).
4) Meet the needs of certification agencies such as the Forest Stewardship Council (FSC). (Data currently provided by new CMI program being pilot tested – not yet implemented).
5) Developing G&Y models; including survivor growth, mortality, and ingrowth. (Data currently provided by the MOF PSP program and a variety of experimental plots).
6) Meet or contribute to secondary information needs as described in Section 2.2. (Data currently provided by stand alone and special programs).

4.3 TARGET POPULATION

The target population in a management unit will be defined by the specific objectives for the unit. It is possible that the three primary data needs (VRI Phase II, monitoring, and model development) will define differing target populations. For example, the target population for VRI is typically the Vegetated Treed area [BC Land Cover Classification Scheme (BCLCCS)] or the timber harvesting landbase, while for monitoring and model development, the target population may only be PHR stands. Alternatively, the target population for monitoring may be the entire landbase as required by some certification protocols. Under Option A, the broadest definition of the target population must be selected to ensure all objectives are met. Under Option B, different target populations could be chosen for the PSPs and the TSPs.

4.4 STRATIFICATION

Stratification prior to sampling is feasible under both design options; however, the short-term advantages of stratification must be considered against the potential long-term negative impacts. If done properly, stratification will improve sample efficiency and reduce field costs, but will result in increased complexity in planning, documentation, and analysis. Stratification creates problems when strata definitions and boundaries change over time making the negative aspects a much larger issue for PSPs than for TSPs.

Given the different information requirements, the most likely stratification will be between mature and PHR stands. Monitoring, and in particular model development needs, will generally focus on PHR stands, while VRI Phase II needs generally focus on mature timber. Ideally, these needs should be assessed on an individual management unit basis in conjunction with a provincial strategy for collecting model development data.

For any strata defined prior to sampling, boundaries should be well defined spatially (landscape level) and stored in GIS. Stratification should be minimized to reduce problems related to subdividing a population before sampling (e.g., mapping errors, increased documentation costs, and increased complexity in sample selection, remeasurement, and analysis).

Post-stratification is also feasible under Options A and B. The potential for post-stratification should be considered when designing a system to ensure a large enough sample to provide adequate observations in key strata. Post-stratification should be based on variables other than those observed on the plots.

4.5 SAMPLE SELECTION

A sample of PSPs or TSPs can be selected using either systematic sampling or PPSWR (Table 1). Theoretically, both methods are similar, however, there are practical difference and advantages of each method in relation to the type of plot used (TSPs or PSPs) and the intent to sample over time.

Table 1. Potential Sample selection methods.

Resource	Phase II	Monitoring
Systematic		
Fixed-grid	√	√*
Sorted List	√	√
PPSWR		
Stratified	√*	√*
Random	√	√

* Methods in use or being pilot tested.

In systematic sampling, either a fixed grid (e.g., 1 km) or a sorted list is used to generate the sample points. In the PPSWR method, plot locations would be selected using pre-stratified or random PPSWR. While there may be little difference between these two options for integrated sampling, the systematic selection may distribute the sample better than the PPSWR. Conversely, in a CFI framework with a fixed set of plots and a population that is changing over time, the PPSWR may give more long-term flexibility than the systematic selection. Furthermore, the stratified PPSWR may not be suitable and we suggest that pre-stratification should be avoided where possible.

Table 2. Resources currently measured in VRI Phase II and monitoring.

Resource	Phase II		Monitoring	
	Timber	*Non-timber*	*G&Y*	*Non-timber*
Trees				
Large Trees (≥ 4 cm)	√		√	
Small Tree (< 4 cm)	√		√	
Site Trees	√		√	
Stumps	√		√	
Ecology				
Shrub height		√		√
Plant lists		√		√
Soil		√		√
CWD	√*			√
Range		√		√

* CWD can be included in the non-timber category.

4.6 VARIABLES TO MEASURE

An integrated sampling approach will likely require measuring both timber and non-timber resources (Table 2) as defined by the management unit objectives. For timber resources, attributes often include top height, age of top height trees, site height, age of site trees, site index, average height, number of trees, species, diameter at breast height, basal area, net volume, and incidence and severity of damage agents. Non-timber attributes often include plant lists, ecological site series descriptions, soil properties, coarse woody debris (CWD) volume, and range resources.

We suggest that the minimum variables to measure under either design option are set by the MSRM. The decision of what to include in a minimum set is a function of the business needs of the MSRM and funding available for data collection. We recommend that a minimum set of variables is kept as small as possible. For example, the large list of variables included in the VRI and CMI standards greatly increase the time to sample plots, and many of these measurements will not be of much value to some users. For example, the tree attributes listed in the CMI manual are a reasonable starting point for a minimum set of variables to measure; however, we suggest that several of these attributes should be options that depend on the specific needs of a management unit, and should not be part of the minimum requirements. For example, CMI attributes that could be optional include call grading, net factoring, wildlife attributes, and stump measurements.

4.7 PLOT DESIGN

Permanent Plots

The TIB requested that a single plot configuration for integrated sampling be investigated. We considered the current Phase II and monitoring sampling systems that include fixed-area and variable radius plots (Table 3), and other plot configurations to collect information to meet the objectives specified in Section 4.2.

Table 3. Current sampling methods.

Resource	Sampling Unit	No. Units	Phase II	Monitoring G&Y	Monitoring Non-timber
Trees					
Large (≥ 4 cm)	VRP	5	√		
Large (> 9 cm)	FAP 400 m^2	1		√	
Medium (4-9 cm)	FAP 100 m^2	1		√	
Small (< 4 cm)	FAP 19.6 m^2	1	√	√	
Stumps	FAP 19.6 m^2	1	√	√	
Site Trees	FAP 100 m^2	5	√	√	
Ecology					
Shrub cover	Line intercept	2	√		√
Shrub height	FAP 314 m^2	1	√		√
Herb, moss	FAP 100 m^2	1	√		√
Soil	Soil pit	1	√		√
CWD	Line intersect	2	√		√
Range	FAP 50 m^2	4	√		√

Given current protocols (Table 3), and that TIB restricted the focus to fixed-area plots, a slightly modified CMI plot may be the best design (Table 4, Figure 1). The slight modification changes the diameter-tagging limit for the 400 m^2 fixed-area plot to sample trees 12.5 cm+, which is the lowest diameter utilization limit currently considered in most timber supply analyses. This will reduce the number of trees to be tagged and measured and will allow most plots to be established in a single day. The smaller diameter trees would still be measured in the 100 m^2 and 19.6 m^2 subplots.

We also suggest that field procedures be modified to map plots that straddle more than one polygon or forest type to better track land type changes over time. For sample plots in treated stands (Appendix II), pre-treatment data should be collected, where feasible; otherwise the analysis must rely on history files for the parent polygon of the sample plot. If a plot has been harvested, the plot would be re-established over the sampling point.

Table 4. Proposed plot configuration summary.

Component	Description and purpose
11.28 m radius (400 m^2)	Large tree ($\geq$ 12.5 cm) measurements (dbh, height, mortality, incidence and severity of damage agents related to forest health, timber quality, etc.).
5.64 m radius (100 m^2)	Medium tree ($\geq$ 4 cm) measurements (same as large tree measurements); Site tree measurements; and plant list (herb, moss, and seedlings % cover).
2.50 m radius (19.6 m^2)	Small tree (< 4 cm) count and stump measurements
Line intercept (24 m)	Two line intercepts, each 24 m long; the first one on a random bearing and the second at 90 degrees from the first, to measure shrub foliar cover and CWD.
10.00 m radius (314 m^2)	Plant list (shrubs & trees % cover) and shrub height.
Soil pit	Pit close to the IPC, but outside the 5.64-m radius plot.
3.99 m radius (50 m^2)	Four small plots located along the CWD line intercepts, to measure range attributes.

Temporary Plots

Design Option B includes establishing TSPs for VRI Phase II in a simplified SPR design. This independent sample could be established using one of several plot design options and it is likely prudent to allow flexibility given the wide range of stand types in the province.

