Statistical Estimation Methods for Provincial Change Inventory

Version 2.0

Prepared for
Jon Vivian, RPF
Resources Inventory Branch
Ministry of Forests
P.O. Box 9516 Stn Prov Govt
Victoria, BC V8W 9C2

Project: MFI-401-055

March 31, 2000

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

Executive Summary

This document describes several statistical methods for estimating change and yield from ground plots. In this report we assume that a fixed set of 314 ground plots are selected with equal probability from a list of 2,367 photo plots on a 20x20-km grid over the entire province. To estimate annual yield, all 314 plots will be measured in the base year (Year 0) and approximately 10% will be re-measured each year following (rotating panel). Change will then be estimated as the difference in yield between two time points (change estimates for very short time periods [e.g., 1 year] may not be desirable as they would have relatively high variances). The main advantages of a rotating panel approach are that yield estimates are current and costs are distributed over time.

The estimation methods for mean or total and their associated variances, are (in increasing degrees of complexity):

1. Estimates based solely on units measured in year *k.*
2. Estimates determined using the ratio between base year and year *k* measurements.
3. Estimates based on adjusted growth model predictions.

These estimates are provided under the simple random sampling (SRS) and stratified random sampling methods of selecting plots to be re-measured. Under the assumed annual sampling plan described above, a ratio estimator with simple random sampling and a combined ratio estimator in stratified random sampling would be the best choices for estimating yield or change.

The estimated mean and variance for the 314 ground plots can be scaled up to the population of 2,367 photo plots and to the province. The estimated mean of the photo plots or of the province is the same as that of the ground plots. To scale up the variance estimate to the population of photo plots (or to the province, assuming the photo plot's systematic sample is equivalent to a SRS), we add a component of variance in the case of yield, or multiply by a finite population correction factor in the case of change.

This document also demonstrates with a simulated example how yield and change in averages or totals and their associated variances for ground measured continuous attributes could be calculated.

Table of Contents

List of Tables

1. INTRODUCTION

1.1 TERMS OF REFERENCE

J.S. Thrower & Associates Ltd. (JST) prepared this report for the Ministry of Forests (MOF) Resources Inventory Branch (RIB). Our team includes Ian Cameron, *MSc, RPF*, Eleanor McWilliams, *MSc*, *RPF*, A.Y. Omule, *PhD, RPF*, Christie Staudhammer, *MSc*, Guillaume Thérien, *PhD*, Jim Thrower, *PhD, RPF*, and Bill Warren, *PhD*. This paper has been reviewed by the Expert Review Panel (ERP) that includes John Barker, *PhD, RPF*, Jim Flewelling, *PhD*, Kim Iles, *PhD*, Peter Marshall, *PhD, RPF*, and Don Munro, *PhD*, *RPF*. Mike Bonnor, *PhD* (Executive Secretary of the Growth and Yield Monitoring Task Force) has also reviewed it.

1.2 BACKGROUND

In 1998 the MOF awarded a two-year contract to JST to develop growth and yield monitoring protocols. The second year (2000) of this contract (Schedule A) involves three tasks to be completed:

1. Conceptual sampling methods – Develop and recommend conceptual sampling methods to estimate change[1] at the provincial level and monitor estimates of change at the management unit level. This task was completed in December 1999 and the report was submitted to the RIB.[2]
2. Detailed procedures – In a detailed report, describe estimators, procedures, standards and field protocols, based on approved conceptual methods. Demonstrate the use and application of the recommended methods.
3. Pilot strategic sampling plan – Complete a strategic sampling plan for a pilot study of the change inventory design at the provincial level, including the study objectives and desired products, and a general strategic direction for implementing the pilot project.

This report, in part, addresses Task B. The statistical estimation procedures for the provincial change inventory are described. The remaining deliverables for Task B, including statistical estimation procedures for monitoring estimates of change at the management unit level and detailed field procedures for both provincial and management unit level, and the pilot strategic sampling plan (Task C), are reported separately.

1.3 DOCUMENT OBJECTIVES

This document provides several statistical estimation methods for estimating change and current yield to meet the National Forest Inventory (NFI) goals and the provincial reporting needs. It

[1] Definition of change and other terms are given in Appendix IV.

[2] J.S. Thrower and Associates Ltd. 1999. Conceptual sampling methods for change inventory and monitoring of vegetation resources, version 2.0. Unpublished report. December 13, 1999.

also demonstrates how a provincial average or total and their associated variances for ground measured continuous attributes could be determined.

2. GROUND SAMPLING METHODS

2.1 OVERVIEW

This section summarizes the ground sampling methods for the provincial change inventory as outlined in the earlier report titled, *Conceptual sampling methods for change inventory and monitoring of vegetation resources version 2.0*[2]. This section will provide background reference for the statistical estimation methods to be developed.

2.2 GOAL

The goal of the provincial-level ground sampling is to:

> *Estimate current yields and change over time of timber and non-timber attributes to meet national and provincial reporting needs.*

The timber and non-timber attributes of interest are those defined by VRI ground sampling methods and include the NFI attributes.

2.3 GROUND SAMPLING OBJECTIVES

The objective is to install and re-measure ground sample plots province-wide to provide change and current yield data for: [3]

1. *National reporting of the Criteria and Indicators (C&I) of sustainable forest management by ecozone, and other NFI goals (Canadian Forest Service, CFS).*
2. *Provincial annual reporting of vegetation resources status and change (MOF Annual Report).*

2.4 TARGET POPULATION

The population of interest is the BC landbase (approximately 95.2 million ha) including Vegetated and Non-Vegetated (BC Landcover Classification Scheme, [BCLCS]) areas. The target population for ground sampling is the Vegetated Treed (VT) areas, Vegetated non-Treed areas with the potential to grow trees, and areas where tree cover has been temporarily removed. According to the NFI[4], the Non-Vegetated areas, and Vegetated non-Treed areas that

[3] The inventory information required to meet national and provincial needs will be obtained from the provincial 2,367 2x2 km photo plots. Data from the photo plots may be obtained from new photography or existing inventories. The ground plots will provide information not available from the photo plots and will not be used to adjust photo estimates, (i.e. use them as the second phase in a two-phase sample). Furthermore, for provincial reporting, specific objectives outlining which data will be used from the ground plots have yet to be defined.

[4] A plot based National Forest Inventory design for Canada, An interagency partnership project, version 2.0, March 1999. Canadian Forest Service, Victoria, BC.

are unlikely ever to have tree cover are excluded from ground sampling. However, their landcover classification will be ground-checked.

2.5 Sample Location

The NFI 20-km grid has been placed Canada wide creating 2,367 photo locations in BC. These photo locations will be sub-sampled and approximately 314 ground plots will be established within the target population for sampling (section 2.4). Choices for ground sampling methods include simple random sampling, systematic sampling from a sorted list, and stratified random sampling by terrestrial ecozones.[5] The provincial objectives, reconciling provincial and federal objectives, and determining sample selection probabilities still need to be clarified.

This report assumes that ground plots will be selected with equal probability from the list of 2,367 photo plots. Further, it is assumed that all 314 ground plots will be established and measured in a single year (Year 0).

2.6 Annual Sampling Strategy

According to NFI design, plots will be re-measured annually to estimate change based on a rotating panel. The re-measurement cycle has not been confirmed but we assume that the ground plots will be re-measured over a 10-year period. Approximately 31 units per year for six years, and approximately 32 units per year for the remaining four years, will be selected at random, without replacement, from the 314 units for re-measurement. The main advantages of this rotating panel approach are that yield estimates are current, and costs are distributed over time.

