Canada's National Forest Inventory

Estimation Procedures

3 September 2004
Version 1.13

File: C:\Documents and Settings\aomule\My Documents\NFI\Estimation\NFI estimation methods_AllAttributes_ver1.13.doc

Table of Contents

1. Introduction

1.1 Overview

This document outlines procedures for the statistical estimation of area totals and totals of tree and other attributes based on Canada's National Forest Inventory (NFI) design. These totals are to be reported by national terrestrial ecozones (strata) and classifiers (domains), at a given point in time. Classifiers include, among others, land cover type, stand disturbance and treatment, land use, ownership and protection status. The estimation methods are described in sufficient detail so they can be readily implemented using common computer software.

Background information, including the NFI target population, stratification, sampling design and data sources, is summarized in sections 1.2 to 1.4, for easy reference. An overview of the estimation methods is given in section 2. Detailed methods to estimate attribute totals and averages by NFI unit using photo plots alone, ground plots alone and combined photo and ground plots are given in sections 3, 4 and 5, respectively. Computational examples are also provided in each section. A method to sum the NFI unit totals to the ecozone and national levels is provided in section 6. Finally, some issues related to the estimation, including assumptions, sample size, and alternative estimators are discussed in section 7.

1.2 Target population, stratification and sub-populations

The target population for the NFI is Canada's land mass, whether vegetated or not. This target population, and thus sampling units, is assumed to consist of infinitesimal points that are stratified by terrestrial ecozone for reporting purposes. Ecozones are partitioned into sub-populations called NFI units for estimation purposes. An NFI unit is defined as an ecozone within the boundaries of a province or a territory. It may consist of two or more subunits, which are non-overlapping subdivisions of the unit. The partitioning of an ecozone into units or subunits is necessitated by the jurisdictional differences in sampling schemes or different data collection methods that may otherwise occur within an ecozone. For example, photo plot data may come from satellite imagery for a portion of an ecozone or unit, and from aerial photos in another portion. See the NFI *Design Overview* document for further details.

1.3 Sampling design

The overall NFI sampling scheme is a probability sample of points in Canada (NFI *Design Overview* document). It is composed of two components (Table 1):

1. A single systematic sample of points, with a photo plot installed at each sample point.
2. A simple random sub-sample of the systematic sample within each selected ecozone, with a ground plot established at or adjacent each sub-sample point.

Table 1. Overview of the NFI sampling scheme.

Inventory Component	*Sampling design*	*Plot design*	*Attributes estimated*
Remote sensing	A single systematic sample of locations (sampling points) located at the nodes of a national 20 km x 20 km fixed grid over the entire land mass; the sample locations are partitioned into ecozones	A 2 km x 2 km photo plot centered at a sample location. The photo plot is partitioned into polygons or pixels using various classifiers based on aerial photo interpretation, other remotely sensed data such	Primary source of the NFI data. Attributes include area totals (summation of polygon or pixel areas) by various classifiers; and stand attributes (e.g., volume) by land cover

Inventory Component	*Sampling design*	*Plot design*	*Attributes estimated*
	and NFI units for implementation, analysis and reporting.	as the data from the Earth Observation for Sustainable Development (EOSD) project , and other data sources.	polygon within a plot, or EOSD summary table by plot.
Ground sampling	A simple random sub-sample of photo plot locations over an entire province or within an NFI unit. The nominal ratio of forested ground plots to photo plots is 1:10, with a minimum of 50 forested locations per ecozone, and no sampling in the three arctic ecozones.	A ground plot established at or adjacent the photo plot center. It actually consists of single set of plots including concentric circular plots, line transects and a soil pit.[1]	Attributes such as carbon mass; woody-debris and tree volume and biomass; plant lists; understory vegetation biomass; site classification, and soil descriptions.

The photo plots and the ground plots are all permanent and established following national guidelines as described in the NFI *Ground Plot Guidelines* and *Photo Plot Guidelines* documents. Ground plots are established in only the forested locations, which are areas classified as Vegetated Treed (see NFI *Land Cover Classification* document), have the potential to be classified as Vegetated Treed, or where tree cover is removed temporarily (e.g. harvested areas). Non-forested ground plot locations are tracked over time and plots established when they become forested. Unavailable (inaccessible) sites are replaced with a suitable subjectively-selected match, and difficult-access sites are sub-sampled, following the NFI *Guidelines for handling unavailable and difficult-access ground plots* document. Methods for photo plot and ground plot periodic remeasurement and handling over time are under development.

1.4 Data sources and issues

There are 56 NFI units (about 2-7 units per ecozone) in Canada formed from intersecting the 15 national ecozones with the 13 Canadian provinces/territories, with varying data collection methods in each unit. There are a total of about 25,646 photo-plot locations in the country, ranging from 345 to 4,225 locations per ecozone (Table 2). Photo plot data are obtained from aerial photos in about one-half of these locations (ecozones 4, 6-9 and 14) and partially from the EOSD data or other data sources in the remainder. The EOSD data may also be used as auxiliary information in the estimation of areas, where both aerial photos and EOSD data exist. Other auxiliary data sources such as CanFI (Canada's Forest Inventory) tree volume and biomass data may also be used as auxiliary variables in the estimation of tree attributes.

[1] The concentric plots include a 0.04-ha large tree plot (LTP) for measuring trees at least 9.0 cm dbh and a 0.005-ha small tree plot (STP) nested within the LTP for measuring trees less than 9.0 cm dbh. There are transect segments for measuring small woody debris (SWD), nominally 20 m, and for measuring coarse woody debris (CWD), nominally 40 m or 60 m (depending on the size of the CWD). For more details, refer to the NFI *Ground Plot Guidelines* document.

Table 2. Number of photo plots (20 km x 20 km grid points) by province/territory and ecozone.

Prov./	*Ecozone*															*Total*	
Terr.	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*(Number)*	*(%)*
BC				164					102			469	515	1,120		*2,370*	*9.2*
AB				147	22	20			940	402				125		*1,656*	*6.5*
SK					104	466			442	616						*1,628*	*6.3*
MB			4		313	637			313	183					169	*1,619*	*6.3*
ON						1,623		260							638	*2,704*	*10.5*
QC	28	87	385		1,307	1,629	169	85							91	*3,781*	*14.7*
NB							188									*188*	*0.7*
NS							142									*142*	*0.6*
PE							12									*12*	*0.0*
NL	45				570	411										*1,026*	*4.0*
YT			12	47							465	700	12			*1,236*	*4.8*
NT		597	461	1,314	891				44		218	14				*3,537*	*13.8*
NU	614	3,541	1,305	2	272										11	*5,747*	*22.4*
Total	*687*	*4,225*	*2,167*	*1,674*	*3,479*	*4,786*	*511*	*345*	*1,841*	*1,201*	*683*	*1,183*	*527*	*1,245*	*909*	*25,646*	

The nominal number of forested ground plots ranges from 0 to 264 per province, for a national total of about 1,135. On average, there are about 30 forested ground plots per NFI unit (excluding the NFI units in the arctic ecozones 1 – 3 (Table 2) and the Yukon territory). In some NFI units, existing permanent plots adjacent to the selected sample locations are used instead. The ground plot data are not used for area estimation, and the small number of ground plots may also prevent the use of double-sampling estimators for tree attributes. The ground plot data, if assumed to be the 'truth', can, however, be used to provide a level of confidence on the accuracy of the photo-plot based estimates. This could be achieved by constructing "error matrix" or "confusion matrix" [see, e.g., Congalton and Mead (1983)[2]] for qualitative attributes, and ratios between photo estimates and ground observations for quantitative attributes.

Due to cost and other considerations, there is some variation between the theoretical sampling scheme outlined in section 1.3 and the data collection methods chosen by the provinces/territories. This variation raises some issues that will affect estimation and reporting. These issues include:

1. There are various mixtures of data sources and coverage. For example, only satellite imagery is used in the northern areas (arctic ecozones) with little or no vegetation, and aerial photos and ground plots are used in the southern areas. This implies that the number of attributes to be estimated and reported will vary: fewer attributes can be reported in the north compared to the south.
2. The period of ground plot establishment varies among provinces/territories, with the overall period of plot establishment (Time 0) expected to last about 5 years. The currency of the aerial photo interpreted and EOSD data also varies among provinces/territories by up to 10 years. Thus, some assumptions have to be made about ground and photo plot changes during these periods (5 or 10 years) when compiling the Time 0 estimates.
3. Incomplete data records cover some NFI units. Some of these missing data may be replaced with actual data in the future, imputed data, or data that are judiciously extrapolated from neighboring units. No statistics will be produced for those attributes or areas where replacement of the missing data is not feasible.

[2] Congalton, R.G. and R.A. mead. 1983. A quantitative method to test for accuracy and correctness in photointerpretation. *Photogrammetric Enginering and Remote Sensing* 49:69-74.

2. Overview of the estimation

2.1 Methods

The main objective is to estimate area and attribute totals and their associated variances by classifier class at the ecozone and national levels. This will be achieved by estimating the totals for individual NFI units by classifier class, and then summing these individual-unit totals and their associated variances to the ecozone and national levels. For example, the totals from the Pacific Maritime/BC and the Pacific Maritime/Yukon units are summed to provide totals for the Pacific Maritime ecozone. An NFI unit with an insufficient number of observations will be pooled with an adjoining unit within the same ecozone.