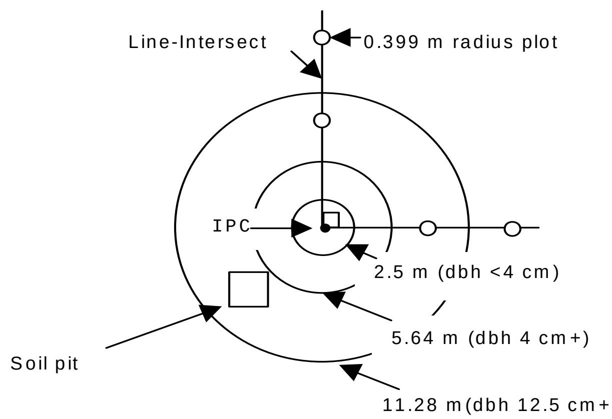


Figure 1. Proposed plot configuration for Phase II monitoring timber & non-timber attributes

For improved efficiency in providing data for statistical adjustment of the Phase I inventory, the Phase II plot samples should be as highly correlated with the Phase I polygon estimates as possible (Appendix V). The higher the correlation, the fewer polygons that need to be sampled. The degree of correlation depends on within polygon variability and the number of plots established in the polygon. Ideally, more plots are established in stand types with higher within polygon variation. The value of an increased number of plots must be weighed against the increased cost - and cost is a function of the plot design. We propose five options for consideration (this is not an exclusive list):

1. The same slightly modified CMI plot as proposed above (Appendix V).
2. The CMI plot surrounded by a cluster of small fixed-area (100 m^2) plots.
3. The CMI plot surrounded by a cluster of variable radius plots. A variable radius plot could be added at the IPC to measure VBAR.
4. The usual VRI cluster and measuring only tree attributes.
5. The usual VRI cluster with adding as many count plots (prism sweeps) as can be established in one day's fieldwork.

Clusters versus Single Plots

The consideration of clusters versus single plots is really a question of sampling efficiency and cost. Plot clusters are often used in forest sampling to increase efficiency, which occurs when the increased cost of measuring a cluster of plots is more than offset by a decrease in variance. This question must consider the cost of sampling, the type of plot and associated variance, spatial variation in the population, and the reduction in variance from the clustering. Plot clusters are often more efficient in forestry sampling applications when they can be completed in one day. However, there are some very heterogeneous stand types in BC where a cluster requiring two days may be more efficient than two independent smaller clusters or two individual plots.

Sampling efficiency is also related to the type of plot and the attribute of interest. For example, empirical evidence from TFL 37 suggests that the single CMI plot is approximately 30% less efficient for estimating current volume than the VRI Phase II prism plot cluster.[8] That is, sampling for net volume using the single CMI plot would require 30% more plots than would the five-point VRI cluster, to attain the same target sampling error.[9] The analysis of the data is the same for clusters as for a single plot, since the values at a sampling location are averaged on a per-hectare basis.

We expect that sampling efficiency can be improved in many stand types in BC by using clusters larger than the current Phase II cluster of five prism plots; however, this should be examined using computer simulation. There are now sufficient Phase II samples from a wide variety of stand types from across the province to test different clustering options using simulation methods with appropriate cost functions (Appendix V).

4.8 SAMPLING OVER TIME

Option A

Different approaches could be used to establish and re-measure plots over time in the TIB CFI-type approach to integrated sampling. The simplest option for planning and data analysis is to establish and remeasure all plots on the same cycle. However, the logistics of securing funding and field crews to measure a large number of plots in one year is often not possible.

Another option for remeasurement is to use the rotating panel approach, where a panel (or subset) of the plots is measured periodically.[10] One simple example of this approach (Figure 2) involves dividing an entire set of plots into a number of groups or panels (e.g., N1, N2, and N3) and measuring all of them in the initial year, and again at subsequent occasions remeasuring only a few groups (e.g., one group per occasion).

---

[8] The CV of the ratio (ground volume to Phase I volume) was 68% using VRI compared to 77% using the original CMI plot. The average ground volumes from the two types of plot were, as expected, virtually identical. The VRI and CMI plots were established at the same 39 sample locations.

[9] Note that the impact of variable radius plots versus fixed-area plot and 5-point cluster versus single CMI plot are mixed together, and not possible to separate in this comparison.

[10] J.S. Thrower & Associates Ltd. 1999. Conceptual methods for change inventory and monitoring of vegetation resources. Version 2.0. Contract report prepared for the BC MOF. Dec. 1999.

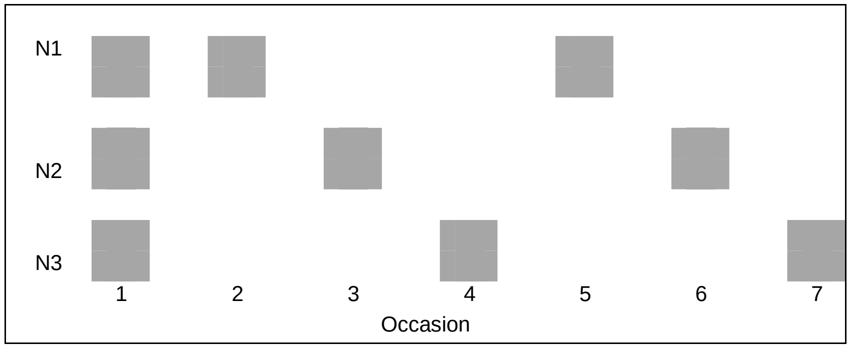


Figure 2. Example of a rotating panel CFI (3 groups of plots measured over a period of 7 years).

Options to implement the rotating panel approach are shown in Appendix V.[11] The advantage of this approach is that remeasurement costs and other resources can be spread over a number of years. The disadvantage is that the estimates of change may be less precise than the traditional approach and may fluctuate annually. The rotating panel approach merits further consideration, although the periodic approach may be preferable where the total number of plots is relatively small (e.g., < 30 plots).

Option B

The intent in Option B is to keep the sample size of PSPs small enough to allow the plots to be established and measured on the same cycle. The independent sample of TSPs could then be established whenever to update the inventory. The independent sample could also be used to derive NVAF estimates. The NVAFs from the independent sample could be used to compile remeasurement data, even on occasions where independent sample were not taken.

4.9 STREAMLINING DATA COLLECTION

Pilot projects have shown that CMI plot installation using the current standards generally requires more than one day to complete a plot. So that one CMI plot can be completed in a day, we suggest considering eliminating call grading, net factoring, wildlife codes, and stump attributes. If net factoring is retained, the rule-based process could be simplified.

There is also an opportunity to further reduce costs by not measuring all resource values on all plots. Timber attributes will be measured on all plots, but non-timber attributes may only need to be measured on a subset of the PSPs (and TSPs under Option B) depending on objectives and precision requirements. Eliminating ecological classification on some or all plots would also greatly reduce the time to establish and measure a CMI plot. This information can be taken at subsequent remeasurements if needed at that future time.

4.10 SAMPLE SIZE

Estimating an appropriate sample size for an integrated approach to data collection is more complicated than for a single attribute. For example, in each management unit the desired level of precision for both estimates of current values and change need to be considered along with the level of confidence in the G&Y predictions to be checked with the monitoring data. Also, the sample size should be statistically efficient over successive occasions. Finally, larger samples may be required to allow for future post-stratification of the data.

---

[11] Scott, C.T., Kohl, M. and Schnellbacher, H.J. (1990). A comparison of periodic and annual forest surveys. Forest

Option A

Variability in current values (e.g., volume) is likely higher than that for change, thus it seems logical to base the initial sample size on the precision required for current estimates (i.e., VRI Phase II). Other more sophisticated approaches using optimization techniques for multiple objectives are available (e.g., linear programming or dynamic programming) but these are probably not warranted in this case.

Current VRI Phase II projects are showing CVs of the ratio of ground volume to Phase I volume estimates of about 60-70%.[12] These CVs result in sample sizes of 140 to 220 VRI plots, to achieve a sampling error of 10% (95% probability). Larger sample sizes would be required with CMI plots.

Option B

Fixed-area plots are established in option B to estimate growth; thus sample size should be based on the expected variability in growth estimates and desired precision. The TSP sample size will be based on the expected variability in current yield and the desired precision. Also taken into consideration will be the number of PSPs that can be incorporated into the estimate of current yield. Ideally, TSPs would be established in a year where the PSPs are re-measured so that the two samples can be combined.

4.11 ANALYSIS

Several design-based estimation methods have been discussed for estimating change and current yield.[13] These include three different estimation methods for data collected using a rotating panel approach.

Option A

Analysis would be simple and straightforward. For monitoring and model development, the data could be used as in current practice. For VRI Phase II, the compiled data from the fixed-area plot would simply replace the compiled data from the current VRI cluster in the ratio and regression estimation procedure.