2.7 Field Procedures

Data definitions, standards, measurements, and field cards for the initial measurement and re-measurements will be based on the VRI ground sampling procedures manual. However, a fixed area plot will replace the variable radius plots for large trees (≥ 4.0 cm dbh), polygon boundaries straddling plots will be mapped, and there will be no auxiliary plots. The purpose of the mapping is to calculate the area for the land types sampled on a plot. As well, the plots will be permanent and inconspicuous, and subject to operational activities. The new VRI procedures for establishing plots in parks and private land will be used. All trees will be stem mapped (therefore identifiable) such that diameter and height growth on each tree are known. Details of the field procedures modifications are described in the report titled *Vegetation Resources Inventory Change Measurement: Preliminary Field Procedures* (not yet completed).

---

[5] Ecological Stratification Working Group. 1995. A national ecological framework for Canada. Agriculture and Agri-Food Canada, Ottawa/Hull.

3. STATISTICAL ESTIMATION

3.1 Summary of Relevant Assumptions

We summarize assumptions for the statistical estimation of change and current yield as follows:

1. The 314 ground plots are selected with equal probability from the list of 2,367 photo plots.
2. Under the assumed annual sampling regime, approximately 1/10 of the 314 units (N=314) will be measured annually, with all units measured in the base year (Year 0).
3. There will be a sample size of approximately 31 plots per year, from which inferences about the N locations can be made. The N locations form the fixed population, for which we estimate a mean (or total), or change in the mean (or total) with associated standard errors.
4. The annual estimates for a population of 314 selected units will be assessed. In the base year there is no sampling error (the whole population has been sampled). There may be measurement error but this is assumed to be negligible. (Measurement error could be formally included in the estimation if necessary. However, it may be impossible to separate measurement error from natural variability).
5. The observations from the photo plots are not used in the estimation of yield or change.[3]
6. The estimators are needed at the provincial level, not by ecozone.

3.2 Overview of Estimation Methods

In all the estimation methods outlined in the following sections, the quantity y_k can be regarded as either the value of y for the 314 plots at time k (e.g. yield) or the measure of change between time k and the base year. The quantity y at the base year, y_0, is always yield.

Following Scott et al.[6], there are three estimation methods for the mean or total and their associated variances for the 314 plots (increasing degrees of complexity):

1. Estimates based solely on units measured in year k.
2. Estimates determined using the relationship between base year and year k measurements.
3. Estimates based on adjusted growth model predictions.

[6] Scott, C.T., M. Kohl, and H.J. Schnellbacher. 1999. *A comparison of periodic and annual forest surveys.* For Sci. 45 (3):433-451.

The specific form of the variance estimator for each method will depend on how the annual sample is selected from the 314 plots for re-measurement. Three annual sampling selection methods are considered: simple random, stratified random, or systematic sampling.

The estimated mean and variance for the population of 314 ground plots can be scaled up to the population of 2367 photo plots and to the province. The estimated mean of the photo plots or of the province is the same as that of the 314 plots. To scale up the variance estimate of the annual sample to the population of photo plots, we will add a component of variance in the case of yield, or multiply by a finite population correction factor in the case of change. The yield variance component can be estimated from the measurements in Year 0 with the estimate depending on how the 314 plots were selected from the photo plots. If, for example, the 314 ground locations are selected as a stratified random sample from the 2,367 photo locations, then the variance component is calculated for that stratified random sample of 314 out of 2,367.

To scale up of the variance estimate of the annual sample to the province, we need to make some assumptions since the 2367 photo plots are a systematic sample of the province. For example, it might be possible to assume that the systematic sample was equivalent to a simple random sample. In this case, the variance estimate for the provincial mean and change would be identical to that for the photo plots.

The scaling up of the variance estimates is not discussed further in this report as the methods for selecting the 314 plots from the 2367 photo plots, and the assumptions about the systematic photo sample, have not been decided by the CFS and the MOF.

The concepts of the three estimation methods are outlined below. Detailed formulae for the estimates and their variance estimates are provided in Appendix I. Examples of calculations are provided in Appendix II.

3.3 ESTIMATES BASED SOLELY ON UNITS MEASURED IN YEAR k

3.3.1 Simple Random Sampling

Yield estimates in a specific year (year k) are formed from those units measured in year k. The change between years k_1 and k_2 is estimated directly from units measured in both k_1 and k_2. This implies that for change, either $k_1 = 0$ (the base year in which all 314 units are measured) or that $k_2 = k_1 + 10$. However, the change between any two years, k_1 and k_2, can be indirectly estimated by the difference between year k_1 and the base year and k_2 and the base year, that is, $(k_2 – k_0) – (k_1 – k_0)$.

In their example, Scott et al[6], assume that the average, at time k, of the plots in each partition or group has the same expectation. If, in our situation, the units are randomly selected (i.e. each has an equal chance of selection from the units remaining), any estimates obtained as outlined above will be unbiased. This applies to the conventional estimates of variance as well as that of the population mean or total. Section 4.2 (Appendix I) has a detailed description of the estimates based on simple random sampling.

3.3.2 Stratified Random Sampling

The population of 314 units may be sorted under some criterion (perhaps based on the Year 0 measurements), and the ordered list divided into 15 strata of 20 units each. In Year 1, two units are chosen at random from each of the strata and form the first group. In Year 2, two more units are chosen at random from the remaining 18 units in each stratum, forming the second group. In Year 3, two more units are chosen at random from the remaining 16 units in each stratum, etc. Group 1 would be measured in Years 1, 11, 21, etc., while Group 2 would be measured in Years 2, 12, 22, etc. This ensures that each group will represent the population (of 314) units as a whole, and the groups will be similar in both mean and variance.

Unbiased estimates of current yield in year k and change between years k_1 and k_2 can be obtained for stratified sampling in a manner analogous to that discussed in Section 3.4.1. Section 4.3 (Appendix I) provides a detailed description of the estimators based on stratified random sampling.

3.3.3 Systematic Sampling

Unbiased estimates of current yield and change could also be obtained by systematic sampling using a random start from the ordered list. Such estimates should have precision comparable to that of stratified random sampling. Sample-based estimates of precision can be obtained if more than one systematic sample is taken. The averages from the systematic samples are treated as observations and used to estimate precision based on simple random sampling, with or without ratio estimation.

3.4 Estimates Determined Using the Ratio of Base Year And Year K Measurements

3.4.1 Simple Random Sampling

The estimation methods in Section 3.4 generate unbiased estimates, but do not make use of all the available information. It is more reasonable to assume that the observations in year k are related to the observations in year $k-1$. For example: $y_{ki} = f(y_{0i}) + e_{ki}$, where y_{ki} refers to the i^{th} observation at time k.

If $y_{1i} = a + by_{0i} + e_{1i}$, the parameters, a and b, could be estimated from those units measured at both $k = 0$ and $k = 1$, and applied to the remaining units (i.e. regression estimation). More complex relationships could be assumed, with the number of parameters being limited by the number of sample units in a group. If, in the above, $a = 0$, the approach reduces to ratio estimation. Indeed, ratio estimation may give reasonable results when $f(y_{0i})$ remains unspecified. In general, ratio (and regression) estimates are somewhat biased, though such bias may be more than compensated for by the increase in precision.