The estimated area total for a classifier class is obtained by multiplying the estimated proportion of area in a given classifier class (based on photo plots) by the NFI unit total area. Similarly, the estimated tree or other attribute total is obtained by multiplying the population average or per-hectare estimate in a classifier class by the NFI total area. The population average or per-hectare value is calculated as the ratio of the sum of the (photo or ground) plot attribute totals in a classifier class to the sum of all the plot areas in the NFI unit. The attribute average over the classifier-class area is obtained by dividing the estimated population total by the estimated area total in a classifier class. In the estimation, ground plots in matched sites are treated just like a regular ground plot[3], and plots in difficult-access sites that were sub-sampled are weighted appropriately using $w_j = \frac{m_d}{m'_d}$, where w_j is the weight of the jth plot in the measured sub-sample, m_d is the number of difficult-access plots in the NFI unit, and m'_d is the number of the measured plots.

The area proportion and the attribute average or per-hectare value are both similar to a ratio-of-means estimator, which is a biased estimator, but whose bias is usually trivial with large sample size ($\geq$ 30) [Cochran (1977)][4]. This common approach to estimation results in consistency and ease of implementation of the estimation procedures. This ratio approach is used because of the likely differences in plot sizes or line-transect lengths within an NFI unit. Although the ground plots, line transects, or the photo plots are each nominally the same size, partial plots and line transects arise from only portions of ground plots and line transects being accessible or inside the population of interest, and from the partitioning of the photo plots into ecozones and NFI units.

The estimation outputs are attribute totals and their associated variances by classifier class and ecozone. A schematic illustration of the estimation output by classifier classes is given in Table 3.

Table 3. A schematic illustration of the area estimation ouptut.

NFI unit within Ecozone	*Classifier class*	
	Vegetated	*Non-Vegetated*
1	99.9 (.9) *	99.9 (.9)
2	99.9 (.9)	99.9 (.9)
3	99.9 (.9)	99.9 (.9)
4	99.9 (.9)	99.9 (.9)
5	99.9 (.9)	99.9 (.9)

[3] Data for inaccessible sites with no suitable matches are imputed.
[4] Cochran, W.G. 1977. Sampling techniques (3rd edition). John Wiley & Sons, Inc., New York.

NFI unit within Ecozone	Classifier class	
	Vegetated	*Non-Vegetated*
Ecozone Total	*9999.9 (9.9)*	*999.9(.9)*

* Area total and sampling error in brackets.

Estimation methods for attribute totals and averages by classifier class within an NFI unit using photo plots and ground plots are outlined in section 3 and 4, respectively. These estimation methods are applicable to any area or other attributes, including the 25 key NFI attributes (see the NFI *Design Overview* document). Double sampling estimators that use both photo plots and ground plots are described in section 5.

2.2 Assumptions

The following assumptions are implicitly required for the ratio estimation of area and other attribute totals and their associated variances and confidence limits by NFI unit:

1. There are at least 30 photo plots in the NFI unit for estimating photo-plot based attributes. About 10 NFI units have fewer than 30 photo plots; data from these units may have to be pooled with that from other units prior to analysis, or other methods of estimation such as model-assisted estimation may have to be used. When plot size is constant over all photo plots in a unit, the requirement for 30 photo plots can be waived. This is so because in this case the ratio-of-means estimator reduces to the usual SRS expansion estimator.
2. There are at least 30 ground plots in the NFI unit for estimating ground-based attributes. The number of NFI units with less than 30 ground plots will be higher than for photo plots. Thus, the ground observations may also have to be pooled prior to analysis. When plot size is constant over all ground plots, which will likely be the case in most units, the requirement for 30 ground plots can be waived.
3. The relationship between the photo plot or ground plot attribute total and the plot size is well approximated by a straight line through the origin. (Recall that plot sizes and transect lengths can vary within an NFI unit for the reasons given in section 2.1.)
4. The variance of the attribute plot total along this straight line is proportional to the plot size. This assumption, along with assumptions 1, 2 and 3, are particularly important for reliable estimates of the confidence limits of population totals [Cochran (1977)].
5. The photo plots from the systematic grid that fall in an NFI unit will be regarded as a simple random sample (SRS) from the NFI unit.
6. The photo plot and ground plot sampling fractions are expected to be small, so that the finite population correction factors can be ignored.
7. All photo plot data inside an NFI unit will be used in the estimation, including data from partial photo plots that straddle the NFI unit boundary (ecozone, provincial/territorial, national or salt-water boundaries). Use of all the sample data, including data from the partial plots (line transects), results in efficiency. It also, with proper weighting of the partial-plot data, eliminates potential "edge-effect" bias associated with approximate methods that exclude partial plots.
8. Ground plots do not straddle NFI unit boundaries since in most cases the unit boundaries are typically fuzzy and not mapped.
9. All the ground plots are accessible; however, if there were a sub-sample of difficult-access plots, the estimation methods can be easily modified by introducing weights as discussed in section 2.1.
10. The photo interpretation and satellite classification errors associated with photo plots, and the measurement errors associated with the ground plots, are negligible. For some attributes,

ground data (assumed to be the "truth") can be used to approximately verify the assumption about the photo plot errors, as discussed in section 1.4.
11. Temporal differences among the photo plot data (due to varying ages of photography or satellite images) and among the ground plot measurements (due to varying establishment dates) are ignored.
12. Total area for the NFI unit is determined from GIS without error.
13. Ground plots (or portions of ground plots) that fall in the non-forest areas are part of the sample and they are included in the estimation and assigned a value of zero (since they are not measured).

2.3 Notation

Consider an NFI unit that is the geographic area of interest, and let:

A = Total area (ha) of the NFI unit.

A_k = Total area (ha) in the *k*th classifier class within the NFI unit. For example, A_k could be the area of the Vegetated landbase in the NFI unit (Table 3).

$\hat{A}_k$ = Estimated total area in the *k*th classifier class within the NFI unit

p_k = Population proportion of area in the *k*th classifier class within the NFI unit.

$\hat{p}_k$ = Estimated proportion of area in the *k*th classifier class within the NFI unit.

a_{is} = *A*rea (ha) of the *s*th polygon in the *i*th photo plot within the NFI unit.

a_{ki} = *A*rea (ha) classified into the *k*th classifier class in the *i*th photo plot within the NFI unit.

$= \sum_{s=1}^{S} a_{is}$, where S is the total number of polygons in the *i*th photo plot, and a_{is} is set equal to zero if the *s*th polygon does not belong to the *k*th classifier class.

a_{Ti} = Total area of the *i*th photo plot (typically 400 ha, except for plots that straddle a unit, a national or a salt-water boundary).

n = Total number of complete and partial photo plots inside the NFI unit.

t''_{is} = Attribute stand average (e.g., site age) or per-hectare value (e.g., volume/ha) for the *s*th polygon in the *i*th photo plot in the NFI unit.

x_{ki} = Attribute total from the *i*th photo plot in the *k*th classifier class within the NFI unit.

$= \sum_{s=1}^{S} t''_{is} a_{is}$, where the t''_{is} has value zero if the *s*th polygon does not belong to the *k*th classifier class.

$\bar{a}_T$ = Average area of all the photo plots in the NFI unit.

$$= \frac{1}{n} \sum_{i=1}^{n} a_{Ti}$$

X_k = Photo based attribute population total for the *k*th classifier class within the NFI unit.

$\hat{X}_k$ = Photo-plot based estimated attribute total for the *k*th classifier class within the NFI unit.

$\bar{x}_k$ = Estimated attribute photo-plot based unit average or per-ha estimate for the *k*th classifier class within the NFI unit.

$\hat{\bar{x}}_k$ = Estimated attribute classifier-area average or per-ha estimate using photo plots for the *k*th classifier class within the NFI unit.

Y_k = Ground based attribute population total for the *k*th classifier class within the NFI unit.

$\hat{Y}_k^{\,c}$ = Ground-plot (combined LTP and STP) based estimated attribute total for the *k*th classifier class within the NFI unit.

$\hat{Y}_k$ = Ground-plot (LTP) based estimated attribute total for the *k*th classifier class within NFI unit.

$\hat{Y}_k^{\,r}$ = Ground-plot (LTP) and photo plot based estimated attribute total for the *k*th classifier class within NFI unit.

m = Number of all forested and non-forested sample ground locations in the NFI unit.

b_j = Area of the *j*th LTP (large tree plot) in the NFI unit. Variable LTP plot sizes arise from only portions of ground plots and line transects being accessible or inside the population of interest.

$\bar{b}$ = Average area of the LTPs in the NFI unit.

$$= \frac{1}{m}\sum_{j=1}^{m} b_j$$

b'_j = Area of the *j*th STP (small tree plot) in the NFI unit. Variable LTP plot sizes arise from only portions of ground plots and line transects being accessible or inside the population of interest.

$\bar{b}'$ = Average area of the STPs in the NFI unit.

$$= \frac{1}{m}\sum_{j=1}^{m} b'_j$$

t_j = Attribute stand average (e.g., site age) or per-hectare value (e.g., volume/ha) for the *j*th LTP in the NFI unit.

y_{kj} = Attribute total from the *j*th LTP for the *k*th classifier class within the NFI unit. (Note: the y_{kj} is set equal to zero if, according to photo plot information or other verification, the *j*th plot enter does not belong to the *k*th classifier class, or is non-forested.)

$= t_j\, b_j$.[5]

t'_j = Attribute stand average (e.g., site age) or per-hectare value (e.g., volume/ha) for the STP in the NFI unit.

y'_{kj} = Attribute total from the *j*th STP for the *k*th classifier class within the NFI unit (Note: the y'_{kj} is set equal to zero if the *j*th plot does not belong to the *k*th classifier class, or is non-forested).

$= t'_j b'_j$.

$\bar{y}_k$ = Estimated attribute unit average or per-ha estimate from the LTP for the *k*th classifier class within the NFI unit.