Option B

The data from PSPs could be analyzed and used for monitoring and model development as in current practice. If different plot configurations are used for the TSPs (e.g., variable radius plots), then combining yield estimates from the TSPs and PSPs for use in Phase II adjustments may require weighting.[14]

Some practitioners do not support using variance estimators based on random sampling for systematic samples. We do not believe this is a problem because the usual practical impact appears to be a possible over-estimation of the resulting confidence interval. Therefore, we suggest that data from systematic samples are treated as if from a random sample and the variance is calculated accordingly. This assumes the sampled population is in “random” order. Cochran (page 213) states, “In this situation we would expect systematic sampling to be essentially equivalent to simple random sampling and to have

---

Science 45: 433-451.

[12] Embedded in these CV estimates from recent VRI projects is the impact of using VDYP to assign volume to polygons in Phase I.

[13] J.S. Thrower & Associates Ltd. 2000. Statistical estimation methods for provincial change inventory. Version 2.0; and Graphical & statistical analysis for monitoring estimates of change at the management-unit level. Contract reports to the BC MOF. March 31, 2000.

[14] Typically when combining two independent estimates they are weighted inversely proportional to their variances.

the same variance."[15] Practical experience and simulation tests show that systematic sampling is generally more efficient that random sampling.

We also suggest that other estimation methods are explored and tested. For example, the Kalman filter approach can be used in CFI-type designs as demonstrated by Dixon and Howitt[16] and Kangas.[17] Estimation could also be based on model-based inference. This approach and its associated estimation methods have already been adapted for G&Y estimation in some areas of the world and should be considered for application in BC (Appendix V). Considering these estimation methods does not impact the basic implementation of the sampling methods discussed here.

4.12 Sample Planning

A sample plan outlining the integrated sampling components needs to be developed for each management unit project. The sample plan could include activities associated with implementation and documentation of the project including:

1. Sample plans for the VRI Phase II, monitoring, and model development as deemed necessary.
2. Installing plots for the activities concurrently.
3. Using the same measurement protocols and definitions for core variables.

[15] Cochran, W.G. 1977. Sampling Techniques, third edition. John Wiley & Sons. New York, NY.

[16] Dixon, B.L. and R.E. Howitt. 1979. Continuous forest inventory using a linear filter. For. Sci. 25 (4): 675-689.

[17] Kangas, A. 1991. Updated measurement data as prior information in forest inventory. Silva Fenn. 23:181-191.

5. SUMMARY OF KEY ISSUES

5.1 Could a Single Standard Fixed Area Plot Design Be Used?

A single standard fixed-area plot design could be used to meet the needs of VRI Phase II, monitoring, and model development. However, the advantages and disadvantages of doing so must be carefully considered.

Advantages

1) Simplified field crew training, production of field cards, and manuals.
2) Simplified data compilation, management and analysis.
3) Cost reductions associated with these two items.

Disadvantages

1) Field costs likely to be prohibitively high for VRI Phase II yield estimates if current precision targets are to be maintained. Fixed-area plots are not as efficient for inventory as variable radius plots.
2) Lack of flexibility in collecting data across the wide variety of stand types in the province.

5.2 Will A CFI Approach Meet The Combined Data Needs?

The CFI design proposed by TIB (Section 3.2) could meet the combined needs of VRI Phase II, monitoring, and model development. However, the advantages and disadvantages of using this single approach must be carefully considered.

Advantages

1) Maintaining a fixed set of plots over time provides the best estimates of change and mortality (assuming there is persistence [similarity] in the spatial pattern in the population over time [Section 0]). This in turn provides better data for monitoring and model development.
2) Using a single fixed-area plot design will simplify the production of field manuals, training of crews, and data entry, compilation, storage, and analysis. This will result in some cost savings.

Disadvantages

1) The cost of establishing the number of PSPs required to estimate VRI Phase II yields with currently accepted precision targets would be prohibitively expensive in most cases. In addition, this approach is not sensitive to budgets fluctuating over time.
2) Due to the initial sample locations, the PSPs may provide an estimate of yield significantly different from the true value for the area, and this will persist over time because the same plots are measured each time.
3) The spatial pattern in the population will change over time and this cannot be addressed with a fixed set of plots. This is an important issue for estimates of yield (VRI Phase II) but not as important for estimates of change (monitoring and model development).

5.3 Can Monitoring Data Also Be Used For Model Development?

We suggest that data collected for monitoring can also be used for model development in BC. Using the same data for monitoring and model development can be risky because this could result in incorrect conclusions from monitoring. This risk is a function of the degree to which the monitoring data have influenced the model. For example, the risk would be highest where most of the same data used to develop a model (e.g., VDYP) were also used to the check the estimates from the model. Ideally, completely independent PSPs would be used to develop and check models; however, the costs of maintaining two independent sets of plots is likely prohibitively expensive and is not necessary.

In BC, most G&Y models (e.g., TASS, VDYP7, PrognosisBC, and SPS) are developed using data from a wide range of areas and stand conditions. Consequently, the risk of a model projection being largely influenced by the data from any given management unit is low. However, it would be prudent to examine the degree to which data from a given area contributed to a model before using the monitoring data from that area to check the model projections.

It should be noted that other types of modeling data (e.g., experimental plots to determine treatment responses) will still be needed in addition to plots established for monitoring and model development.

5.4 CONSIDER ALTERNATIVE DESIGNS FOR INTEGRATED SAMPLING

For discussion, we proposed an alternative sampling design for addressing the integrated needs of VRI Phase II, monitoring, and model development (Section 3.3). This consisted of using a single standard fixed-area plot design for monitoring and model development data and allowing a variety of TSP designs for VRI Phase II. This alternate design addresses some disadvantages of the CFI design proposed by TIB, but also has it's own disadvantages. The primary trade-off between the two options is cost-effectiveness and flexibility versus simplicity.

5.5 PILOT TESTING

This integrated sampling approach should be pilot tested. Possible objectives for a pilot project include determining how well the various system components work (e.g., planning, plot establishment, and data analysis), identifying possible shortcomings, and suggesting necessary improvements to the proposed systems. These pilots could be conducted in one or two management units. As well, these pilots could test different plot cluster configurations for VRI Phase II to address the issue of high within polygon variation. Possible designs include the fixed-area CMI plot with either small fixed-area (100 m^2) or variable radius auxiliary plots, and the current Phase II IPC with as many count plots (variable radius sweeps) that could be done by a field crew in a single day.

5.6 ALTERNATIVE ESTIMATION METHODS

The Kalman filter estimation methods and model-based inference should also be considered and illustrated, perhaps using the data from the Woodlot CFI (Appendix III) or simulated data. These methods have been used successfully in forestry and other disciplines. There is a need to develop local expertise on these topics.

5.7 MEETING SECONDARY DATA NEEDS

Data from both designs discussed in this paper will help address several of the secondary business needs (Section 2.2). However, we suggest that the management unit level data collection systems are used for passive data collection and are not changed to accommodate secondary needs. An integrated management unit sampling system will not contain enough samples nor will the plots be appropriate to address these secondary needs.

Site Index Adjustment (SIA)

An integrated sampling approach on a management unit will not provide enough samples in suitable stand types to complete an SIA project. However, data from some of the fixed or variable radius plots (that include a 100 m^2 fixed-area subplot) can provide data for this purpose. The only issue is that using data from the fixed area PSPs to develop the SIA theoretically compromises the independence of the PSPs to monitor the site index estimates over time. However, this potential problem is probably not of practical significance if the proportion of PSPs in the overall sample used for SIA is low.

Site Index-Biogeoclimatic Ecosystem Classification (SIBEC)

Data from all suitable PSPs should be used to contribute to the provincial SIBEC database. Data from PSPs in an integrated sampling approach are probably more suitable for the SIBEC database as they will be from randomly located plots as opposed to the current policy of subjectively locating SIBEC plots.

Forest Health Monitoring

The plots established under an integrated management unit level sampling program to meet the three primary business needs, likely will not provide much useful information for forest health monitoring. The main reason is that most forest health issues operate at a larger spatial resolution than can be adequately captured in these plots. Forest health indicators can be measured in the PSPs or TSPs and they will provide some data to help estimate G&Y impacts; however, a more specific and focused program will likely be needed to monitor most forest health issues. This could possibly be achieved by establishing much larger forest health plots centered at plots established as part of an integrated management unit level sampling program.

Linking VRI with Predictive Ecosystem Mapping (PEM)

Plots established under an integrated management unit level sampling program could be used to provide point estimates of site series to check PEM estimates of site series. However, again the observations from this program would have to be supplemented with many more samples to achieve the objective.