A detailed description of ratio estimation with simple random sampling is available in Section 4.4 (Appendix I).

3.4.2 Stratified Random Sampling

The ratio (or regression) estimation can also be applied in stratified random sampling. A stratified random sample ensures that the range of initial conditions is represented in each group, thus aiding in parameter estimation. Therefore, any relationship estimated from the measured group (at time *k*) should be reasonably applicable to the non-measured groups.

A more detailed description of ratio estimation using stratified random sampling is available in section 4.5 (Appendix I).

3.5 Estimates Based on Adjusted Model Predictions

Substitution of a known (or assumed) growth model[7] for the relationship between observations in year *k* and observations in year *k–1* is a process similar to those outlined in section 3.5. The main difference is that the parameters of the relationship are, or are assumed to be, known rather than estimated from the data. Growth models could be used to project values for all groups. The projected values could then be adjusted on the basis of the comparison between the projected and measured values of the measured group. For example, let $\tilde{Y}_j$ be the projected value for group j and assume that group k was measured with the outcome $\hat{Y}_k$. We may then take:

$$\hat{Y} = \frac{\tilde{Y}_j \hat{Y}}{\tilde{Y}_k}$$

This estimation method is not considered further at this time. However, it should be more fully explored as time and finances permit since the adjusted model predictions may prove to be superior to ratio estimation.

[7] It is not necessary to have the growth model, but something to make a reasonable projection that is modified on the basis of the data. Of course, the more accurate the projection, the better the final estimate.

3.6 DISCUSSION

3.6.1 Estimation Methods

Several methods have been discussed for estimating change and current yield. These can be viewed as specializations of the Kalman filter (Appendix III). The Kalman filter takes into account the relationship between successive values of the population mean. More complex and, perhaps, more realistic versions could be considered but would likely be more sophisticated than demanded by the present situation. Nevertheless, as has been demonstrated by Dixon and Howitt (1979)[8] and Kangas (1991)[9], there is potential for the use of the Kalman filter in continuous forest inventory. In particular, the Kalman filter approach can be used to separately estimate changes due to harvesting or fire (Appendix III).

The choice of method depends on several factors, including the sample selection strategy for annual sampling, desired precision, and desired degree of sophistication. Under the assumed annual sampling plan described in this report, the ratio estimation under simple random sampling and the combined ratio estimator in stratified random sampling would be the best choices. These two estimation methods provide very precise estimates with little bias, and there does not seem to be much to choose between them.

The estimates of change discussed in this report were *net change* including mortality. It may be desirable to estimate change of survivor trees separately from the mortality estimate, to get a more precise and interpretable estimate of change. Formulae that can estimate annual change have been provided. However, change estimates involving very short time periods (e.g. 1 year) may not be desirable. Such estimates would have high relative variances (although small in absolute terms), especially those estimates based on simple random sampling.

The estimation methods described are for estimating overall provincial statistics of change and current yield. If provincial estimates are needed by ecozone, then there would be five populations instead of one, with the methodology described in this report applied separately to each. However, there would be a need to increase the size of the annual sample to obtain reasonable estimates by ecozone. For the NFI, the BC sample will be combined with those of other provinces to get national statistics by ecozone.[10]

---

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

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

[10] It is unclear how the CFS will achieve this task.

3.6.2 Assumptions

It is preferable to choose the 314 ground plots from the 2,367 photo plots with equal selection probabilities. The MOF may decide to select these plots with unequal probabilities, however, we strongly recommend against it. These unequal selection probabilities will have to be accounted for in the variance component for scaling up to the 2,367 photo plots.

We have assumed that the population of the 314 ground plots is fixed. This assumption is plausible since it is expected that the original target population for ground sampling (Section 2.4) will not change. If changed, the estimation techniques will have to be modified if additional samples are allowed into the original sample.

The base (Year 0) measurements should ideally all be done in the same year. The ratio estimation approach can be extended to cope with multiple starting years but with added variability

3.7 RECOMMENDATIONS

We recommend that:

1. The ratio estimation method (Section 3.4) under simple random sampling or the combined ratio estimation under stratified random sampling be considered for estimating yield or change of the ground plots. (This stratification is for the 314 ground plots, not the photo plots).
2. To make yield or change estimates by ecozone, treat the ecozones as separate populations and apply the recommended approach to each.
3. The variance formulae we have provided should be modified to scale up the variance estimates for the population of photo plots and for the province (after the CFS and MOF decide on ground plot selection from the photo plots and on assumptions about the systematic photo sample).
4. Use of the Kaman filter (Appendix III) and the adjusted model prediction methods (Section 3.5) should be fully explored if time and funding permit.

4. APPENDIX I – DETAILED ESTIMATION FORMULA AND EXAMPLES

4.1 OVERVIEW

Five estimation methods are discussed:

1. simple random sampling,
2. stratified random sampling,
3. ratio estimation under simple random sampling,
4. ratio estimation under stratified random sampling (separate ratio estimator), and
5. ratio estimation under stratified random sampling (combined ratio estimator).

In the descriptions below, the quantity y_k can be regarded as either the yield at year *k* or the change between year *k* and the base year (Year 0). In either case, the annual change between year *k* and *k-1* can be estimated as:

$$y_k - y_{k-1}$$

The variance estimates presented assume the 314 plots represent the population of interest. Variance estimates for yield at year *k* or change between year *k* and Year 0 will be equal, as we assume no sample variance at Year 0 (all units of the population are measured). Accordingly, the variance of $y_k - y_{k-1}$ can be estimated as:

$$\mathrm{Var}(y_k) + \mathrm{Var}(y_{k-1})$$

To estimate the total variance of the estimate of the mean (or total) of the population of 2,367 photo plots, a component of variance must be added; this component can be estimated from the measurements in Year 0 (Section 3.2).

4.2 SIMPLE RANDOM SAMPLING – NO AUXILIARY INFORMATION

Over 10 years, sampling the population without replacement divides the population of N =314 plots into j = 10 possible samples with size n_j, where $N = \sum_{j=1}^{10} n_j$. There are $\frac{N!}{(n_1!n_2!...n_{10}!)}$ possible outcomes with equal probability. Sample *j* will be measured in year *k*, and every ten years thereafter. Thus, $n_j = n_k$. Let y_{ijk} be the i^{th} observation of the j^{th} sample for year *k*. The population total in year *k* is:

$$Y_k = \sum_{j=1}^{10} \sum_{i=1}^{n_j} y_{ijk}$$

and the population mean in year *k* is:

$$\overline{Y}_k = \frac{Y_k}{N}$$

Let the sample total in year *k* be $y_k = \sum_{i=1}^{n_j} y_{ijk}$ so that the sample mean in year *k* is[11]:

[1] $$\overline{y}_k = \frac{y_k}{n_j}$$

Then $E(\overline{y}_k) = \overline{Y}_k$ where expectations are taken over all possible outcomes (i.e. $\overline{y}_k$ is an unbiased estimator of $\overline{Y}_k$), and also:

$$\hat{Y}_k = N\overline{y}_k$$

is an unbiased estimator of Y_k.