$\bar{y}'_k$ = Estimated attribute unit average or per-ha estimate from the STP for the *k*th classifier class within the NFI unit.

$\bar{y}_k^{\,c}$ = Estimated attribute unit average or per-ha estimate using combined ground plots (LTP and STP) for the *k*th classifier class within the NFI unit.

[5] For woody debris, the average is multiplied by T'_j or T_j, where T'_j or T_j are the transect lengths for coarse and small woody debris sampling, respectively, in the *j*th plot location.

$\bar{y}_k^{\ r}$ = Estimated attribute unit average or per-ha estimate using the LTP and the photo plots for the kth classifier class within the NFI unit.

$\hat{\bar{y}}_k$ = Estimated attribute classifier-area average or per-ha estimate using the LTP for the kth classifier class within the NFI unit.

$\hat{\bar{y}}_k^{\ c}$ = Estimated attribute classifier-area average or per-ha estimate using the LTP and STP for the kth classifier class within the NFI unit.

$\hat{\bar{y}}_k^{\ r}$ = Estimated attribute classifier-area average or per-ha estimate using the LTP and the photo plots for the kth classifier class within the NFI unit.

3. Estimates from photo plots

3.1 Area totals

3.1.1 Estimation inputs

Inputs for estimating area totals are obtained from the photo plot summary data tables in the NFI database or EOSD tables. These data include plot areas by land use types; plot areas by ownership; plot areas by protection status; and land-cover polygons (within plot) descriptions. The full list of photo plot classifiers and attributes is available in the NFI *National Standards for Photo Plots Data Dictionary* document.

Area estimation methods are outlined below for a classifier class within an NFI unit. These estimation methods are applicable to any area classifier, including those defined for the key NFI attributes 1-18, and 20-21.[6] The ratio-of-means estimator that uses no auxiliary information is described in section 3.1.2, and a potential alternative estimator that uses classified satellite imagery as auxiliary information is presented in section 3.1.4.

3.1.2 Estimator 1: No auxiliary information

The approach to estimating classifier class areas is to estimate the proportion of area of the kth classifier within an NFI unit, and then multiply this estimated proportion by the total land area of the unit. This estimation approach can be implemented using the following steps:

1. Obtain from photo plot summary tables or EOSD tables the a_{ki}, the area classified into the kth classifier class in the ith photo plot of total land area a_{Ti} ($i = 1,2,3,\ \dots,\ n$) within an NFI unit.

2. Calculate the average proportion of area of the kth classifier class within the NFI unit:

[1]
$$\hat{p}_k = \frac{\sum_{i=1}^{n} a_{ki}}{\sum_{i=1}^{n} a_{Ti}}$$

3. Estimate the variance of the estimated area proportion of the kth classifier class [Cochran (1977, p. 66, eq. 3.34)]:

[6] Attribute #9 involves change estimation, however, the methods given here are applicable with appropriate definition of the variables; for Attribute #20, the classifier class is either 'native' or 'exotic', and the same formulae are applicable using species count instead of area; and for Attribute #21, classifier class is a seedling batch and we estimate area per batch. The interpretation of the remaining key attributes is straightforward.

[2]
$$\hat{\text{var}}(\hat{p}_k) \doteq \frac{1}{\bar{a}_T^{\,2}}\left(\frac{\left[\sum_{i=1}^{n} a_{ki}^{\,2} - 2\hat{p}_k \sum_{i=1}^{n} a_{ki}a_{Ti} + \hat{p}_k^{\,2}\sum_{i=1}^{n} a_{Ti}^{\,2}\right]}{n(n-1)}\right)$$

4. Estimate the relative standard error (*SE%*) (or coefficient of variation, CV%) of the estimated area proportion of the *k*th classifier class:

[3]
$$SE\%(\hat{p}_k) = \frac{\sqrt{\hat{\text{var}}(\hat{p}_k)}}{\hat{p}_k}100$$

5. Obtain independently the total area of the NFI unit from GIS:

[4]
$$A = NFI\_unit\_Total\_area(ha)$$

6. Estimate the total area A_k (ha) in the *k*th classifier class in the NFI unit:

[5]
$$\hat{A}_k = A\hat{p}_k$$

7. Estimate the variance of the estimated total area in the *k*th classifier class:

[6]
$$\hat{\text{var}}(\hat{A}_k) = A^2\ \hat{\text{var}}(\hat{p}_k)$$

8. Estimate the relative standard error (SE%) of the estimated total area in the *k*th classifier class:

[7]
$$SE\%(\hat{A}_k) = \frac{\sqrt{\hat{\text{var}}(\hat{A}_k)}}{\hat{A}_k}100$$

9. Calculate the approximate (1-α)100% confidence limits of A_k as:[7]

[8]
$$\hat{A}_k \pm t_{\alpha/2,n-1}\sqrt{\hat{\text{var}}(\hat{A}_k)}$$

where α is a stated probability level of a Type I error (usually 0.05), and $t_{\alpha/2}$ is a value from the *t*-distribution with *n*-1 degrees of freedom.

3.1.3 Example 1: Estimating area totals with no auxiliary information

A sample dataset from the Province of Nova Scotia (an NFI unit) is used to illustrate area estimation procedures described in section 3.1.2. It was embellished by including partial plots. There are *n* = 15

[7] Lower confidence limits that go beyond zero (likely with small classifier areas) should be set to zero; more elegant ways to do this could be considered during implementation of these procedures.

complete and partial plots. The data consisted of polygon area and volume estimates by land cover class from the 15 photo plots; these data were summarized by plot and land cover class (Table 4).

Table 4. Example data from Nova Scotia: areas and volumes by land cover class.

NFI Plot No.	*Land Cover Class (lcc*) (k)*	*No. of Polygons*	*Sum of Polygon Areas (ha) (a_{ki})*	*Polygon Total Volume (m^3) (x_{ki})*	*"Poly"*
985121	NLU	16	159.026	0	
985121	NWW	5	1.465	0	
985121	VNU	21	116.394	0	
985121	VNW	6	36.252	0	no
985121	VTU	36	86.865	36547.925	no
985126	NWW	3	0.628	0	
985126	VNU	18	47.923	0	
985126	VNW	6	30.123	0	
985126	VTU	76	321.327	142932.725	
985131	NWW	5	52.474	0	
985131	VNU	15	57.154	0	
985131	VNW	10	54.130	0	
985131	VTU	60	236.243	112128.950	
985136	NLU	2	0.055	0	
985136	NWW	2	60.679	0	
985136	VNW	17	15.829	0	
985136	VTU	52	323.435	187001.275	
991996	MMM	2	162.182	0	
991996	NWW	1	0.348	0	
991996	VNU	22	93.678	0	
991996	VNW	2	2.953	0	
991996	VTU	47	140.843	57245.750	
992001	VNU	6	19.729	0	
992001	VNW	3	5.405	0	
992001	VTU	158	374.863	168305.450	
992006	NLU	5	6.725	0	
992006	NWW	5	7.562	0	
992006	VNU	7	24.244	0	
992006	VNW	8	17.228	0	
992006	VTU	85	344.237	148000.600	no
992011	NWW	5	23.535	0	
992011	VNU	13	26.206	0	
992011	VNW	9	67.878	0	
992011	VTU	58	282.381	139959.925	
992016	NLU	5	13.802	0	
992016	NWW	2	82.600	0	
992016	VNU	21	90.169	0	
992016	VNW	4	6.613	0	
992016	VTU	76	206.821	108154.200	
998876	NLU	3	8.162	0	
998876	NWW	1	0.016	0	
998876	VNU	25	134.196	0	
998876	VTU	68	257.627	125008.425	
998881	NWW	1	3.111	0	
998881	VNW	6	4.426	0	
998881	VTU	73	392.461	167425.325	
998886	NLU	2	3.816	0	

NFI Plot No.	Land Cover Class (lcc*) (k)	No. of Polygons	Sum of Polygon Areas (ha) (a_{ki})	Polygon Total Volume (m^3) (x_{ki})	"Poly"
998886	NWW	3	9.175	0	
998886	VNU	16	99.215	0	
998886	VTU	82	287.792	127176.575	
998891	NLU	2	5.403	0	
998891	NWW	7	0.278	0	
998891	VNU	23	152.908	0	
998891	VNW	1	1.233	0	
998891	VTU	60	240.182	104159.500	
998896	NLU	22	41.947	0	
998896	NWW	15	11.378	0	
998896	VNU	24	53.021	0	
998896	VNW	4	14.236	0	
998896	VTU	74	279.41	132637.100	
998901	NLU	2	0.223	0	
998901	NWW	1	0.090	0	
998901	VNU	7	13.478	0	
998901	VNW	13	52.916	0	
998901	VTU	60	333.293	156920.250	

***Notes:**

1. ***lcc*** = combination of first, second and third level of the NFI land cover classification scheme (see NFI *Land cover classification* document). The MMM stands for missing lcc classification labels.
2. ***Polygon area*** = total area (ha) of all polygons for each *lcc*= *k* within the NFI plot ($= a_{ki} = \sum_{i=1}^{S} a_{is}$), where a_{is} has value zero if polygon does not belong to *k*.
3. ***Polygon total volume*** (**x_{ki}**) = sum of [volume/ha of the *s*th polygon x area of the *s*th polygon] for each *lcc* = *k* within the NFI plot (Equation 18: $\sum_{i=1}^{S} t''_{is} \times a_{is}$), where a_{is} has value zero if a polygon does not belong to *k*.
4. ***Poly*** = indicator variable indicating whether all polygons in the *lcc* are in the NFI unit (*yes* or *no*; *yes* = blank); delete all polygons with ***Poly*** = *no* from the analysis.