APPENDIX I – GLOSSARY OF TERMS

We define the following terms for this report:

Continuous forest inventory (traditional) is a forest inventory in which the same plot locations, usually installed on a systematic grid, are measured over successive occasions.

Change inventory is the process of observing changes and trends over time in the level of the resource and change in land cover classification between two or more time points.

Change projection is the process of predicting the difference in future level or classification of the resource between two or more time points in a management unit.

Change monitoring is an independent check on the projected change or growth in a management unit.

Forest inventory is the procedure for obtaining information on the quantity and quality of trees and stands at a given point in time.

G&Y monitoring is the process of observing the G&Y of a forest and comparing this with the predicted G&Y of that forest. Growth monitoring is a specific type of *change monitoring*.

Model development includes model estimation, validation, and calibration.

Model estimation is the process of deriving basic relationships that comprise the model, (e.g., model coefficients to define growth rates as functions of some measured variables).

Model validation is the process of checking a model to see how well it predicts growth using whenever possible data not used in the model estimation (where possible). It involves comparing projected and actual G&Y to identify any discrepancies.

Model calibration is the physical process of adjusting model relationships or coefficients to improve performance in local conditions that may differ from those on which the model was based or meant.

Monitoring, in general, is the process of observing the status, changes, and trends over time in the levels of the resources and changes in land cover classification between two or more occasions. Growth and Yield monitoring, as defined by the G&Y Monitoring Task Force, is then a special case of the general "monitoring" where the observed growth of a forest is compared with the predicted growth of that forest. Also, change inventory and repeated forest inventories may be regarded as special cases of monitoring.

VRI Phase I: the delineation of polygons from aerial photography and the estimation of resource attributes.

VRI Phase II: the establishment of plot clusters in selected polygons to measure timber, ecological, and/or range attributes.

APPENDIX II – CURRENT PROGRAMS

VRI Phase II

- Typically 100-200 plot clusters established within an inventory unit (e.g., TSA).
- Plots selected using PPSWR
- The Phase II data are needed to:
  - Adjust the Phase I estimates
  - Estimate overall totals and averages, and associated confidence intervals, of various resources
- Average unit cost for a plot cluster establishment is $1,000 - $1,500.

Monitoring

- Monitoring may be needed for:
  - Checking G&Y estimates. Managers need a reasonable level-of-comfort that the G&Y models they are using adequately reflect the actual G&Y achieved on the ground. They want to base decisions on predictions that they can defend. To do this requires feedback to the basic inputs and assumptions in the timber supply analysis. This feedback mechanism includes monitoring.
  - Tracking non-tree vegetation to detect changes in presence or proportion of plant species as an early warning indicator of climatic or other changes.
  - Meeting certification requirements of external agencies such as the Forest Stewardship Council (FSC), where it may be necessary to demonstrate that certain targets are met.
- Monitoring G&Y estimates requires the sample that is being used is representative of the population and independent of any samples used to derive the G&Y estimates.
- Specifically monitoring forest short-term growth projections:
  - Checks for aggregate error due to model implementation, model application, and other unknown sources of error. Model application error can result from incorrect inputs being supplied to the model or model results being extrapolated to situations the model has not been calibrated for.
  - This requires re-measured plots that may be independent of the VRI plots be used to determine the yield at time 1 and other inputs to the model.
- Management unit level monitoring procedures are not formalized but are being pilot tested in various management units.
- Some management unit pilots are testing a uniform fixed grid and others are overlaying fixed-area plots over the VRI TEPs.
- Establishment of monitoring plots at the management unit level is costing a minimum of $2,500 per plot.
- Provincial level monitoring procedures have been formalized.[18]
- Establishment of the provincial monitoring plots is costing approximately $4,000 - $7,500 per plot.

Growth Model Development

- Inventory PSP data concentrated in natural (unmanaged) stands on a matrix.
- The current natural stands PSP matrix is based on BEC zone. Only species that are abundant or common for the BEC zone are considered. Three species groups per species are considered (pure stands, species mixed with conifers, and species mixed with deciduous). There are five, 20-year age classes (except last one) for deciduous and seven for conifers. Within each age class there are three density classes and three site index classes.
- The Silviculturally Treated Matrix that was dropped was stratified by treatment types as follows: Natural + Spacing, Natural + Fertilized, Natural+ Spaced+ Fertilized, Natural+ Spaced+ Pruned, Natural+ Spaced+ Fertilized+ Pruned, Planted, Plant + Spaced, Plant + Fertilized, Plant + Pruned, Plant + Spaced+ Fertilized, Plant + Spaced+ Pruned, and Plant + Spaced+ Fertilized+ Pruned.
- The inventory PSPs are different from the research installations (G&Y experiments, or EPs); these EPs are not discussed further here.
- The cost for PSP measurement in the interior is $2,300, and $2,700 for the coast. These costs include reconnaissance (38%), ecological classification (10%), and measurement (52%). These costs exclude office time.
- Time required to measure a PSP is approximately 10 hours (without the ecological classification that is usually done separately).
- Inventory PSPs can be used for model estimation or validation.
- Data-based (semi-empirical) stand-level models such as VDYP7 use inventory PSPs in model estimation – derivation of model coefficients.
- Models such as TASS can use inventory PSPs for validation – e.g., overlay simulated yields on comparable PSP data – followed by calibration.
- Model validation:
  - May localize relationships or components to a specific management unit or stratum; data is specific to the management unit.
  - It uses data independent of the data used to develop the model.
  - Data used for validation may be representative of the population the model is applied to, or purposively selected.
  - Re-measured Phase II plots provide a representative sample of the population for model validation. They provide feedback on how well the model is projecting that set of plots.
  - Re-measured Phase II plots do not provide a completely independent estimate of the average growth for the management unit to check against the average projected growth.
- There may be benefits to grid or random location of inventory PSPs. These include the ability to estimate actual growth with confidence limits and sample most common types in landbase.

---

[18] B.C. Ministry of Forests. 2001. Change Monitoring Inventory Ground Sampling Procedures for the Provincial Change Monitoring and the National Forest Inventory, Version 1.1 March 2001. MOF, Victoria, B.C. 311 pp.

- Location of inventory PSP on matrix may be futile. Extreme conditions should be generated from research experiments or research case studies.
- Plot size: large plots are needed if spatial attributes are of interest.
- Plot uniformity is not a problem if spatial data collected; prefer not to straddle different stand conditions.
- Inventory PSPs could also be used for G&Y monitoring, especially if the model's dependence on inventory PSPs is low. Therefore, no need for separate programs.
- Plots in treated stands: are pre-treatment data compulsory for model development?

APPENDIX III – EXAMPLE OF A WOODLOT CFI (WOODLOT LICENCE W0038)

CONTINUOUS FOREST INVENTORY

Prepared by Dennis Swensson, Blue Mountain Woodlot Ltd.

Introduction

Blue Mountain Woodlot is located 15 km east of Maple Ridge on the lower slopes of Blue Mountain. The forest cover consists of second growth Western Hemlock, Douglas-fir and Western Red Cedar. Small amounts of Birch, Alder, Cottonwood and Bitter Cherry are also found throughout the Woodlot. The average age of the forest is 45 years.

The Continuous Forest Inventory was begun in March, 1987 at the time of the awarding of the Licence. The initial area of the Woodlot was 374 hectares but it has been increased recently to 400 hectares.

Objectives

The objective for this Management Inventory is two-fold:

- to supply the basic forest resource information needed to complete a Management Plan for the Woodlot; and
- to establish the foundation for a long-term forest resource monitoring and control system for the Woodlot.

Specifications

- A north-south, square grid was placed over the entire Woodlot.
- Ninety-two variable radius plots were established in March, 1987.
- A Basal Area Factor (BAF) of 6.25 (2.5 Relaskop Bars) was used for all plots. The target average tree count was 6-8 trees, live or dead.
- All live and dead standing trees were tallied.
- The minimum diameter was 7.5 cm DBH (diameter at breast height).
- The heights of all trees on the plots were measured.
- Plot centres were marked with a wooden stake. The centre was referenced to two adjacent, live trees. A metal tag, showing bearing and distance to plot centre, was attached to the base of the trees. These measurements were also noted on the tally card.
- Cruise lines between plots were well flagged with blue flagging.
- Two site index trees were measured for breast height age and height on each plot. The largest diameter trees of the leading species were selected in an 8-metre radius from plot centre. In mixed stands, additional site trees of secondary species were measured. This procedure is very similar to current Ministry of Forests' Site Index Tree procedures.