Further, an unbiased estimator of the variance of $\overline{y}_k$ is:

[2] $$\text{Var}(\overline{y}_k) = \frac{\frac{1}{n_j}\left(1 - \frac{n_j}{N}\right)\sum_{i=1}^{n_j}(y_{ijk} - \overline{y}_k)^2}{n_j - 1} = \frac{1}{n_j}\left(1 - \frac{n_j}{N}\right)s_k^2$$

and an unbiased estimator of the variance of $\hat{Y}_k$ is:

$$\text{Var}(\hat{Y}_k) = \frac{N^2}{n_j}\left(1 - \frac{n_j}{N}\right)s_k^2$$

Note that the n_j will not be equal but each will be approximately $N/10 \cong 31$. Accordingly,

$$\text{Var}(\hat{Y}_k) \approx 9Ns_k^2$$

4.3 Stratified Random Sampling – No Auxiliary Information

To permit the calculation of a standard error, consider the case where there are L strata, and the *h*th stratum has m_h=2 observations. It is assumed that N is even, as are the strata sizes, N_h, with:

$$\sum_{h=1}^{L} N_h = N .$$

[11] Numbered equations will be referred to in the numerical example.

The N=314 measurements of the base year are ordered and the first 20 form the first stratum, etc. With N=314, there will be 15 strata with N_h=20 and one with N_h=14. The number of strata is therefore L=16. In each year, m_h=2 units are selected at random without replacement from each stratum so that, after 10 years, each unit will have been measured at least once. In the case of N_h=14, all units will have been measured after 7 years. The easiest option is to continue sampling in this stratum with a new randomization. This stratum may best be placed in the middle of the ordered set. Under these conditions, the following estimations should be effectively unbiased.

The sample mean in year k is calculated as:

[3] $$\bar{y}_k = \frac{\sum_{h=1}^{L} N_h \bar{y}_{hk}}{N}$$

where:

$$\bar{y}_{hk} = \frac{\sum_{i=1}^{m_h} y_{hki}}{m_h}$$ with $m_h = 2$ and L=16.

The sample total in year k is calculated as:

$$\hat{Y}_k = N\bar{y}_k$$

The variance of the sample mean is:

[4] $$Var(\bar{y}_k) \hat{=} \frac{1}{N^2} \sum_{h=1}^{L} \frac{N_h (N_h - m_h) s_{hk}^2}{m_h}$$

where:

$$s_{hk}^2 = \frac{\sum_{i=1}^{m_h} (y_{hki} - \bar{y}_{hk})^2}{m_h - 1}$$

Since $m_h = 2$ for all h:

$$\text{Var}(\bar{y}_k) \hat{=} \frac{\frac{1}{N^2} \sum_{h=1}^{L} N_h (N_h - 2) s_{hk}^2}{2}$$

The variance of the sample total is then:

$$\mathrm{Var}\left(\hat{Y}_k\right) \hat{=} \frac{\sum_{h=1}^{L} N_h \left(N_h - 2\right) s_{hk}^2}{2}$$

Then, if $s_{hk}^2 \approx s_k^2$ for all *h*:

$$\mathrm{Var}\left(\hat{Y}_k\right) \hat{=} \left(\frac{15(20)(18)}{2} + \frac{14(12)}{2}\right) s_k^2 \;\hat{=}\; 2784 s_k^2 \;\hat{=}\; 9N s_k^2$$

Thus the estimates via stratified random sampling will be more precise than those via simple random sampling if $s_k^2 < s_{hk}^2$. Given the method of stratification, we should have $s_k^2 << s_{hk}^2$.

4.4 Ratio Estimation with Simple Random Sampling

Using the notation from Section 4.2, define:

[5] $$\hat{R}_k = \frac{y_k}{y_0}$$

Then, let $y_{j0} = \sum_{i=1}^{n_j} y_{ij0}$, where the summation is over n_j units measured at time 0. Note that all groups are measured in the base year.

The overall total at the base year is then: $Y_0 = \sum_{j=1}^{10} y_{j0}$. An estimate of the population total at year *k* is then:

$$\hat{Y}_k = Y_0 \hat{R}_k$$

And an estimate of the population mean in year *k* is:

[6] $$\bar{y}_k = \bar{Y}_0 \hat{R}_k$$

These estimators are slightly biased.

As an approximation to the variance of $\hat{Y}_k$, we take:

$$\begin{aligned}\text{Var}\left(\hat{Y}_k\right) &\hat{=} \frac{N\left(N-n_j\right)}{n_j\left(n_j-1\right)}\sum_{i=1}^{n_j}\left(y_{ijk}-\hat{R}_k y_{ij0}\right)^2 \\ &\hat{=} \frac{N\left(N-n_j\right)}{n_j\left(n_j-1\right)}\left[\sum_{i=1}^{n_j}y_{ijk}^2+\hat{R}_k^2\sum_{i=1}^{n_j}y_{ij0}^2-2\hat{R}_k\sum_{i=1}^{n_j}y_{ijk}y_{ij0}\right]\end{aligned}$$

An approximation to the variance of $\bar{y}_k$ will then be:

[7]

$$\text{Var}\left(\bar{y}_k\right)=\frac{\text{Var}\left(\hat{Y}_k\right)}{N^2}$$

The ratio estimate will be more precise than the simple random sample estimator if the correlation between the y_{ijk} and y_{ij0} is greater than:

$$\frac{\text{Coefficient of Variation}(y_{ij0})}{2\cdot\text{Coefficient of Variation}(y_{ijk})}$$

These coefficients of variation should be roughly equal, so the ratio estimator should be the more precise if the correlation is greater than approximately ½.

4.5 Ratio Estimation with Stratified Random Sampling

The two potential options are separate ratio estimates for each stratum, or a combined estimator for all strata.

4.5.1 Separate estimator

Using the notation of Section 4.3 let:

$$y_{h0}=\sum_{i=1}^{2}y_{ih0} \quad ,$$

$$y_{hk}=\sum_{i=1}^{2}y_{ihk} \quad \text{, and}$$

$$Y_{h0}=\sum_{i=1}^{N_h}y_{ih0}$$

Define the separate ratio as:

$$\hat{R}_{S_{hk}}=y_{hk}/y_{h0}$$

Then, the estimated total for stratum h is:

$$\hat{Y}_{hk} = Y_{h0} y_{hk} / y_{h0} = Y_{h0} \hat{R}_{S_{hk}}$$

with variance approximated by:

$$\text{Var}\left(\hat{Y}_k\right) = \frac{\left(N_h^2 - m_h\right)}{N_h\left(m_h - 1\right)}\left[\sum_{i=1}^{N_h} y_{ihk}{}^2 + \hat{R}_{C_k}^2 \sum_{i=1}^{N_h} y_{ih0}^2 - 2\hat{R}_{C_k} \sum_{i=1}^{N_h} y_{ihk} y_{ih0}\right]$$

Then the estimate of the total is:

$$\hat{Y}_k = \sum_{h=1}^{L} \hat{Y}_{hk}$$

with variance estimated by:

$$\text{Var}(\hat{Y}_k) \hat{=} \sum_{h=1}^{L} s_{hk}^2$$

The estimate of the mean is then:

$$\bar{y}_k = R_{S_k} \overline{Y}_0$$

with variance estimated by:

$$\text{Var}(\bar{y}_k) = \frac{Var\left(\hat{Y}_k\right)}{N^2}$$

4.5.2 Combined estimator

Using the notation of Section 4.3, let the combined ratio $\hat{R}_C$ in year *k* be:

[8]

$$\hat{R}_{C_k} = \frac{\sum_{h=1}^{L} N_h \bar{y}_{hk}}{\sum_{h=1}^{L} N_h \bar{y}_{h0}}$$

Then the estimate of the total is:

$$\hat{Y}_k = \hat{R}_{C_k} Y_0$$

with variance estimated by:

$$\text{Var}\left(\hat{Y}_k\right) = \frac{\left(N_h^2 - m_h\right)}{N_h\left(m_h - 1\right)}\left[\sum_{i=1}^{N_h} y_{ihk}{}^2 + \hat{R}_{C_k}^2 \sum_{i=1}^{N_h} y_{ih0}^2 - 2\hat{R}_{C_k} \sum_{i=1}^{N_h} y_{ihk} y_{ih0}\right]$$

The estimate of the mean is then:

[9] $$\bar{y}_k = R_{C_k} \bar{Y}_0$$

with variance estimated by:

[10] $$\mathrm{Var}(\bar{y}_k) = \frac{Var(\hat{Y}_k)}{N^2}$$

4.6 DISCUSSION

Unless R_h is constant from stratum to stratum, the separate estimator is likely to be more precise. However, with $n_{hk}=2$ the approximation to the variance of the separate estimator is questionable. In this situation, the combined estimator would be generally recommended unless there is strong empirical evidence to the contrary. Another reason for preferring the combined ratio approach is that the predicted growth from a growth model (used as the auxiliary variable to estimate change) for some strata could be zero, making the separate ratio formulae unusable. The way that the strata have been constructed suggests that appreciable variation in the R_h is a real possibility. The choice of estimator should therefore be left until the data have been collected and the variability of the R_h explored.

For the ratio estimation one would have ideally $E(y_{hki}) = R_h\ E(y_{hoi})$. Regression estimation would be preferred if some other linear relationship existed, however, the advantage of using regression estimation over ratio estimation, even in non-linear cases, may be minor. One would depart from ratio estimation only if the data showed a well-defined relationship other than proportionality. Regression estimation formulae parallel those of ratio estimation and can be presented if required. Use of a known (or assumed) growth model can be treated similarly to ratio estimation with known parameters (i.e. not estimated from the data).

Note that with stratification, unless all the s^2_h can be assumed equal (and thus estimated jointly) for the construction of confidence intervals, the "effective degrees of freedom" will likely need to be calculated.[12]

[12] See Cochran, W.G. 1962. Sampling Techniques, 2nd Ed., p.94.

5. APPENDIX II – EXAMPLE OF CALCULATIONS AND SIMULATION RESULTS

5.1 SIMULATED DATA

A simulated data set of 314 plots was developed to demonstrate the different options for calculating estimates and to assess the relative performance of the different estimators. For ease of simulation, the variable of interest was assumed to be volume/ha. The following steps were taken to generate the 314 plots:

1. Each plot was randomly assigned a leading species based on the estimated proportion of the provincial area covered by that leading species.
2. Each plot was randomly assigned as mature or immature based on the assumption of 60% mature and 40% immature.
3. Each plot was randomly assigned an age. For mature plots, a uniform distribution between either 80 or 120 and 340 was assumed. For immature plots, a skewed distribution towards the younger ages was assumed.
4. VDYP yield curves were generated for each leading species assuming an average site index.
5. Each plot was randomly assigned an initial volume (Year 0) by selecting the appropriate yield curve and assuming a CV of 30% around that curve.
6. Volumes for Years 1 through 10 were generated by assuming a CV of 15% around the yield curve.[13] Growth was estimated by taking the difference between this new volume and the VDYP volume at time zero. Growth was then added to the generated volumes for the previous year. For mature plots, there was also a 0.5% chance that their volume would go to zero, to represent harvesting.

Simulated data for 314 plots are shown in Table 1. The plots are shown sorted by Year 0 volume to illustrate how stratification could be done. For this example, 15 strata of 20 plots and one stratum of 14 plots were assumed. The stratum with 14 plots was placed in the middle of the sorted list (stratum 8). Plots randomly selected under simple random sampling (SRS) and stratified random sampling (STRS) are flagged.

[13] This CV could be interpreted to account for the effects of natural variability and measurement error.

Table 1. Year 0 and 1 measurements for the 314 simulated plots.

Stratum – Plot #	Plot #	Year 0 – Age	Year 0 – Volume	Year 1 – Volume	SRS	STRS	Stratum – Plot #	Plot #	Year 0 – Age	Year 0 – Volume	Year 1 – Volume	SRS	STRS
1-1	100	0	0.0	0.0			3-1	274	6	0.0	0.0		
1-2	156	8	0.0	0.0			3-2	304	0	0.0	0.0		
1-3	213	0	0.0	0.0			3-3	2	28	0.0	0.0		
1-4	232	11	0.0	0.0	x		3-4	11	9	0.0	0.0		
1-5	267	20	0.0	0.0			3-5	20	11	0.0	0.0		
1-6	189	7	0.0	0.0	x		3-6	40	15	0.0	0.0		
1-7	204	17	0.0	0.0			3-7	45	23	0.0	0.0		x
1-8	261	19	0.0	0.0			3-8	66	30	0.0	0.0		
1-9	305	12	0.0	0.0		x	3-9	70	12	0.0	0.0		
1-10	16	13	0.0	0.0			3-10	107	7	0.0	0.0		
1-11	75	3	0.0	0.0			3-11	128	9	0.0	0.0		
1-12	288	3	0.0	0.0			3-12	164	20	0.0	0.0		
1-13	9	0	0.0	0.0			3-13	188	0	0.0	0.0		
1-14	118	12	0.0	0.0		x	3-14	223	22	0.0	0.0		
1-15	170	16	0.0	0.0	x		3-15	311	28	0.0	0.0		x
1-16	218	14	0.0	0.0			3-16	312	8	0.0	0.0		
1-17	195	16	0.0	0.0			3-17	94	4	0.0	0.0		
1-18	273	7	0.0	0.0			3-18	97	30	0.0	0.0		
1-19	13	2	0.0	0.0	x		3-19	235	11	0.0	0.0		
1-20	116	11	0.0	0.0			3-20	269	14	0.0	0.0		
2-1	123	6	0.0	0.0			4-1	292	11	0.0	0.0		
2-2	194	24	0.0	0.0			4-2	296	31	0.0	0.0		x
2-3	240	1	0.0	0.0			4-3	12	28	0.0	0.0		
2-4	254	5	0.0	0.0			4-4	57	15	0.0	0.0		x
2-5	6	9	0.0	0.0			4-5	67	6	0.0	0.0		
2-6	22	1	0.0	0.0			4-6	101	8	0.0	0.0		
2-7	51	26	0.0	1.5		x	4-7	133	12	0.0	0.0		
2-8	58	18	0.0	0.0			4-8	143	0	0.0	0.0		
2-9	79	5	0.0	0.0		x	4-9	198	8	0.0	0.0	x	
2-10	83	10	0.0	0.0			4-10	96	34	0.1	8.5		
2-11	88	18	0.0	0.0			4-11	64	35	7.1	13.3		
2-12	98	1	0.0	0.0			4-12	209	29	10.6	16.8		
2-13	117	6	0.0	0.0			4-13	228	29	11.6	16.8		
2-14	187	6	0.0	0.0	x		4-14	46	40	13.5	19.1		
2-15	205	16	0.0	0.0			4-15	275	35	18.4	25.5		
2-16	207	2	0.0	0.0			4-16	7	23	19.9	23.3		
2-17	210	1	0.0	0.0			4-17	263	30	23.0	28.5		
2-18	215	12	0.0	0.0			4-18	258	23	25.5	35.6		
2-19	217	8	0.0	0.0			4-19	121	31	27.9	32.5		
2-20	270	21	0.0	0.0			4-20	19	30	32.9	37.1		

Table 1 (cont.) Year 0 and 1 measurements for 314 simulated plots.