The area calculations using Microsoft Excel (Office 2000) are shown in detail for lcc = *k* = VTU (Vegetated, Treed, Upland) below. Copies of the workbooks used here and in the rest of the examples are available upon request from the NFI Office, Pacific Forestry Centre, Victoria, BC.

a) Calculate the following statistics from the data in Table 4:

$$\sum_{i=1}^{n} a_{ki} = 0 + 321.327 + \ldots + 333.293 = 3676.678$$

$$\sum_{i=1}^{n} x_{ki} = 142932.725 + 112128.950 + \ldots + 156920.250 = 1729055.450$$

$$\sum_{i=1}^{n} a_{ki}^{2} = 1096608.077810$$

$$\sum_{i=1}^{n} x_{ki}^{2} = 142932.725^2 + 112128.950^2 + \ldots 156920.250^2 = 243760047729.853729$$

n = 15

a_{Ti} = 276.885, 400.001, … 399.992, 400.000

$$\sum_{i=1}^{n} a_{Ti} = 5532.643$$

$$\bar{a}_T = \frac{5532.643}{15} = 368.84286666666$$

$$\sum_{i=1}^{n} a_{Ti}^{2} = 276.885^2 + 400.001^2 + \ldots 400.000^2 = 2159773.569451$$

$$\sum_{i=1}^{n} a_{ki} \times a_{Ti} = 0 \text{ x } 276.885 + 321.327 \text{ x } 400.001 + \ldots + 333.292 \text{ x } 400.0 = 1470669.206157$$

$$\sum_{i=1}^{n} x_{ki} \times a_{Ti} = 36547.925 \text{ x } 276.885 + 142932.725 \text{ x } 400.001 + \ldots + 156920.250 \text{ x } 400.000$$

$$= 691621217.242600$$

b) Estimate the area totals and their associated variances and confidence limits (the equation numbers correspond to those in section 3.1.2):

[1] $\hat{p}_k = \frac{3676.678}{5532.643} = 0.664542787$

[2] $\hat{\text{var}}(\hat{p}_k) = 0.003351685$

[3] $SE\%(\hat{p}_k) = 8.71\%$

[4] A = 5549000 ha (Approximate area of the Province of Nova Scotia)

[5] $\hat{A}_k = 0.664542787236 \times 5549000 = 3687547.9264$ ha

[6] $\hat{\text{var}}(\hat{A}_k) = 103203062555.263$

[7] $SE\%(\hat{A}_k) = 8.71\%$

[8] Approximate confidence limits for A_k (95%, $t_{15\text{-}1}$) = 2998529.58 ha to 4376566.27 ha

c) Summarize the area estimates and statistics for the *lcc* = NWW and the remaining classifier classes that were calculated also using Microsoft Excel (Table 5).

Table 5. Example area estimates by landcover class (*k*).

Lcc (k)	$\hat{p}_k$	$\hat{\text{var}}(\hat{p}_k)$	$\hat{A}_k$ *(ha)*	$\hat{\text{var}}(\hat{A}_k)$	*Confidence limits for A_k (ha)*	
					Lower	*Upper*
MMM	0.02931	0.000852502	162661.48	26249724359.695	0.00	510155.13
NLU	0.04323	0.000876914	239866.06	27001407442.633	0.00	592299.98
NWW	0.04579	0.000335967	254087.98	10344892735.093	35941.87	472234.10
VNU	0.16779	0.001249976	931059.52	38488513224.612	510284.39	1351834.65
VNW	0.04934	0.000250833	273777.02	7723492395.763	85285.70	462268.34
VTU	0.66454	0.003351685	3687547.93	103203062555.26	2998529.58	4376566.27

3.1.4 Estimator 2: Using auxiliary information

An alternative estimator that may be more precise uses observations from the aerial photos and satellite imagery corresponding to a set of photo plots, to provide the basis for estimation through ratio estimation. This is potentially possible for a limited number of classifiers that can be seen on both the photo and satellite images. (This would be the case where we have photo plots in the EOSD area.)

This proposed ratio estimator would be appropriate under the following conditions [Cochran (1977)]:

1. The relationship between the area estimates based on aerial photo interpretation and the area estimates based on EOSD is well approximated by a straight line through the origin.
2. The variance of the photo-based area estimates along this straight line is proportional to the area estimates based on EOSD.

Further to the prior notation and definitions (section 2.2), let:

A'_k = Total area in the *k*th classifier class within NFI unit estimated from the EOSD census (wall-to-wall) data.

a'_{ki} = *A*rea (ha) classified as the *k*th classifier class (by EOSD) in the *i*th photo plot within the NFI unit.

Then, this alternative estimator could be implemented as follows:

1. Obtain the from the photo plot summary data tables land area in the *k*th classifier class in the *i*th photo plot of total area a_{ki} ($i = 1,2,3, \ldots, n$) within an NFI unit. Obtain from the EOSD tables a corresponding land area in the *k*th classifier class in the *i*th photo plot, a'_{ki} ($i = 1,2,3, \ldots, n$) within the NFI unit. Note that:

[9] a'_{ki} = Sum of areas of pixels in the *k*th classifier class in the *i*th photo plot.

2. Calculate the ratio of estimated area for the *k*th classifier class based on photo plots to estimated area for the same classifier class based on satellite plots:

[10]
$$\hat{R}_k = \frac{\sum_{i=1}^{n} a_{ki}}{\sum_{i=1}^{n} a'_{ki}}$$

3. Estimate the variance of the estimated ratio [Cochran (1977, eq. 6.9-6.13, p. 155)]:

[11]
$$\hat{\text{var}}(\hat{R}_k) \doteq \frac{1}{(\sum_{i-1}^{n} a'_{ki}/n)^2}\left(\frac{\left[\sum_{i=1}^{n} a_{ki}{}^2 + \hat{R}_k{}^2 \sum_{i=1}^{n} a'_{ki}{}^2 - 2\hat{R}_k \sum_{i=1}^{n} a_{ki} a'_{ki}\right]}{n(n-1)}\right)$$

4. Estimate the relative standard error of the ratio estimate:

[12]
$$SE\%(\hat{R}_k) = \frac{\sqrt{\hat{\text{var}}(\hat{R}_k)}}{\hat{R}_k} 100$$

5. Obtain the total area for the kth classifier class based on satellite census data:

[13]
$$A'_k = EOSD\_census\_data \text{ ha}$$

6. Estimate the total area A_k (ha) in the kth classifier class in the NFI unit:

[14]
$$\hat{A}_k = A'_k \hat{R}_k$$

7. Estimate the variance of the estimated total area in the kth classifier class:

[15]
$$\hat{\text{var}}(\hat{A}_k) = (A'_k)^2 \ \hat{\text{var}}(\hat{R}_k)$$

8. Estimate the relative standard error of the estimated total area in the kth classifier class:

[16]
$$SE\%(\hat{A}_k) = \frac{\sqrt{\hat{\text{var}}(\hat{A}_k)}}{\hat{A}_k} 100$$

9. Calculate the approximate (1-α)100% confidence limits of A_k are:

[17]
$$\hat{A}_k \pm t_{\alpha/2,n-1} \sqrt{\hat{\text{var}}(\hat{A}_k)})$$

where α is a stated probability level of a Type I error (usually 0.05), and $t_{\alpha/2}$ is a value from the t-distribution with n-1 degrees of freedom.

Note that it may be necessary to condition the area estimates [14] for all the classifier classes K in the unit to sum to A, using, e.g.:

[17a]
$$\hat{A}_k^* = \hat{A}_k \frac{A}{\sum_{k=1}^{K} \hat{A}_k}$$

where $\hat{A}_k^*$ is the adjusted estimated area for the kth classifier class.

This conditioning will add more error to the estimates. If this is significant, other alternative possibly more complex ways to ensure additivity [e.g., use of multivariate ratio estimators; Cochran (1977, p. 184-186)] could be investigated.

3.1.5 Example 2: Estimating area totals with auxiliary information

A sample dataset provided by the EOSD Project[8] is used to illustrate area estimation procedures described in section 3.1.4. The data consists wall-to-wall EOSD classification in a study area (NFI unit) of A = 13,331,832.01 ha near Prince George, BC, and a sample of 320 photo plots, each with area classifications based on photo interpretation, a_{ki}, and EOSD satellite image classification, a'_{ki}.

[8] These data were provided by Joanne Webber and Mike Wulder of the Pacific Forestry Centre, Canadian Forest Service, Victoria, BC, on 26 November 2003.

The a_{Ti} ($i = 1,2,3, \ldots, 320$) are all equal to 400. The EOSD data are the auxiliary information. The sample area data for the classifier class k = Vegetated Treed are displayed in Figure 1.