Results of Inventory

The first remeasurement of the 92 CFI plots was made in February 1992 (five years after the initial measurement). The second remeasurement was carried out in October, 1996 (also after 5 more growing seasons).

In 1992, 5 plots had been logged. By 1996, a total of 9 plots had been logged. These logged plots are included in the remeasurement compilations as having no volume.Volumes were compiled to a utilization level of 17.5 cm DBH to a 10 cm top.

The results of the two remeasurements are shown in the following table:

Items	Initial 1987	5-year 1992	10-year 1996	Remarks
Area (ha)	374.4	374.4	374.4	total Licence area
Volume (m^3)	152,612	163,148	176,483	10 years of harvest
Trees/ha	664	640	611	less due to logged plots
Avg. DBH (cm)	29.5	30.3	31.7	larger diameter trees logged
Avg. Ht. (m)	27.6	28.6	30.6	taller trees logged

The results show that the volume of timber on the Woodlot continued to increase even after 10 years of harvesting (over 23,000 m3). The CFI results have assisted the Licensee in obtaining an increase in the Annual Allowable Cut.

The samples will be re-measured for the third time in the winter of 2001- 2002.

Problems Encountered

- The BAF of 6.25 was very suitable for the initial measurement. By the second remeasurement, using the same BAF, the trees had grown considerably in diameter. Due to ingrowth, up to 18 trees per plot were often re-measured on the plots. A fixed radius plot would have eliminated this problem. The Licensee is considering converting the samples to fixed radius plots on the third remeasurement.

- Wooden plot centre stakes have rotted during the past 14 years. Many have been replaced with new wooden or aluminum stakes. The Licensee recommends that aluminum stakes be used in initial plot establishment.

Summary

The Continuous Forest Inventory, while costly to establish and maintain, has proven to be a very valuable tool to monitor the growth of the forest. It has proven that the allowable harvest levels are conservative.

APPENDIX IV – PROJECT TERMS OF REFERENCE (SCHEDULE "A")

File: 500252JVT010GY

Attachment to the Agreement with J S THROWER AND ASSOCIATES LTD for Integrated Sampling Approach – Phase One.

The Contractor shall provide the following Services:

The contractor will work closely with the Forest Inventory and Monitoring Branch (FIMB) to develop simple, cost-effective, and flexible integrated sampling approaches to collect data at the management unit level, to meet several specific business needs. The main specific business needs are:

1. Vegetation Resources Inventory (VRI) Phase II plots needed to adjust the VRI Phase I estimates,
2. Monitoring, and
3. Growth model development.

Other business considerations include operational adjustment site index sampling; SIBEC sampling; VRI/PEM linkages and forest health monitoring.

Specifically, the contractor will first focus on the full continuous forest inventory (CFI) integrated sampling approach, and conduct the work sequentially as follows:

1. Sampling unit:

Examine the feasibility of adopting a single standard plot configuration, similar to the Change Monitoring Inventory (CMI) plot design, to meet the stated main business needs. This will include:

a) Examining the use of the data currently collected in the three separate programs.
b) Recommending a minimum set of variables to be collected on all plots, and if necessary, any additional variables to be collected for special purposes.
c) Recommending approaches to streamline the data collection at a sample location, to allow plot completion in one day on average.
d) Examining the implications on sample size and variances of using a single plot instead of a cluster of plots at a sample location.

2. Sampling approach:

Investigate the suitability of the full continuous forest inventory (CFI) approach, for integrated sampling for forest inventory and monitoring.

In this work, the contractor will consider and discuss issues related to the CFI sampling approach, including:

a) Using a grid for sample selection as opposed to the current VRI PPWSR (probability proportional to size with replacement) approach. This includes consideration of the type of grid and the method of selection.
b) The need for pre- or post-stratification.

c) Selection of sub-samples and remeasurement schedules for monitoring or growth model development.
d) Statistical estimators and variance estimates.
e) Costs of operation (plot installation and remeasurement).

The FIMB may request the contractor to investigate other sampling approaches if the CFI approach is found to be unsuitable. The FIMB will also approve the next Project steps.

Note: In the fulfillment of this project, the contractor is not fettered or bound by previous decisions arising from recommendations from the Vegetation Inventory Working Group, the Peer Review Group, or from other sources, standards or procedures currently in place for the VRI. The contractor is free, however, to utilize such information in the pursuit of this contract.

Further, recommendations arising from this work are not be in the public domain until sanctioned by the Director of the Forest Inventory and Monitoring Branch of the Ministry of Sustainable Resource Management.

Deliverables

This contract focuses on the first part of the overall goals of this project. The deliverables for this first part are:

1. A preliminary discussion paper on the feasibility of the full CFI integrated sampling option for inventory and monitoring, due October 31, 2001.

APPENDIX V – SOME THOUGHTS ON CFI BY W. G. WARREN

Some Thoughts Relevant to the Topic "Common Ground Plot and Continuous Forest Inventory Designs for VRI Phase II Monitoring, and Inventory PSPs".

William G. Warren, Ph.D.
August 2001

Notes on Two- Phase Sampling

Currently, the area to be surveyed is divided into a set of non-overlapping polygons and a Phase I, or photo-interpreted, estimate is made for each polygon. Accordingly, at this level, there is no sampling error per se, the only error stemming from the ability of the photo-interpreter to make an accurate estimate for each polygon. Ground, or Phase II, samples are used to adjust the Phase I estimates. Only a fraction of the polygons are selected for ground sampling and the ground samples themselves also cover only a fraction of their respective polygons. As representative of their respective polygons, the ground samples are, thus, subject to sampling error. They may also be subject to measurement error, although, with proper auditing, the latter should be small and unbiased, i.e. equally likely to be greater or less than the true plot value. In practice, the measurement error cannot be separated from the sampling error and is embodied with it.

The accuracy of the adjusted estimate for the area of interest is therefore dependent on one's ability to estimate the relationship between the ground-sampled and associated photo-interpreted values. Consequently, with the proviso that the ground samples reasonably cover the range of conditions existing in the area of interest, the mechanism of their selection will be relatively unimportant. Both methods that have been used, namely systematic sampling from a sorted list (SSSL) and selection with probability proportional to size with replacement (PPSWR) should be satisfactory, with little to choose between them. One might expect SSSL to be more reliable in attaining coverage of the existing conditions since, under PPSWR, there will always be chance instances where the latter will result in an incomplete coverage of conditions. However, in a CFI framework, with fixed plots and a population that is changing over time, PPSWR may, in the long term, exhibit greater flexibility than SSSL.

The overall accuracy is also contingent on the ground samples being an accurate representation of what is on their respective polygons. There is, as noted above, sampling error, although random location of plots within polygons will unsure unbiasedness. In practice, where centres are placed at the nodes of a grid, total unbiasedness is not attained; the reason for this will be addressed in Section 0 below.

Under both SSSL and PPSWR, with the exception of the larger polygons, there will generally be, at most, one ground-sample plot cluster in a polygon. The plot cluster consists of a central plot and 4 satellite plots, i.e. 5 plots in all in relatively close proximity. In any situation where one is sampling at locations in space there is likely to be some degree of spatial correlation, i.e. locations that are closer in space tend to be more similar than locations that are further apart. Thus, with a plot cluster, we may wonder whether we have 5 "units of information"; for example, given the data from some of the 5 plots, could we make a better guess of what was on the remainder than if we did not have that information? There is no doubt that increasing the number of sample plot clusters in a polygon would give a more accurate representation of what was on the polygon but the cost of doing so would soon become prohibitive. But what about two or more single-plot "clusters", or two or three 2-plot clusters? If the relationship between the ground and photo-interpreted values can be tightened by obtaining a better ground-sample estimate of what is on a polygon, any additional cost of sampling within polygons might well be compensated for by a reduction in the number of polygons required. This is an aspect that, I feel, deserves examination. A proper study of cost cannot be made without knowledge of the degree of spatial correlation. (See also Section 0 below).

Note that it is only when the ground samples (Phase II) are used in conjunction with the photo-interpreted (Phase I) polygons that the sampling design for the ground samples is relatively unimportant. The choice of sampling design would be crucial if the ground samples were to be used on their own. For example, the use of SSSL with a single random start would not, without additional assumptions, permit the estimation a sampling error.

Sampling Over Time

Let me now turn to the question of obtaining estimates of yield over time. Firstly, we must expect a similarity of spatial pattern in successive surveys. This comes about from the fact that, over the area of interest, the forest will, to some extent, be exposed to generally similar environmental conditions, in particular climatic. Admittedly, the longer the time period between surveys the less similar the patterns will likely be, but, given the nature of forest growth, one might expect a strong similarity for periods of, at least, 5-10 years. (The effect of human intervention will be addressed below).