Stratum – Plot #	Plot #	Year 0 – Age	Year 0 – Volume	Year 1 – Volume	SRS	STRS	Stratum – Plot #	Plot #	Year 0 – Age	Year 0 – Volume	Year 1 – Volume	SRS	STRS
5-1	54	32	32.9	39.1			7-1	280	88	250.2	254.0		x
5-2	246	23	37.2	53.0		x	7-2	141	175	253.3	254.2		
5-3	60	33	41.3	46.1			7-3	279	32	253.6	264.2		
5-4	39	40	43.9	48.2			7-4	76	114	255.4	258.8		
5-5	140	39	48.8	53.1		x	7-5	184	226	259.6	260.6		
5-6	56	35	49.7	54.6			7-6	222	168	262.0	262.0	x	
5-7	37	39	54.2	63.1			7-7	192	37	263.2	278.1		
5-8	268	34	56.4	73.1	x		7-8	129	98	267.0	268.9		
5-9	41	27	61.7	75.0			7-9	71	90	273.2	276.1		
5-10	29	24	65.1	79.9			7-10	251	40	276.8	291.3		
5-11	259	38	69.0	74.5			7-11	25	253	281.6	281.9		x
5-12	289	81	79.0	82.7			7-12	183	172	284.1	284.3		
5-13	212	30	80.6	97.8			7-13	127	107	287.5	290.9		
5-14	91	52	88.4	92.4			7-14	52	86	288.5	292.2		
5-15	168	50	93.1	97.9			7-15	5	255	299.6	299.9		
5-16	157	26	93.4	110.2			7-16	163	142	301.5	302.1		
5-17	243	26	95.7	110.8			7-17	31	276	302.5	302.5		
5-18	86	42	120.3	127.1			7-18	15	315	303.1	303.1		
5-19	221	31	130.0	144.5			7-19	150	288	304.3	304.3		
5-20	300	109	136.0	137.6			7-20	78	91	305.5	308.0		
6-1	219	27	142.8	160.6			8-1	272	104	310.6	312.9		
6-2	148	61	152.9	156.5			8-2	200	177	312.7	312.7		
6-3	283	36	153.8	164.7			8-3	33	243	314.7	315.1		
6-4	145	183	154.8	154.9		x	8-4	211	256	318.3	318.5		x
6-5	38	39	158.6	172.4	x		8-5	50	224	323.0	323.5		
6-6	4	160	168.5	169.2			8-6	43	118	326.8	329.0		
6-7	242	159	169.3	169.3			8-7	114	239	330.1	330.1		
6-8	202	31	172.1	185.8			8-8	233	85	338.9	342.5		
6-9	167	169	190.8	191.8			8-9	146	174	340.1	340.2		
6-10	282	28	192.4	205.0	x		8-10	81	142	344.6	345.6		
6-11	302	35	196.4	208.1			8-11	155	113	347.3	349.6		x
6-12	93	49	198.6	203.2			8-12	28	203	351.0	352.6		
6-13	175	322	211.0	211.0			8-13	110	192	355.8	356.7		
6-14	276	77	223.0	226.4			8-14	17	253	362.6	363.0		
6-15	27	112	229.2	231.3									
6-16	265	220	232.2	232.5	x								
6-17	248	137	234.4	236.2									
6-18	142	231	237.0	237.0		x							
6-19	135	121	241.9	244.6									
6-20	185	100	249.7	252.4									

Table 1 (cont.) Year 0 and 1 measurements for 314 simulated plots.

Stratum – Plot #	Plot #	Year 0 – Age	Year 0 – Volume	Year 1 – Volume	SRS	STRS	Stratum – Plot #	Plot #	Year 0 – Age	Year 0 – Volume	Year 1 – Volume	SRS	STRS
9-1	191	298	368.2	368.4			11-1	55	268	449.5	449.5		
9-2	297	194	368.4	368.4			11-2	73	216	451.2	452.0		
9-3	147	102	370.0	373.1		x	11-3	308	310	453.5	453.6	x	
9-4	111	110	371.1	373.4			11-4	112	236	454.0	454.2		
9-5	165	232	372.3	372.5			11-5	239	36	456.1	469.6		
9-6	92	118	376.8	380.1	x		11-6	245	131	457.4	458.8		
9-7	197	267	377.9	377.9			11-7	144	291	457.7	457.9		
9-8	247	239	378.6	378.6			11-8	234	261	464.1	464.1		
9-9	227	215	380.4	381.0			11-9	14	139	468.7	471.9		x
9-10	293	296	380.6	380.9			11-10	301	125	470.1	470.5		
9-11	99	129	383.1	385.1			11-11	162	124	471.1	477.6		
9-12	61	143	389.7	391.0		x	11-12	32	277	472.6	472.8		
9-13	253	292	390.4	390.6			11-13	307	298	478.5	478.7		
9-14	120	307	392.6	392.6			11-14	87	183	481.5	481.5		
9-15	176	101	394.0	396.9			11-15	152	151	487.5	0.0		
9-16	35	309	394.8	394.8			11-16	125	129	488.8	492.0		
9-17	65	133	394.9	396.1			11-17	10	216	490.3	490.9		
9-18	80	263	397.2	397.5	x		11-18	124	319	495.3	495.4		
9-19	306	169	397.3	397.6			11-19	231	185	499.6	499.6	x	x
9-20	84	333	397.3	397.4			11-20	229	131	500.9	502.8	x	
10-1	34	269	403.1	403.1			12-1	26	259	501.4	501.4		
10-2	199	300	405.2	405.3			12-2	287	112	506.4	508.7		
10-3	249	183	410.8	410.8			12-3	256	160	508.2	509.0		x
10-4	122	254	414.0	414.3			12-4	89	217	511.8	512.1		
10-5	154	277	416.1	416.7		x	12-5	271	167	523.8	524.2		
10-6	214	128	420.0	422.2			12-6	85	147	523.9	524.8	x	
10-7	291	146	423.1	424.5			12-7	303	251	524.5	524.8		
10-8	201	127	427.7	430.2			12-8	277	154	525.1	525.1	x	
10-9	42	159	429.6	429.6			12-9	180	313	536.3	536.3		
10-10	53	241	429.7	430.0		x	12-10	108	153	538.9	540.1		
10-11	299	268	430.0	430.0			12-11	244	293	540.6	540.6		
10-12	136	174	431.1	0.0			12-12	174	73	544.9	550.6	x	
10-13	314	169	432.2	433.1			12-13	23	190	551.9	552.6	x	x
10-14	262	308	432.9	433.1			12-14	82	158	554.1	555.4		
10-15	220	332	433.0	433.1	x		12-15	62	156	563.4	564.6		
10-16	206	175	433.6	433.6	x		12-16	173	210	566.7	566.9		
10-17	74	194	440.4	440.4			12-17	252	175	570.3	570.3		
10-18	266	141	441.3	442.1			12-18	309	159	577.8	578.8		
10-19	48	179	447.0	449.0	x		12-19	193	218	578.7	580.3		
10-20	295	335	447.7	447.8			12-20	30	276	579.3	579.5		

Table 1 (cont.) Year 0 and 1 measurements for 314 simulated plots.