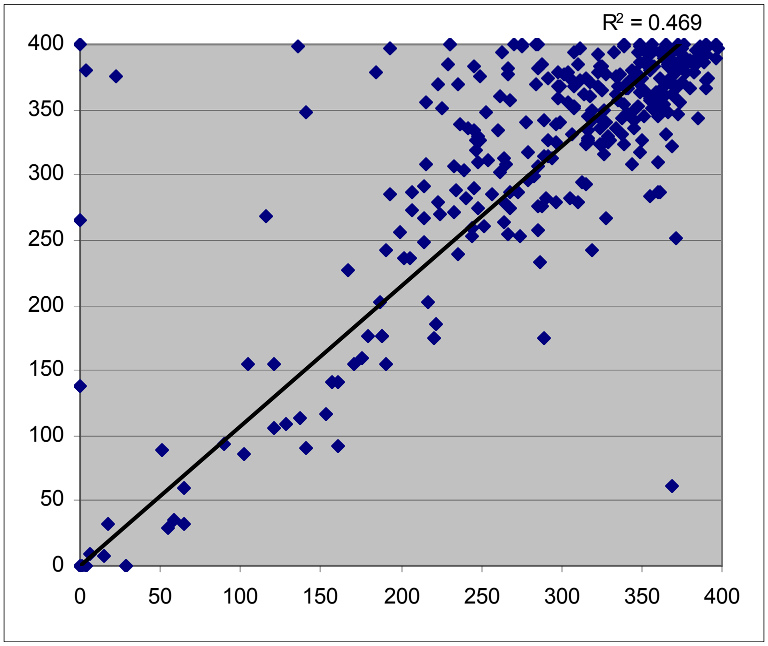

Figure 1. Relationship between the area estimates based on the EOSD (ha; x-axis) and photo-interpreted area estimates (ha; y-axis) for the Vegetated Treed (VT) land cover class, for a pilot project area in British Columbia with N = 320 sample locations. The data trend line through the origin is also shown.

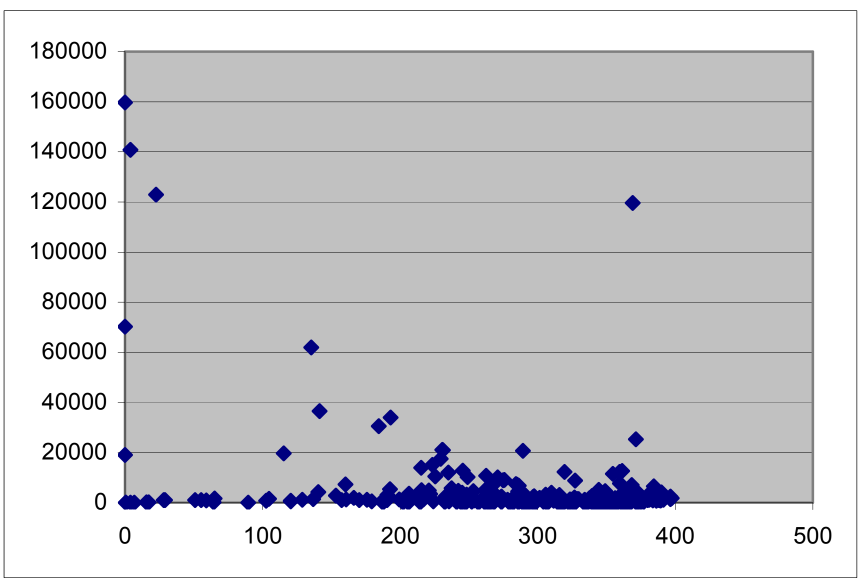

Figure 2. A scatter plot of the squared residuals of a_{ki} (assuming a ratio-of-means model) (y-axis; ha^2) and a'_{ki} (x-axis; ha) for the data in Figure 1. The scatter pattern should ideally resemble an ice cream cone lying on its side (x-axis) and vertex at the origin.

The following procedure was used to check whether the ratio-of-means estimation was appropriate:

a) Plot a_{ki} against a'_{ki}, to check assumption 1 (section 3.1.4).

b) Calculate the ratio-of-means $\hat{R} = \sum_{i=1}^{n} a_{ki} / \sum_{i=1}^{n} a'_{ki}$ and the predicted a_{ki}'s: $\hat{a}_{ki} = \hat{R}a'_{ki}$.

c) Calculate the squared residuals $e_i^2 = (a_{ki} - \hat{a}_{ki})^2$.

d) Plot e_i^2 against a'_{ki}, to check assumption 2 (section 3.1.4).

The relationship between a_{ki} and a'_{ki} appears to be reasonably well approximated by a straight line through the origin (Figure 1). It is unclear whether the variance of the area a_{ki} along this straight line is proportional to a'_{ki} (Figure 2). However, when the observations with $e_i^2 > 60000$ are excluded, (possibly errors in classification) then assumption 2 seems plausible. Despite this uncertainty, we used these data to illustrate the use of the EOSD data for estimating area totals for the Vegetated Treed class in the NFI unit.

a) Calculate the following summary statistics from the data shown in Figure 1:

$$\sum_{i=1}^{n} a_{ki} = 101913.9375000$$

$$\sum_{i=1}^{n} a'_{ki} = 92330.7500000$$

$$\sum_{i=1}^{n} a_{ki}^2 = 35225050.6289062$$

$$\sum_{i=1}^{n} a'^2_{ki} = 29410419.8828125$$

n = 320

$$\sum_{i=1}^{n} a_{ki}a'_{ki} = 31508185.1289062$$

$$\bar{a}_T = 400.0$$

b) Calculate the area total and its associated variance and confidence limits for the Vegetated Treed classifier class in the study area. The equation numbers correspond to those in the section 3.1.4.

[10] $$\hat{R}_k = \frac{101913.9375}{92330.75000} = 1.10379193822$$

[11] $$\hat{\mathrm{var}}(\hat{R}_k) = 0.000176560589$$

[12] $$SE\%(\hat{R}_k) = 1.20\%$$

[13] $$A'_k = 9617934.94 \text{ ha}$$

[14] $$\hat{A}_k = 1.10379193822 \times 9617934.94 = 10616199.04912 \text{ ha}$$

[15] $$\hat{\mathrm{var}}(\hat{A}_k) = 1633679446.5909$$

[16] $$SE\%(\hat{A}_k) = 1.20\%$$ [9]

[17] Approximate confidence limits of A_k (95%, $t_{320\text{-}1}$) =10364817.68 ha to 10867580.41 ha

3.2 Tree attribute totals

3.2.1 Estimation inputs

Inputs for estimating tree attribute totals are obtained from the photo plot summary data tables in the NFI database or possibly EOSD tables. These data include stand age, stand height, crown closure, volume and biomass. The full list of photo plot attributes is available in the NFI *National Standards for Photo Plots Data Dictionary* document.

Attribute estimation methods are outlined below for a classifier class within an NFI unit. These estimation methods are applicable to any attribute, including those defined for the key NFI attributes 22-24.

3.2.2 Estimator

Estimation of attributes totals based on data from photo plots is similar to estimation of area totals (section 3.1.2). This estimation can be implemented as follows:

1. Obtain from the photo plot summary data tables or EOSD tables the t''_{is}, the polygon average or per-hectare values and corresponding polygon areas a_{is} ($s = 1,2,3, \ldots, S$) from the ith photo plot of area a_{Ti} ($i = 1,2,3, \ldots, n$). Estimate the attribute plot total, x_{ki}, for the ith plot within an NFI unit as (Note: the x_{ki} is set equal to zero if the polygon does not belong to the kth classifier class):

[18] $$x_{ki} = \sum_{s=1}^{S} (t''_{is} a_{is})$$

2. Calculate the estimated unit average or per-hectare value for the kth classifier class from all the photo plots within the NFI unit:

[19] $$\bar{x}_k = \frac{\sum_{i=1}^{n} x_{ki}}{\sum_{i=1}^{n} a_{Ti}}$$

3. Estimate the variance of the estimated photo plots average or per-hectare value [Cochran (1977, eq. 6.9 - 6.13, p. 155)]:

[20] $$\hat{\text{var}}(\bar{x}_k) \doteq \frac{1}{\bar{a}_T^{\,2}} \left(\frac{\left[\sum_{i=1}^{n} x_{ki}^{\,2} + \bar{x}_k^{\,2} \sum_{i=1}^{n} a_{Ti}^{\,2} - 2\bar{x}_k \sum_{i=1}^{n} x_{ki} a_{Ti} \right]}{n(n-1)} \right)$$

[9] The SE% when auxiliary information is used is slightly less than that with no auxiliary information (1.63%).

4. Estimate the relative standard error of the estimated average or per-hectare value for the *k*th classifier class:

[21] $$SE\%(\bar{x}_k) = \frac{\sqrt{\hat{\text{var}}(\bar{x}_k)}}{\bar{x}_k}100$$

5. Estimate the photo-based total in the *k*th classifier class in the NFI unit:

[22] $$\hat{X}_k = A\bar{x}_k$$

6. Estimate the variance of the estimated total for the *k*th classifier class:

[23] $$\hat{\text{var}}(\hat{X}_k) = A^2\ \hat{\text{var}}(\bar{x}_k)$$

7. Estimate the relative standard error of the estimated total for the *k*th classifier class in the NFI unit:

[24] $$SE\%(\hat{X}_k) \doteq \frac{\sqrt{\hat{\text{var}}(\hat{X}_k)}}{\hat{X}_k}100$$

8. Calculate the approximate (1-α)100% confidence limits of X_k:

[25] $$\hat{X}_k \pm t_{\alpha/2,n-1}\sqrt{\hat{\text{var}}(\hat{X}_k)}$$

where α is a stated probability level of a Type I error (usually 0.05), and $t_{\alpha/2}$ is a value from the *t*-distribution with *n*-1 degrees of freedom.

3.2.3 Example 3: Estimating tree volume totals with no auxiliary information

The embellished sample dataset from the Province of Nova Scotia (an NFI unit) shown in Table 4 is again used to illustrate the procedures for estimating tree volume totals estimation described in section 3.2.2. These data include polygon area and volume estimates by land cover class from 15 photo plots.

a) Recall the statistics estimated in Example 1.

b) Estimate volume totals and their associated variances and confidence limits.
The computational steps for estimating tree volume totals are shown for lcc = *k* = VTU (Vegetated, Treed, Upland). The equation numbers correspond to those in section 3.2.2.