In the extreme case where all possible plots experience the same relative (or absolute) change - say increase by 5% - one sample plot would suffice for the estimation of change (growth). However, it would be foolhardy to use that single plot for an estimate of yield. Although a randomly located plot would provide an unbiased estimate of yield, in the sense that the average over all possible plots that could be selected would be the actual yield, the value obtained for that one plot could be well removed from that average. In practice, things will not be so extreme, but the fact remains that, as long as there is any degree of persistence in the spatial pattern, change or growth will be more efficiently estimated from a collection of fixed plots, relatively few in number, however, there will always be the chance that those plots will give an estimate of yield significantly removed from the true value for the area, and this situation will persist from survey to survey, unless the yield is estimated by augmenting the fixed plots by a changing set of temporary plots.

This then was the motivation for sampling with partial replacement (SPR). ("Sampling with partial replacement *(*SPR) was introduced by Ware and Cunia. (1962) as a further enhancement of CFI for use in extensive forest surveys when efficient estimates of current values and change in those values are of Interest" - Scott and Kohl (1994)). While temporary plots may provide an overall estimate of change, they provide little, if any insight to its cause**,** while the cost of fixed (permanent) plots prohibits the establishment of the number necessary for an acceptably precise estimate of yield. SPR thus appears as a logical compromise. Note that, since over time the spatial pattern will change, there is little advantage in maintaining the same set of fixed plots indefinitely. What is needed are sets that remain common over at least two surveys (cf. Warren (1994) for an example showing the gradual breakdown of spatial pattern over time).

Incorporation of Growth Models

As early as 1979, Dixon and Howitt considered using a growth model in conjunction with continuous forest Inventory (CFI). They state "A timber growth model can be utilized to generate estimates of timber stocks conditioned on past samples before a current sample is taken. Neglecting this source of information ignores the benefits of recursive estimators that combine past information about timber levels and growth with current sample observations". Kangas (1991) observes that "The growth model used [by Dixon and Howitt] is crude; it simply gives the proportional change of the state vector over time. It may be assumed, that the efficiency and usability can be improved, when more detailed models are incorporated". Kangas goes on the say that "Often forest Inventory and forest management are treated as separate problems. The data measured should be used *as* efficiently as possible. Thus, the same data should be applicable for both objectives. In the approach of Dixon and Howitt (1979) information about all control actions (harvests, cultural treatments, etc.) is used in the forest inventory. However the use of these inventory results in forest management systems is limited, because only means and sums are known. For most management systems more detailed information is needed. The required, information can be obtained when treewise growth models are used for updating the data".

The formulation of both Dixon and Howitt and Kangas utilizes the Kalman filter. (The SPR estimator of Ware and Cunia can be shown to be a limiting case of the Kalman filter - Dixon and Howitt (1979)).

This is not the place to become too technical, but an outline of the procedure may aid in understanding its nature. Following Kangas, let the yield at time t for s timber classes (e.g. species) be denoted by the vector x and the control actions (human intervention) be denoted by the vector u. The model of yield evolution through time can be written as

$$x_{t+1} = A_t x_t + B_t u_t + e_t$$

The subscript t on the matrices A and B indicates that these matrices can be time varying. Matrix A gives the proportional change of the state vector x over time (can be derived from a growth model) and the matrix B gives the impact of the control actions. The error vector e represents the fact that even though A_t, B_t, x_t and u are known, x_{t+1} cannot be predicted perfectly. It is also assumed that x_t is never observed perfectly; what is observed is y_t where

$$y_t = C_t x_t + w_t$$

where C_t is a known matrix and w_t is again a random error vector.
If $x_{t|t}$ is the estimate of yield given all the information up to and including time t then the yield at time t+1 is predicted as

$$x_{t+1|t} = A_t x_{t|t} + B_t u_t$$

A sample is then taken to observe y_{t+1} and it can be expected that

$$y_{t+1} = C_{t+1} x_{t+1|t}$$

but the prediction will almost always involve error so a residual vector is defined as

$$v_{t+1} = y_{t+1} - C_{t+1} x_{t+1|t}$$

If $v_{t+1} = 0$, the predicted and observed values are identical.
Since this is unlikely, an updated estimate is formed by combining the predicted and observed values, thus

$$x_{t+1|t+1} = x_{t+1|t} + K_{t+1} v_{t+1}$$

where K_{t+1} involves not only the C_{t+1} but also the variance of the error vectors e and w. The Kalman filter can be seen to be an empirical Bayesian procedure.

In simplistic terms, given the yield, or an estimate thereof, at time 1, we may use a growth model to project the yield at time 2. When the data at time 2 become available, they may be incorporated with the projected value to obtain a better (i.e. smaller mean square error) estimate of the yield at time 2.

With fixed plots and an individual-tree growth model, projections could be made on a plot-by-plot basis, however, unless one believes that the plot data are subject to appreciable measurement error, there would seem to be no reason to incorporate the projected values at this level. In contrast, since the overall estimate at time 2 is subject to both sampling and measurement error, incorporating the growth-model projection with the time 2 estimate at this level may well result in an improved estimate.

The tacit assumption here is that there is no systematic difference between the projected yield and that estimated at time 2. A discrepancy between the projected and estimated values, the vector v in the above formulation, in excess of what might be expected would suggest that the growth model is inappropriate and/or that the sample estimate is, by chance, well removed from the true value (or that the sample estimate is defective for some other reason). The point is that the incorporation of the projected yield with the time 2 estimate should not he considered until the cause of any such discrepancy has been determined. However, when we can reasonably believe that these are two independent estimates of the same thing, we would surely be remiss in not combining them. The approach provides a means for checking the viability of the growth model.

SPR and the Incorporation of Growth Models

A variation on sampling with partial replacement consists of dividing the entire set of plots into a number of groups. In some variations, in the initial year, all plots would be measured, but in all variations, at subsequent times only a few groups are remeasured, although commonly there will beat least one group common in successive remeasurements; see Scott at al. (1990) from which the example in Figures 4-6 is drawn. Scott at al. observe that, with such systems, "plot and tree data assessed in different years must be updated to a common point in time. This can be done by modifying the CFI and SPR estimators, or by introducing methods such as the Kalman filter or Mixed Estimator (ME), which can incorporate updating and projections using growth functions". There are indeed, a variety of methods that can be used for imputing yields for those groups not re-measured at any time. These range from simply using the ratio of the plots in common to growth-model projections. Scott et al. remark that "One clear design recommendation is not to use a rolling survey without having a group of plots to update all plots to a common year periodically". As above, growth models should be checked and used in updating only if the projected yields were essentially in agreement with those estimated directly.

An advantage of this type of design is that the sampling effort, and thus the costs, can be spread over a number of years. Scott et al. express it this way; "From an organizational point of view, annual assessments are superior to periodic assessments in that a (constant) proportion of the total number of sample plots is visited each year". For this reason alone the approach would seem to merit further consideration.

Sampling Spatially Distributed Individuals

The first step in the sampling of a finite population is the definition of a frame; the population must be divided into parts (called sampling units or simply units) which cover the whole population and must not overlap - i.e. every element in the population must belong to one and only one unit (see e.g. Cochran, 1963). In our context, the collection of polygons provides a frame from which we may select a simple random sample of polygons, a sample selected with probability proportional to size with, or without replacement, a systematic sample selected, with a random start, from a sorted list, for which the sorting could be by polygon size or any other definable characteristic.

We next have to select samples within the chosen polygons, and the choice of a frame at this level is not so straightforward. Ideally, the polygon could be partitioned into a collection of non-overlapping plots, and one or more of these plots be chosen at random and measured. But this is not what appears to be done. Plot centres are limited to the nodes of a grid and it is impossible to totally cover any region with non-overlapping circles. Yandle and Wiant (1991) express it this way; "The creation of a sampling population of mutually exclusive plots when the plot shape is any geometric figure, such as a circle, that does not tessellate the plane, in order to use SRS [simple random sampling] will necessarily result in an unsatisfactory definition of the sampling population because parts of the area, and the population of interest that is contained therein, will be excluded". In our context there will, in general, be areas within a polygon that have no chance of being included in the sample; thus, the estimate of the yield for a polygon, will be, of necessity, biased. Fortunately such bias should be small and, when averaged over a number of polygons, possibly negligible. Nevertheless, this does illustrate a potential limitation of design-based inference.