Stratum – Plot #	Plot #	Year 0 – Age	Year 0 – Volume	Year 1 – Volume	SRS	STRS	Stratum – Plot #	Plot #	Year 0 – Age	Year 0 – Volume	Year 1 – Volume	SRS	STRS
13-1	104	233	579.7	580.1			15-1	69	275	826.0	826.0		x
13-2	203	192	580.7	580.7		x	15-2	106	148	845.7	847.4		
13-3	63	333	586.6	586.8			15-3	149	130	861.8	867.0	x	
13-4	18	300	590.3	590.5			15-4	24	181	862.0	863.2	x	
13-5	237	184	598.7	599.4			15-5	68	190	879.4	881.7		
13-6	138	161	602.0	602.7			15-6	250	328	893.3	894.3		
13-7	8	206	610.6	611.2			15-7	230	275	927.4	927.4		
13-8	181	337	616.2	616.2			15-8	115	311	938.6	938.9		x
13-9	264	308	616.7	616.8			15-9	178	284	944.8	944.8		
13-10	313	275	618.4	618.6			15-10	134	232	979.3	980.5		
13-11	290	287	619.4	620.0			15-11	238	80	994.1	1003.2	x	
13-12	159	162	621.7	624.3			15-12	179	130	1004.0	1008.3		
13-13	190	190	636.9	636.9			15-13	49	123	1015.8	1020.4		
13-14	153	324	643.9	643.9			15-14	105	253	1051.7	1053.3		
13-15	77	56	646.5	656.5			15-15	130	339	1075.9	1075.9		
13-16	225	310	654.1	654.3			15-16	119	161	1089.5	1094.3		
13-17	160	194	654.7	655.3			15-17	161	338	1110.9	1111.0		
13-18	196	303	658.6	658.8			15-18	113	331	1112.2	1112.2		
13-19	284	238	674.1	674.4			15-19	139	239	1123.4	1124.7		
13-20	294	320	674.3	674.3	x	x	15-20	95	200	1152.2	1153.6		
14-1	36	253	689.1	689.3			16-1	151	279	1163.1	1164.5		
14-2	226	83	693.4	700.3			16-2	208	320	1198.2	1198.3		
14-3	286	237	702.0	702.2			16-3	59	315	1209.0	1209.0	x	
14-4	103	338	704.5	704.5		x	16-4	171	226	1248.8	1250.3		
14-5	3	267	716.8	716.8		x	16-5	21	196	1278.2	1279.4		
14-6	44	312	733.9	733.9			16-6	172	143	1304.3	1307.2	x	
14-7	281	171	734.4	736.6			16-7	126	244	1337.1	1338.0		
14-8	241	304	734.5	734.5			16-8	47	314	1351.1	1352.0		
14-9	278	161	737.7	740.5	x		16-9	260	202	1465.0	1467.0		
14-10	90	147	753.4	756.4	x		16-10	158	304	1469.5	1470.5		
14-11	182	107	755.8	760.6			16-11	72	272	1487.2	1487.7		
14-12	257	328	758.7	758.8			16-12	186	260	1501.0	1502.3		
14-13	285	141	771.8	775.1			16-13	169	292	1542.3	1542.7		
14-14	298	306	788.4	788.5			16-14	132	305	1543.2	1543.5		
14-15	131	274	791.2	791.8			16-15	109	265	1576.5	1576.5		
14-16	216	206	791.4	791.8			16-16	102	276	1634.6	1636.2		x
14-17	255	146	793.5	795.9			16-17	1	331	1646.8	1647.7		x
14-18	310	161	800.2	801.5			16-18	166	316	1744.6	1744.6		
14-19	177	294	824.5	824.5			16-19	224	172	1908.0	1911.6		
14-20	236	126	825.8	830.1			16-20	137	311	2108.7	2109.6		

5.2 STATISTICAL ESTIMATION

Only the current yield estimates are demonstrated and extending the estimation to change is straightforward. The calculations below are examples for Year 1 and are based on one set of randomly generated measurements for the 314 plots. The population mean in Year 0 is 396.3 m^3/ha and in Year 1 is 397.9 m^3/ha. Year 1 estimates are based on different sample selection methods presented in Table 2. They are based on the formula described in sections 4.2 to 4.5.

Table 2. Estimates of the average volume in Year 1 based on the data presented in Table 1.

Method	Equation #	Parameter	Estimate
SRS	1	mean	453.80
	2	Var(mean)	3694.11
		SE(mean)	60.78
Ratio with SRS	5	ratio	1.01
	6	mean	398.60
	7	Var(mean)	0.73
		SE(mean)	0.85
STRS	3	Mean	409.90
	4	Var(mean)	30.07
		SE(mean)	5.48
Ratio with STRS	8	Ratio	1.00
	9	Mean	397.40
	10	Var(mean)	0.15
		SE(mean)	0.39

5.3 RELATIVE PERFORMANCE OF ESTIMATORS

The various estimators are compared in terms of precision (standard error); the errors of estimation (due to the simulation sampling) are also provided. The results of 20 independent simulations (i.e. 20 generations of 314 plots, their growth and the sample estimates based upon them) are presented below. For each estimate of the mean, the error of estimation was calculated as the difference between the sample mean and the population mean. Minimum, maximum and average percent errors of estimation for each method for Years 1 through 10 are summarized in Table 3 and Table 4. The minimum, maximum and average standard errors for each method for Years 1 through 10 are summarized in Table 5 and Table 6.

Table 3. Percent error of estimation for simple random sampling and simple random sampling with a ratio estimator. Results of 20 simulations over 10 years of sampling.

Year	SRS			SRS Ratio		
	Min (%)	Max (%)	Ave (%)	Min (%)	Max (%)	Ave (%)
1	-22.4	41.2	2.4	-0.2	2.7	0.5
2	-23.2	26.6	-1.3	-3.3	2.6	0.0
3	-25.6	37.5	8.4	-10.8	1.4	-0.5
4	-44.4	32.9	-6.2	-7.7	2.7	-0.4
5	-19.8	11.9	-2.8	-0.2	3.1	0.9
6	-21.2	24.1	0.1	-3.5	2.7	-0.3
7	-43.5	30.8	3.3	-10.6	5.5	-0.6
8	-24.5	28.3	-3.9	-3.5	3.3	0.7
9	-22.8	22.6	0.1	-5.3	12.7	2.0
10	-35.4	38.6	0.1	-6.5	16.1	0.8
Average..			0.0			0.3

Table 4. Percent error of estimation for stratified random sampling and stratified random sampling with a ratio estimator. Results of 20 simulations over 10 years of sampling.