[19] $\bar{x}_k = 312.51889$ m^3/ha

[20] $\hat{\text{var}}(\bar{x}_k) = 784.45804190741$

[21] $SE\%(\bar{x}_k) = 8.96\%$

[22] $\hat{X}_k$ = 312.518890157922 x 5549000 = 1734167321.4863 m^3

[23] $\hat{\mathrm{var}}\left(\hat{X}_k\right) = 24154562136046000$

[24] $SE\%\left(\hat{X}_k\right) = 8.96\%$

[25] Approximate confidence limits for X_k (95%, $t_{15\text{-}1}$) = 1400829896.41 m^3 to 2067504746.57 m^3

c) Summarize the volume statistics as follows (the volume estimates are zero for all other *lcc*):

Lcc (k)	$\bar{x}_k$ *(m³/ha)*	$\hat{\mathrm{var}}(\bar{x}_k)$	$\hat{X}_k$ *(m³)*	$\hat{\mathrm{var}}(\hat{X}_k)$	*Confidence limits of* X_k *(m³)*	
					Lower	Upper
VTU	312.519	784.4580	1734167321	24154562136046000	1400829896	2067504747

We can also estimate the classifier-class average. This can be done as follows:

1. Estimate the average "per-hectare" attributes (e.g. volume/ha) within the *k*th classifier class area in the NFI unit:

[26]
$$\hat{\bar{x}}_k = \frac{\hat{X}_k}{\hat{A}_k} = \frac{A\bar{x}_k}{A\hat{p}_k} = \frac{\bar{x}_k}{\hat{p}_k} = \frac{312.51889016}{0.6645428} = 470.277\ \mathrm{m^3/ha}$$

2. Estimate the variance of the estimated average value within the classifier area [Cochran (1977, eq. 6.95, p. 184)]:

[27]
$$\hat{\mathrm{var}}(\hat{\bar{x}}_k) \doteq \hat{\bar{x}}_k^{\,2}\left[\frac{\hat{\mathrm{var}}(\bar{x}_k)}{\bar{x}_k^{\,2}} + \frac{\hat{\mathrm{var}}(\hat{p}_k)}{\hat{p}_k^{\,2}} - \frac{2\,\hat{\mathrm{cov}}(\bar{x}_k, \hat{p}_k)}{\bar{x}_k\hat{p}_k}\right] = 162.83321851$$

since, using the result in Cochran [(1977, eq. 6.90, p. 181)]:

$$\hat{\mathrm{cov}}(\bar{x}_k, \hat{p}_k) \doteq \frac{1}{n(n-1)\bar{a}_T^{\,2}}\left[\sum_{i=1}^{n} x_{ki}a_{ki} - \hat{p}_k\sum_{i=1}^{n} x_{ki}a_{Ti} - \bar{x}_k\sum_{i=1}^{n} a_{ki}a_{Ti} + \hat{p}_k\bar{x}_k\sum_{i=1}^{n} a_{Ti}^{\,2}\right] = 1.545693454$$

3. Estimate the relative standard error of the estimated average within the *k*th classifier class area in the NFI unit:

[28]
$$SE\%(\hat{\bar{x}}_k) \doteq \frac{\sqrt{\hat{\mathrm{var}}(\hat{\bar{x}}_k)}}{\hat{\bar{x}}_k}100 = 2.71\%$$

4. Calculate the approximate (1-α)100% confidence limits of X_k / A_k:

[29] $$\hat{\hat{x}}_k \pm t_{\alpha/2,n-1}\sqrt{\hat{\mathrm{var}}(\hat{\hat{x}}_k)} = 442.91\ \mathrm{m^3/ha\ to\ } 497.65\ \mathrm{m^3/ha}$$

where α is a stated probability level of a Type I error (here $\alpha = 0.05$), and $t_{\alpha/2}$ is a value from the *t*-distribution with n-1 (15-1) degrees of freedom.

4. Estimates from ground plots

4.1. Estimation inputs

Attribute totals of interest from ground plots include per-hectare woody-debris and tree volume and biomass, and carbon mass of trees, stumps, woody-debris and soil. Inputs for estimating the totals of these attributes are obtained from the ground plot summary data tables in the NFI database. These summaries were obtained from individual tree, woody debris or soil measurements, which were summarized into stand average and per-hectare values using standard algorithms. These attributes are listed in the NFI *National Standards for Ground Plots Data Dictionary* document.

Estimation methods are described with reference to only tree and woody debris attributes. However, these estimation methods are applicable to any continuous attribute, including the key NFI attributes 19 and 22-25.[10] Attribute totals may be estimated using data from:

1. The LTP or STP, or transects (SWD or CWD).
2. Combined ground plots (LTP and STP), or transects (SWD and CWD).

Estimation methods are described below for case 1 (sections 4.2) and case 2 (section 4.3).

4.2 Estimator using data from the LTP

We assume that the attribute ground data come from the LTP, as, for example, we were interested in the volume of large trees only. The estimation methods are identical if we use the STP, or in the case of woody debris we would use the SWD or CWD transects. The situation where data are combined from two line-transects or two plots (e.g., volume of all trees from both the LTP and STP) is discussed in section 4.4.

The approach to estimating attribute averages or totals for a classifier class in an NFI unit is similar to that for estimating tree attributes from photo plots (section 3.2). That is, post-stratify the ground plots into a classifier class, estimate the population average or per-hectare in the classifier class, and then multiply this estimated average by the land area of the NFI unit. The post-stratification is based on photo plot information or other verification about the ground plot center only since the ground plots are not mapped according to photo-plot polygons.

This estimation can be implemented using the following steps:

1. Post-stratify the ground plots based on photo plot information into classifier classes of interest for a given NFI unit.

2. Obtain from the ground plot summary data tables the t_j , the average or per-hectare value for the *j*th ground plot of area b_j ($j = 1,2,3,\ \ldots,\ m$). Estimate the attribute plot total, y_{kj} , for the *j*th plot within an

[10] Attribute #19 is reported as biodiversity indices, which can be treated as the observations in the estimation; and Attribute #25 involves change estimation, however, the methods given here are applicable with appropriate definition of the variables. Interpretation of the remaining attributes is straightforward.

NFI unit as (Note: the y_{kj} is set to equal zero if the jth plot does not belong to the kth classifier class, or is non-forested):

[30]
$$y_{kj} = t_j b_j$$

3. Calculate the estimated unit average or per-hectare value for kth classifier class within the NFI unit:

[31]
$$\bar{y}_k = \frac{\sum_{j=1}^{m} y_{kj}}{\sum_{j=1}^{m} b_j}$$

4. Estimate the variance of the estimated average or per-hectare value [Cochran (1977, eq. 6.9-6.13, p. 155)]:

[32]
$$\hat{\text{var}}(\bar{y}_k) \doteq \frac{1}{\bar{b}^2}\left(\frac{\left[\sum_{j=1}^{m} y_{kj}^{\ 2} + \bar{y}_k^{\ 2}\sum_{j=1}^{m} b_j^{\ 2} - 2\bar{y}_k \sum_{j=1}^{m} y_{kj} b_j\right]}{m(m-1)}\right)$$

5. Estimate the relative standard error of the estimated average or per-hectare value for the kth classifier class:

[33]
$$SE\%(\bar{y}_k) = \frac{\sqrt{\hat{\text{var}}(\bar{y}_k)}}{\bar{y}_k} 100$$

6. Estimate the ground-based population total for the kth classifier class in the NFI unit:

[34]
$$\hat{Y}_k = A\bar{y}_k$$

7. Estimate the variance of the estimated total for the kth classifier class:

[35]
$$\hat{\text{var}}(\hat{Y}_k) \doteq A^2\ \hat{\text{var}}(\bar{y}_k)$$

8. Estimate the relative standard error of the estimated total in the NFI unit:

[36]
$$SE\%(\hat{Y}_k) = \frac{\sqrt{\hat{\text{var}}(\hat{Y}_k)}}{\hat{Y}_k} 100$$

9. Calculate the approximate (1-α)100% confidence limits of Y_k:

[37]
$$\hat{Y}_k \pm t_{\alpha/2,m-1}\sqrt{\hat{\text{var}}(\hat{Y}_k)}$$

where α is a stated probability level of a Type I error (usually 0.05), and $t_{\alpha/2}$ is a value from the *t*-distribution with *m*-1 degrees of freedom.

4.3 Example 4: Estimating tree volume totals using the LTP

An artificial dataset was created from permanent sample data from the Province of Nova Scotia (an NFI unit) to illustrate volume estimation procedures described in section 4.2. There are $m = 15$ plots inside the NFI unit (Table 6). Four of these plots had a different land cover class (*lcc=k=NLU, VNW, and NVU*) than the class of interest (VTU) in the following examples. Blanks in the last column in Table 6 indicate that the ground plot is in the NFI unit.

Table 6. Example dataset of NFI ground plots.