There is an alternative approach, which in some respects, seems better suited to the estimation of spatially distributed quantities. Let x(s) be the value of the property of interest at location s. In the design-based system, the x(s) are regarded as fixed, while the sample locations, s, are randomized.

Expectations are taken over all possible combinations of sample locations, of which there will be $\binom{N}{n}$ when sampling at n out of a possible N locations. To emphasize this we may write x($\underline{s}$). In the alternative, model-based system, the set of sample locations is regarded as fixed, but the observations are regarded as having been generated by some spatial process; expectations are taken with respect to the assumed underlying process. We may then write x(s) (cf. Brus and de Gruijter, 1993).

In the design-based system there is no necessity that the observations be i.i.d. (independent, identically, distributed) random variables; the sampling distribution is generated by the randomization process. Spatial correlation is, then, irrelevant. In the model-based system the observations are treated as random variables, but there is no requirement that they be independent; indeed, the approach involves the estimation and utilization of the correlation structure.

In contrast to the design-based approach, the model-based system permits the spatial interpolation of the quantity of interest, i.e. its estimation at any location in the area (with a measure of its uncertainty) and, in particular, the estimation of the mean and its standard error from a single spatially-systematic sample, There is, however, no free lunch. Assumptions have to be made about the nature of the underlying stochastic process, commonly that it is second-order stationary and isotropic. Such assumptions are often difficult to verify but, if violated raise questions about the viability of the resulting estimates. Notwithstanding, since the approach seems more compatible with the way that the forest is sampled, it merits further consideration. For existing forestry applications we may cite Duplat (1976) and Mandallaz (1993).

Discussion

Many statistics courses begin with, and never get beyond, the assumption of i.i.d. (independent, identically distributed) random variables. This is unfortunate since real-world data commonly involve spatial and/or temporal correlation. This is certainly the case with any investigation of forest growth and/or yield.

Lack of independence, when recognized, has usually been seen as an impediment and methodology, for example, randomization, was developed to get around the problems consequent on not having independent observations. There has been, however, an increasing body of work direct to improving estimates and, subsequently, actions and decisions, by attempting to evaluate and exploit the spatial and/or temporal associations. Thus, we have seen the development of spatial analysis (geostatistics, kriging) and expansion from mining to applications across the whole spectrum of physical, biological and social sciences. Similarly for methodology for studying relationships that evolve over time, e.g. the Kalman filter, the early use of which was primarily in economics.

Since spatial relationships are rarely static but changing with time, the methodologies are beginning to coalesce. Thus, the cover of the Proceedings of the 1996 Conference on Modelling Longitudinal and Spatially Correlated Data, held on the island of Nantucker (Gregoire et al. Eds., 1997) notes that "This volume focuses on the statistical treatment of continuous and discrete data measured at different points in time, locations in space, and/or across combined spatio-temporal dimensions. Linear, nonlinear, and generalized linear models and methods are presented, as are new developments to handle messy data. The volume provides an examination of the historical development of approaches to model spatially and temporally correlated data and the ongoing convergence of these methods". The keynote paper by P.J. Diggle is titled "Spatial and Longitudinal Data Analysis; Two Histories with a Common Future?".

In some parts of the world, these methods have already been adapted for the estimation of forest G&Y. It is, I think, inevitable that they will ultimately be used in British Columbia. It is time that they be given consideration and to start the build up of the expertise that will inevitably be needed.

References

Brus, D.J. and de Gruijter, J.J. (1993). Design-based versus model-based estimates of spatial means. Theory and application in environmental soil science. Environmetrics 4:123-152.

Cochran, W.G. (1963). Sampling Techniques, 2nd. ed. Wiley, N.Y.

Diggle, P.J, (1997). Spatial and longitudinal data analysis: two histories with a common future? In Modelling Longitudinal and Spatially Correlated Data, Methods, Applications and Future Directions. Gregoire, T.G. et al. Eds., see below.

Dixon, B.L. and Howitt, S.E. (1979) Continuous forest inventory using a linear filter. Forest Science 25:675-689.

Duplat, P. (1976) Point de vue forestier sur la theorie des variables regionalisees. Section Technique Office National des Forets, Doc. 76-3.

Gregoire, T.G., Brilliager, D.R., Diggle, P.J. Russek-Cohen, E, Warren, W.G. and Volfinger, R.D. (Editors). (1997). Modelling Longitudinal and Spatially Correlated Data, Methods, Applications and Future Directions. Springer, N.Y,

Kangas, A. (1991) Updated measurement data as prior Information in forest inventory. Silva Fennica 25:181-191.

Mandallaz, D. (1993). Geostatistical methods for double sampling schemes: application to combined forest inventories. Thesis, Chair of Forest Inventory and Planning, Swiss Federal Institute of Technology, Zurich.

Scott, C.T. and Kohl, M. (1994). Sampling with partial replacement and stratification. Forest Science 40:30—46.

Scott, C.T., Kohl, M. and Schnellbacher, H.J. (1990). A comparison of periodic and annual forest surveys. Forest Science 45: 433-451.

Ware, K.B. and Cunia, T. (1962). Continuous forest inventory with partial replacement sampling for growth components. Forest Science Monograph 3.

Warren, W.G. (1994). The potential of sampling with partial replacement for fisheries surveys. ICES Journal of Marine Science 51:315-324.

Yandle, D.O. and Wiant, H.V. Jr. (1981). Comparison of fixed-radius circular plot sampling with simple random sampling. Forest Science 27:245-252.

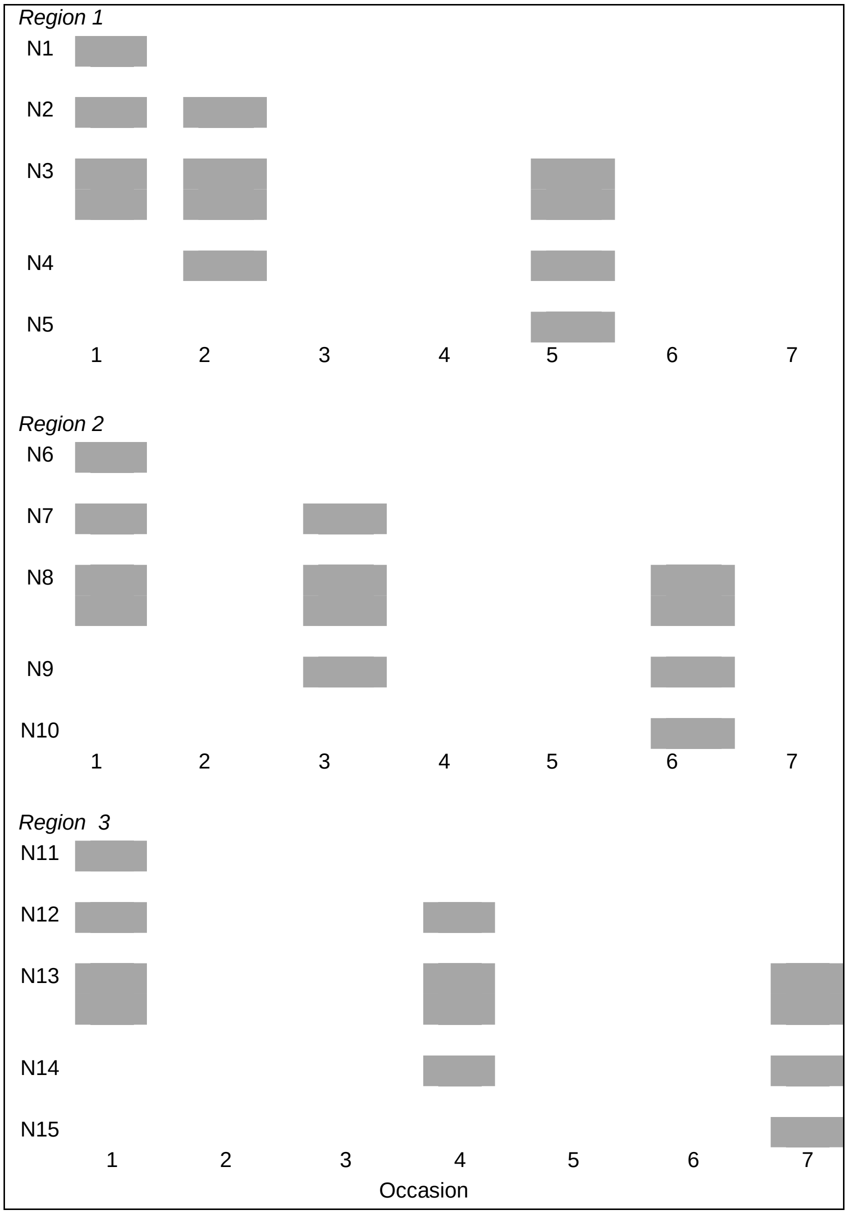


Figure 3. Rotating panel options from Scott el at. (1990)[7] (part 1).