	STRS			STRS ratio		
Year	Min (%)	Max (%)	Ave (%)	Min (%)	Max (%)	Ave (%)
1	-10.0	14.1	1.6	-3.6	0.4	-0.3
2	-27.1	15.5	2.7	-24.4	0.4	-1.2
3	-15.1	19.6	0.0	-3.2	1.0	0.0
4	-13.6	10.8	-0.2	-5.9	1.6	-0.5
5	-25.8	14.2	1.8	-1.9	1.8	0.2
6	-7.6	32.2	2.5	-9.4	2.2	-4.6
7	-1.3	38.1	9.2	-2.2	2.7	0.0
8	-23.3	17.3	5.0	-5.6	4.2	0.1
9	-24.6	26.4	4.9	-4.0	4.3	0.0
10	-5.1	37.8	5.7	-1.8	1.9	0.3
Average			3.3*			-0.6

* This error is high because in the simulation there was no randomization of samples within the strata.

Table 5. Standard errors for simple random sampling and simple random sampling with a ratio estimator. Results of 20 simulations over 10 years of sampling.

	SRS			SRS ratio		
Year	Min	Max	Ave	Min	Max	Ave
1	50.8	96.4	72.4	0.3	1.1	0.7
2	52.8	86.0	70.2	0.8	14.4	2.5
3	58.3	94.6	75.8	1.1	46.3	5.2
4	42.0	100.1	67.8	1.1	29.1	4.7
5	52.3	88.1	72.5	2.4	6.5	3.9
6	51.5	103.4	73.1	1.4	14.1	4.5
7	47.3	102.4	75.5	2.9	39.1	8.6
8	39.7	97.7	71.6	1.6	15.6	6.0
9	52.8	94.9	71.3	4.7	19.3	9.5
10	49.2	84.5	67.0	4.0	15.0	8.0

Table 6. Standard errors for stratified random sampling and stratified random sampling with a ratio estimator. Results of 20 simulations over 10 years of sampling.

	STRS			STRS ratio		
Year	Min	Max	Ave	Min	Max	Ave
1	4.2	68.2	15.6	0.2	14.6	1.8
2	3.5	80.9	22.1	0.3	6.6	1.4
3	2.4	77.3	13.4	0.6	12.0	2.1
4	2.8	72.1	16.3	0.6	26.2	4.5
5	5.1	72.8	14.4	0.8	10.7	2.7
6	5.8	78.2	17.7	1.4	16.0	3.8
7	16.2	93.5	32.0	1.7	14.9	3.9
8	8.7	74.7	20.1	1.9	29.9	5.5
9	6.2	93.0	18.3	2.2	14.4	5.3
10	7.8	79.8	19.6	2.3	9.0	5.0

Under the assumptions used for the simulations, the absolute values are not expected to match those that will be obtained if an actual sample of 314 ground plots is established across the province. However, the relative performance of the estimators illustrated in the simulation is expected to provide a reasonable indication of the relative performance of the estimators when applied to actual data. Given this, as would be expected, the ratio estimators perform better than the simple expansion estimators. The ratio estimator performed only slightly better under stratified random sampling than simple random sampling. This is probably due to the strong correlation between measurements at Year 0 and re-measurements in Years 1 through 10.

6. APPENDIX III – ESTIMATION WITH A KALMAN FILTER

The estimation methods outlined in Section 3 can be put in the framework of a Kalman filter. The generic form of the Kalman filter (in the notation of van Deusen (1989))[14] is as follows:

The relationship between the vector of observations at time **t**, $\mathbf{Y_t}$ and the state parameters a_t , is assumed to be:

$$\mathbf{Y_t} = \mathbf{F_t} a_t + \mathbf{v_t}$$

where:

$\mathbf{F_t}$ is a fixed matrix, $\mathbf{E}(\mathbf{v_t}) = 0, \mathbf{Var}(\mathbf{v_t}) = \mathbf{V_t}$.

This is referred to as the observation or measurement equation. The state parameters are variables that evolve over time, thus:

$$a_t = \mathbf{G_t} a_{t-1} + \mathbf{w_t}$$

where:

$\mathbf{G_t}$ is another fixed matrix, $\mathbf{E}(\mathbf{W_t}) = 0, \mathbf{Var}(\mathbf{w_t}) = \mathbf{W_t}$.

This is referred to as the transition of system equation.

This is essentially the same as the formulation assumed by Dixon and Howitt (1979)15 and Kangas (1991)16. These authors, however, include an extra component in the transition equation, namely:

$$a_t = \mathbf{G_t} a_{t-1} + \mathbf{K_t} \mathbf{u_t} + \mathbf{w_t}$$

where:

[14] Van Dusen, P.C. 1989. Multiple-occasion partial replacement sampling for growth components. For. Sci. 35:388-400.

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

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

the $\mathbf{K_t u_t}$ represent the "impact of control actions on the state variables". In our context this could be, for example, the effect of harvesting or fire that would upset the natural progression of growth.

The simplest situation would be:

$$Y_t = m_t + e_t \quad \text{, and}$$

$$m_t = m_{t-1} + h_t$$

In other words, the state parameters follow a random walk.

More realistically, m_t could be a function of age, species, etc. Then, $m_t = r_t m_{t-1} + h_t$ with r_t also a function of age, species, etc.

Note that the Kalman filter is essentially an empirical Bayes procedure where all prior information is used to project parameter values, which are then modified on the basis of current data. Estimation is as follows:

Let $\mathbf{a_t}$ denote the optimal estimator of a_t based on all information up to and including $\mathbf{Y_t}$, and let $\mathbf{Var}(\mathbf{a_t} - a_t) = \mathbf{P_t}$. The prediction equation for a_t, and the associated variance matrix, conditional on a_{t-1} and $\mathbf{P_{t-1}}$ are:

$$\mathbf{a_{t|t-1}} = \mathbf{G_t a_{t-1}}$$

$$\mathbf{P_{t|t-1}} = \mathbf{G_t P_{t-1} G'_t} = \mathbf{W_t}$$

When $\mathbf{Y_t}$ becomes available, the updating equations for the estimation of a_t and the associated variance matrix, are:

$$\mathbf{a_t} = \mathbf{a_{t|t-1}} + \mathbf{P_{t|t-1} F'_t H_t^{-1} E}_t$$

$$\mathbf{P_t} = \mathbf{P_{t|t-1}} - \mathbf{P_{t|t-1} F'_t H_t^{-1} F_t P_{t|t-1}}$$

where:

$$\mathbf{E_t} = \mathbf{Y_t} - \mathbf{F_t a_{t|t-1}} \text{, and}$$

$$\mathbf{H_t} = \mathbf{F_t P_{t|t-1} F'_t} + \mathbf{V_t}$$

Initial values a_0 and $\mathbf{P}_0$ are required – the setting of these can be discussed elsewhere.

Also required is knowledge of $\mathbf{V_t}$ and $\mathbf{W_t}$. Maximum likelihood can be used for estimating any unknown components. Under the assumption of normality, the log likelihood of the sample is given by:

$$\mathbf{L} = \frac{1}{2}\sum_{\mathbf{t}} \log(\left|\mathbf{H_t}\right| + \mathbf{E}_\mathbf{t}'\mathbf{H_t}^{-1}\mathbf{E_t})$$

Maximizing **L,** provided there are not too many unknown variances, is relatively straightforward.

Once the matrices $\mathbf{F_t}$ and $\mathbf{G_t}$ have been defined, the procedure can be readily programmed, especially if a matrix language, such as SAS IML is available.

7. APPENDIX IV – DEFINITIONS

We define the following terms for this report:

Change is *net change*, which equals survivor growth plus ingrowth less mortality.

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.

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

Yield audit is the process of observing the yield of a forest and comparing this with the predicted yield of that forest.