NFI Plot No.	*Lcc (k)*	*Area of LTP (ha)* (b_j)	*Area of STP (ha)* (b'_j)	*Volume in LTP* (m^3/ha) (t_j)	*Volume in STP* (m^3/ha) (t'_j)	*Photo Plot area inside NFI unit (ha)* (a_{Tj})	*Ground Plot in NFI unit*
985121	NLU	0.0404	0.004	0.0000	0.0000	276.885	
985126	VNW	0.0404	0.004	0.0000	0.0000	400.001	
985131	VTU	0.0404	0.004	319.8678	32.2868	400.001	
985136	VTU	0.0404	0.004	382.7619	38.6762	399.998	
991996	VTU	0.0404	0.004	357.1651	36.2165	400.004	
992001	VTU	0.0404	0.004	618.3421	62.4342	399.997	
992006	NLU	0.0404	0.004	0.0000	0.0000	55.759	
992011	VTU	0.0404	0.004	322.0941	33.0094	400.000	
992016	VTU	0.0404	0.004	300.3875	30.9387	400.005	
998876	VTU	0.0404	0.004	338.4368	34.8437	400.001	
998881	VTU	0.0404	0.004	329.2257	34.0226	399.998	
998886	VTU	0.0404	0.004	378.3308	39.0331	399.998	
998891	VNU	0.0404	0.004	0.0000	0.0000	400.004	
998896	VTU	0.0404	0.004	518.876	53.2876	399.992	
998901	VTU	0.0404	0.004	300.5111	31.5511	400.000	

a) Calculate the following statistics:

t_j = 0, 0, 319.86779, …, 300.51108

b_j = 0.0404, 0.0404, … 0.0404

[30] $y_{kj} = t_j b_j$ = 0, 0, 12.922658716, 15.46357955, … , 12.140647632

m = 15

$\sum_{j=1}^{m} y_{kj}$ = 12.922658716 + 15.46357955 + … + 12.140647632= 168.306349904

$\sum_{j=1}^{m} b_j$ = 0.0404 + 0.0404 … + 0.0404 = 0.6060

$(\sum_{j=1}^{m} a_{Tj})/m$ = [276.885+400.001+ … + 400.000]/15 = 5532.643/15 ~368.843

$$\bar{b} = \frac{0.6060}{15} = 0.0404$$

$\sum_{j=1}^{m} y_{kj}^{2}$ =12.922658716^2 + 15.46357955^2 + … + 12.140647632^2 = 2739.2795056

$\sum_{j=1}^{m} y_{kj} b_j$ = (12.922658716 x 0.0404) + (15.46357955 x 0.0404) + … + (12.140647632 x 0.0404)
= 6.799576536

$\sum_{j=1}^{m} b_j^{2}$ = 0.0404^2 + 0.0404^2 + … + 0.0404^2 = 0.0244824

b) Calculate the volume totals and their associated variances and confidence limits. The equation numbers correspond to those in section 4.2.

[31] $\bar{y}_k = \dfrac{168.306349904}{0.6060} = 277.733250$ m^3/ha

[32] $\hat{\text{var}}(\bar{y}_k) = 2482.2817648$

[33] $SE\%(\bar{y}_k) = 17.94\%$

[34] $\hat{Y}_k = 277.733250 \times 5549000 = 1541141807.9493$ m^3

[35] $\hat{\text{var}}(\hat{Y}_k) = 76432933215783200.0$

[36] $SE\%(\hat{Y}_k) = 17.94\%$

[37] Approximate confidence limits for Y_k (95%, $t_{15\text{-}1}$) = 948182685.38 m^3 to 2134100930.52 m^3.

c) Summarize the volume statistics as follows:

Lcc (k)	$\bar{y}_k$ *(m³/ha)*	$\hat{\text{var}}(\bar{y}_k)$	$\hat{Y}_k$ *(m³)*	$\hat{\text{var}}(\hat{Y}_k)$	*Confidence limits for* Y_k *(m³)*	
					Lower	*Upper*
VTU	277.733	2482.28	1541141808	76432933215783200	948182685	2134100930

Note that we can also estimate the classifier-class average. This can be done as follows:

1. Estimate the average "per-hectare" attributes (e.g. volume/ha) within the *k*th classifier class area in the NFI unit:

[38]
$$\hat{\bar{y}}_k = \frac{\hat{Y}_k}{\hat{A}_k} = \frac{A\bar{y}_k}{A\hat{p}_k} = \frac{\bar{y}_k}{\hat{p}_k} = \frac{277.7332507}{0.664542787} = 417.9313296 \text{ m}^3\text{/ha}$$

2. Estimate the variance of the estimated average value within the classifier class area [Cochran (1977, eq. 6.95, p. 184)]:

[39]
$$\hat{\mathrm{var}}(\hat{\bar{y}}_k) \doteq \hat{\bar{y}}_k^{\,2}\left[\frac{\hat{\mathrm{var}}(\bar{y}_k)}{\bar{y}_k^{\,2}} + \frac{\hat{\mathrm{var}}(\hat{p}_k)}{\hat{p}_k^{\,2}} - \frac{2\,\hat{\mathrm{cov}}(\bar{y}_k, \hat{p}_k)}{\bar{y}_k \hat{p}_k}\right]$$

$$= 751.5896502$$

where, using the result in Cochran [(1977, eq. 6.90, p. 181)]:

$$\hat{\mathrm{cov}}(\bar{y}_k, \hat{p}_k) \doteq \frac{1}{(m)(m-1)\bar{b}\left(\sum_{j=1}^{m} a_{Tj} / m\right)}\left[\sum_{j=1}^{m} y_{kj}a_{kj} - \bar{y}_k \sum_{j=1}^{m} a_{kj}b_j - \hat{p}_k \sum_{j=1}^{m} y_{kj}a_{Tj} + \bar{y}_k\hat{p}_k \sum_{j=1}^{m} a_{Tj}b_j\right]$$
$$= 3.273018851$$

3. Estimate the relative standard error of the estimated average in the classifier class area:

[40]
$$SE\%(\hat{\bar{y}}_k) = \frac{\sqrt{\hat{\mathrm{var}}(\hat{\bar{y}}_k)}}{\hat{\bar{y}}_k}100 = 6.56\%$$

4. Calculate the approximate (1-α)100% confidence limits of Y_k / A_k:

[41]
$$\hat{\bar{y}}_k \pm t_{\alpha/2,m-1}\sqrt{\hat{\mathrm{var}}(\hat{\bar{y}}_k)} = 359.13 \text{ m}^3\text{/ha to } 476.73 \text{ m}^3\text{/ha}$$

where α is a stated probability level of a Type I error (here α = 0.05), and $t_{\alpha/2}$ is a value from the *t*-distribution with *m*-1 (15-1) degrees of freedom.

4.4 Estimator using combined data from the LTP and STP

As indicated earlier, data to estimate tree and other attributes may be a combination of data from the LTP and STP (or the SWD and CWD line transects) at a given sample location. For example, total volume of all trees (> 1.3m) is estimated with data from both the LTP and the STP; and total woody debris biomass is estimated with data from both coarse and small woody debris line-transect portions. In this case, the estimated area-weighted average or per-hectare values from the two plots are summed to get an estimated overall average or per-hectare value, which is them multiplied by the total area of the NFI unit to obtain the overall totals.

Methods to handle these situations are similar to the LTP approach described in section 4.2, and are described below.

1. Post-stratify the ground plots into classifier classes of interest for a given NFI unit. Recall that the post-stratification is based on photo plot information or other verification about the ground plot center since the ground plots are not mapped according to photo-plot polygons.

2. Obtain from the ground plot summary data tables the t_j and t'_j, the average or per-hectare value for the *j*th ground plot of area b_j (LTP) and b'_j (STP),[11] respectively (*j = 1,2,3, ..., m*). Estimate the plot attribute totals, y_{kj} and y'_{kj}, at the *j*th sample location for the LTP and STP, respectively, within an NFI unit as (Note: y_{kj} and y'_{kj} are set equal to zero if the *j*th plot does not belong to the *k*th classifier class, or is non-forested):

[42]
$$y_{kj} = t_j b_j$$
$$y'_{kj} = t'_j b'_j$$

3. Calculate the area-weighted averages or per-hectare values for the LTP and STP for the *k*th classifier class within the NFI unit:

[43]
$$\bar{y}_k = \frac{\sum_{j=1}^{m} y_{kj}}{\sum_{j=1}^{m} b_j}$$
$$\bar{y}'_k = \frac{\sum_{j=1}^{m} y'_{kj}}{\sum_{j=1}^{m} b'_j}$$

4. Estimate the variance of the estimated averages or per-hectare values [Cochran (1977, eq. 6.9-6.13, p. 155)]:

[44]
$$\hat{\text{var}}(\bar{y}_k) \doteq \frac{1}{\bar{b}^2}\left(\frac{\left[\sum_{j=1}^{m} y_{kj}{}^2 + \bar{y}_k{}^2 \sum_{i=1}^{m} b_j{}^2 - 2\bar{y}_k \sum_{j=1}^{m} y_{kj} b_j\right]}{m(m-1)}\right)$$
$$\hat{\text{var}}(\bar{y}'_k) \doteq \frac{1}{\bar{b}'^2}\left(\frac{\left[\sum_{j=1}^{m} y'_{kj}{}^2 + \bar{y}'^2_k \sum_{i=1}^{m} b'^2_j - 2\bar{y}'_k \sum_{j=1}^{m} y'_{kj} b'_j\right]}{m(m-1)}\right)$$

[11] The estimators provided here are also applicable to combined line transects, for small and coarse woody debris.