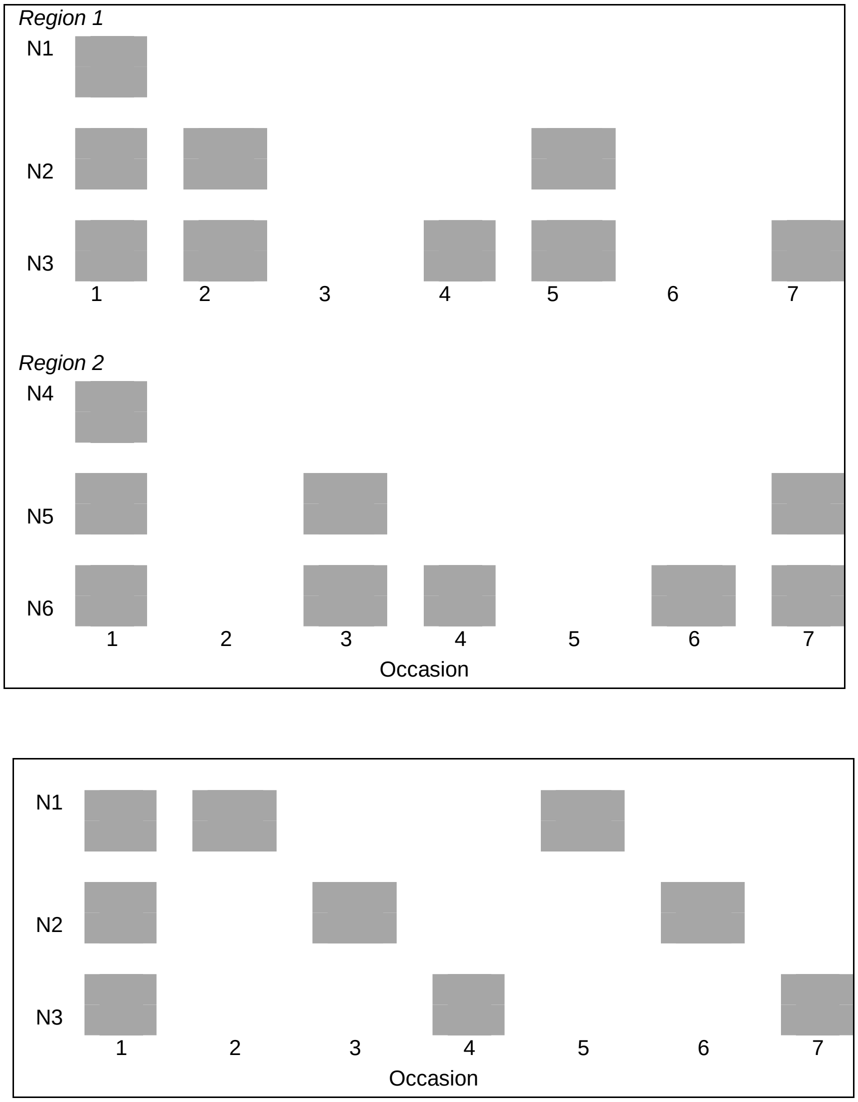
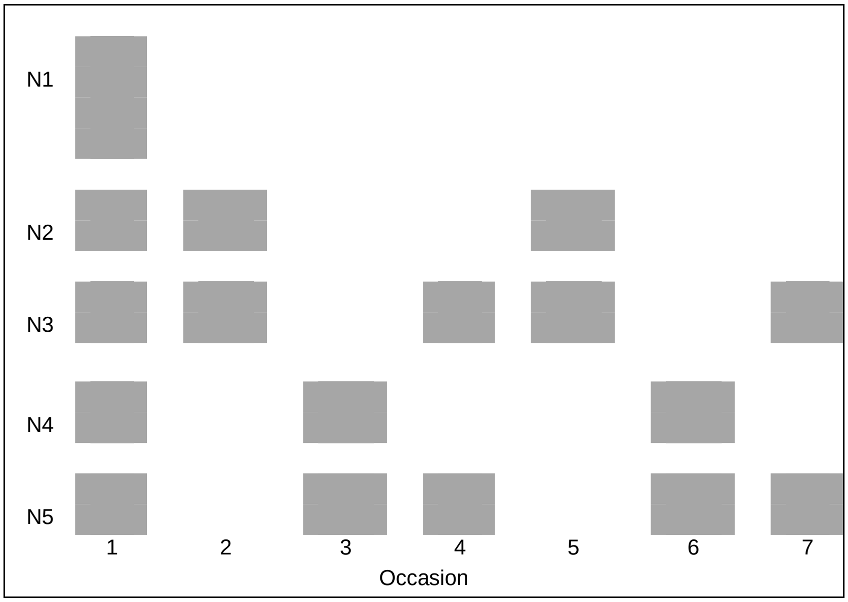


Figure 4. Rotating panel options from Scott el at. (1990)[7] (part 2).



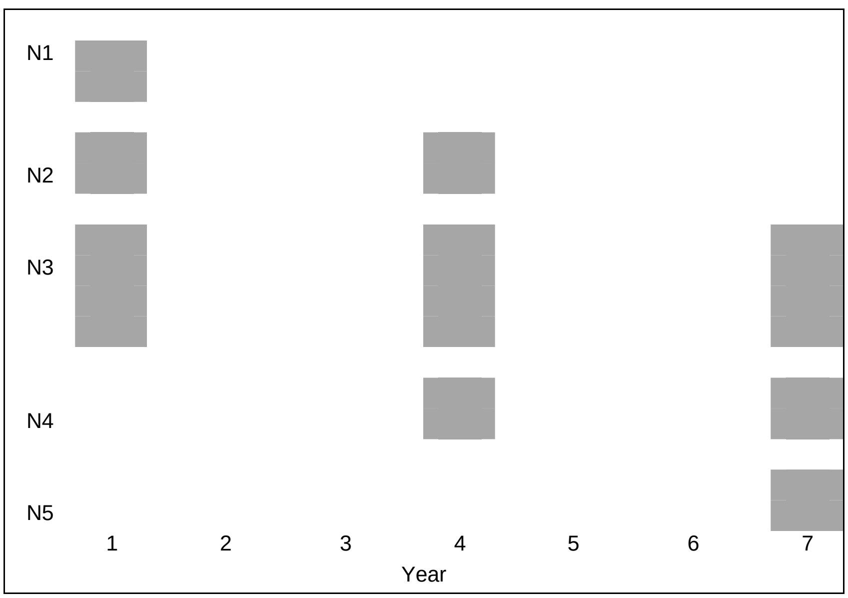


Figure 5. Rotating panel options from Scott el at. (1990)[7] (part 3).

In Figures 4-6, the N*i* refers to the *i*th group of plots, and Region *k*, in British Columbia's case, could refer to the *k*th management unit within a Forest Region. The shading refers to the Occasion or Year a particular group of plot is measured.

TRACKING FORM

Project:	**MSI-002**		
Due Date:	**2001-10-15**		
Manager:	**A.Y. Omule**	**Writer:**	**A.Y. Omule**
Leader:		**Editor:**	
Reference:			

Report History

Date	Notes
Oct 1-01	Jim T - some formatting.
Oct 5-01	Jim T – mods to section 1 and 2.
Oct 6-01	Jim T – more work on sec 2 and 3.
Oct 7-01	Jim T – more clearing. Added three main questions to intro. Removed "BS" acronym and changed to "variable radius plot" (this may not make sense in some places). Replaced manual references to sections, table, and figures with embedded cross references.
October 13	Hamish does a "quick" edit. Sends to Eleanor to review and add Exec Summ and Recommendations
Oct 14	Eleanor re-writes executive summary, section 4.3.2, and recommendations. Makes minor corrections in rest of text.
Oct 15	Jim T – review of paper and eleanor's additions. Some minor editing.
Oct 17/18	EM – full rewrite
Oct 17	Jim T
Oct 19	Jim T
Oct 19	EM - Minor edits

Remove Document Tracking Sheet before final print or creation of e-copy!