5. Calculate the combined estimated unit average or per-hectare values for all the ground sample locations for the *k*th classifier class within the NFI unit ("*c*" stands for combined):

[45]
$$\bar{y}_k^{\,c} = \bar{y}_k + \bar{y}'_k$$

6. Estimate the variance of the combined estimated average or per-hectare value [Freese (1962, p. 18-19) [12]]:

[46]
$$\hat{\text{var}}(\bar{y}_k^{\,c}) = \hat{\text{var}}(\bar{y}_k) + \hat{\text{var}}(\bar{y}'_k) + 2\,\hat{\text{cov}}(\bar{y}_k, \bar{y}'_k)$$

where, using the result in Cochran (1977, eq. 6.90, p. 181):

$$\hat{\text{cov}}(\bar{y}_k, \bar{y}'_k) \doteq \frac{1}{m(m-1)\bar{b}\,\bar{b}'}\left[\sum_{j=1}^{m} y_{kj}y'_{kj} - \bar{y}_k\sum_{j=1}^{m} y'_{kj}b_j - \bar{y}'_k\sum_{j=1}^{m} y_{kj}b'_j + \bar{y}_k\bar{y}'_k\sum_{j=1}^{m} b_j b'_j\right]$$

7. Estimate the relative standard error of the estimated average or per-hectare value for the *k*th classifier class:

[47]
$$SE\%(\bar{y}_k^{\,c}) = \frac{\sqrt{\hat{\text{var}}(\bar{y}_k^{\,c})}}{\bar{y}_k^{\,c}}100$$

8. Estimate the ground-based population total for the *k*th classifier class in the NFI unit:

[48]
$$\hat{Y}_k^{\,c} = A\bar{y}_k^{\,c}$$

9. Estimate the variance estimate of the estimated total for the *k*th classifier class:

[49]
$$\hat{\text{var}}(\hat{Y}_k^{\,c}) \doteq A^2\,\hat{\text{var}}(\bar{y}_k^{\,c})$$

10. Estimate the relative standard error (SE%) of the estimated total for the *k*th classifier class in the NFI unit:

[50]
$$SE\%(\hat{Y}_k^{\,c}) = \frac{\sqrt{\hat{\text{var}}(\hat{Y}_k^{\,c})}}{\hat{Y}_k^{\,c}}100$$

11. Calculate the approximate (1-α)100% confidence limits of Y_k:

[51]
$$\hat{Y}_k^{\,c} \pm t_{\alpha/2,m-1}\sqrt{\hat{\text{var}}(\hat{Y}_k^{\,c})}$$

where α is a stated probability level of a Type I error (usually 0.05), and $t_{\alpha/2}$ is a value from the *t*-distribution with *m*-1 degrees of freedom.

[12] Freese, F. 1962. Elementary forest sampling. Agriculture handbook no. 232. USDA Forest Service.

4.5 Example 5: Estimating tree volume totals using combined LTP and STP

The example data in section 4.3 are used to illustrate the combined estimation approach.

a) Calculate the following statistics:

t'_j = 0, 0, , 32.2868, … , 31.5511

b'_j = 0.0040, 0.0040, … , 0.0040

[42] $y'_{kj} = t'_j b'_j$ = 0, 0, 0.1291472, 0.1547048, … , 0.1262044

m = 15

$\sum_{j=1}^{m} y'_{kj}$ = 0.1291472 + 0.1547048 + 0.144866+ … + 0.1262044 = 1.7051996

$\sum_{j=1}^{m} b'_j$ = 0.004 + 0.004 … + 0.004 = 0.0600

[43] $\bar{y}'_k = \frac{1.70519961}{0.0600} = 28.41999333$

$\bar{b}' = \frac{0.0600}{15} = 0.0040$

$\sum_{j=1}^{m} y'^{2}_{kj}$ = $0.1291472^2 + 0.1547048^2 + 0.144866^2 + + \ldots + 0.1262044^2$ = 0.2804002817544

$\sum_{j=1}^{m} y'_{kj} b'_j$ = (0.1291472 x 0.004) +(0.1547048 x 0.004) + (0.144866 x 0.004) … + (0.1262044 x 0.004)= 0.0068207984

$\sum_{j=1}^{m} b'^{2}_j$ = $0.004^2 + 0.004^2 + \ldots + 0.004^2$ = 0.0002400

$\sum_{j=1}^{m} y_{kj} y'_{kj}$ = 27.7127864591

$\sum_{j=1}^{m} y'_{kj} b_j$ = 0.0688900638

$$\sum_{j=1}^{m} y_{kj} b'_j = 0.6732253996$$

$$\sum_{j=1}^{m} b_j b'_j = 0.0024240$$

[44] $\hat{\text{var}}(\bar{y}'_k) = 25.75989187$

b) Calculate the combined volume totals and their associated variances and confidence limits. The equation numbers correspond to those in section 4.4.

[45] $\bar{y}_k^{\,c} = 28.4199933 + 277.7332507 = 306.1532440$

[46] $\hat{\text{var}}(\bar{y}_k^{\,c}) = 3013.6831636678$

$\hat{\text{cov}}(\bar{y}_k, \bar{y}'_k) = 252.8207534830$

[47] $SE\%(\bar{y}_k^{\,c}) = 17.93\%$

[48] $\hat{Y}_k^c = 306.1532440 \times 5549000 = 1698844350.956$ m^3

[49] $\text{var}(\hat{Y}_k^c) = 92795526779445000$

[50] $SE\%(\hat{Y}_k^{\,c}) = 17.93\%$

[51] Approximate confidence limits for Y_k (95%, t_{15-1}) = 1045491166.08 m^3 to 2352197535.83 m^3

c) Summarize the volume statistics as follows:

Lcc (k)	$\bar{y}_k^{\,c}$ *(m^3/ha)*	$\hat{\text{var}}(\bar{y}_k^{\,c})$	$\hat{Y}_k^{\,c}$ *(m^3)*	$\hat{\text{var}}(\hat{Y}_k^{\,c})$	*Confidence limits of Y_k (m^3)*	
					Lower	*Upper*
VTU	306.15	3013.68	1698844351	92795526779445000	1045491166	2352197536

Note that we can also estimate the classifier-class average. This can be done as follows:

1. Estimate the average "per-hectare" attribute (e.g. volume/ha) within the *k*th classifier class area in the NFI unit:

[52]
$$\hat{\bar{y}}_k^{\,c} = \frac{\hat{Y}_k}{\hat{A}_k} = \frac{A\bar{y}_k^{\,c}}{A\hat{p}_k} = \frac{\bar{y}_k^{\,c}}{\hat{p}_k} = \frac{306.1532440}{0.6645428} = 460.6975651 \text{ m}^3/\text{ha}$$

2. Estimate the variance of the estimated average value within the classifier class area [Cochran (1977, eq. 6.95, p. 184)]:

[53]
$$\hat{\mathrm{var}}(\hat{\bar{y}}_k^{\,c}) \doteq (\hat{\bar{y}}_k^{\,c})^2 \left[\frac{\hat{\mathrm{var}}(\bar{y}_k^{\,c})}{(\bar{y}_k^{\,c})^2} + \frac{\hat{\mathrm{var}}(\hat{p}_k)}{\hat{p}_k^{\,2}} - \frac{2\,\hat{\mathrm{cov}}(\bar{y}_k^{\,c}, \hat{p}_k)}{\bar{y}_k^{\,c}\,\hat{p}_k} \right]$$

$$= 906.625322$$

where

$$\hat{\mathrm{cov}}(\bar{y}_k^{\,c}, \hat{p}_k) \doteq \hat{\mathrm{cov}}(\bar{y}_k, \hat{p}_k) + \hat{\mathrm{cov}}(\bar{y}'_k, \hat{p}_k) = 3.608300988$$

since, using the result in Cochran [(1977, eq. 6.90, p. 181)]:

$$\hat{\mathrm{cov}}(\bar{y}_k, \hat{p}_k) \doteq \frac{1}{m(m-1)\bar{b}(\sum_{j=1}^{m} a_{Tj} / m)} \left[\sum_{j=1}^{m} y_{kj} a_{kj} - \bar{y}_k \sum_{j=1}^{m} a_{kj} b_j - \hat{p}_k \sum_{j=1}^{m} y_{kj} a_{Tj} + \bar{y}_k \hat{p}_k \sum_{j=1}^{m} a_{Tj} b_j \right]$$

= 3.2730188511 (from Example 4, Eq. 39)

$$\hat{\mathrm{cov}}(\bar{y}'_k, \hat{p}_k) \doteq \frac{1}{m(m-1)\bar{b}'(\sum_{j=1}^{m} a_{Tj} / m)} \left[\sum_{j=1}^{m} y'_{kj} a_{kj} - \bar{y}'_k \sum_{j=1}^{m} a_{kj} b'_j - \hat{p}_k \sum_{j=1}^{m} y'_{kj} a_{Tj} + \bar{y}'_k \hat{p}_k \sum_{j=1}^{m} a_{Tj} b'_j \right]$$

$$= 0.335282137$$

3. Estimate the relative standard error of the estimated average in the classifier class area:

[54]
$$SE\%(\hat{\bar{y}}_k^{\,c}) = \frac{\sqrt{\hat{\mathrm{var}}(\hat{\bar{y}}_k^{\,c})}}{\hat{\bar{y}}_k^{\,c}} 100 = 6.54\%$$

4. Calculate the approximate (1-α)100% confidence limits of Y_k / A_k:

[55]
$$\hat{\bar{y}}_k^{\,c} \pm t_{\alpha/2,m-1} \sqrt{\hat{\mathrm{var}}(\hat{\bar{y}}_k)} = 396.12 \text{ m}^3\text{/ha to } 525.28 \text{ m}^3\text{/ha}$$

where α is a stated probability level of a Type I error (here α = 0.05), and $t_{\alpha/2}$ is a value from the *t*-distribution with *m*-1 (15-1) degrees of freedom.

5. Estimates from combined photo plots and ground plots

5.1 Estimation inputs

In NFI units where the ground plots have been established at a subsample of the photo plot locations in a double sampling (two-phase) scheme, it might be possible to increase precision of the estimated attribute totals. This could be achieved by using the ratio-of-means relationship between the ground values (y) and the corresponding photo estimates (x) to correct the (unadjusted) population sample average obtained from photo plots alone for a classifier class (estimator $\overline{x}_k$ of section 3.2.2).[13] The adjusted estimated population average is then expanded to the population total in the usual manner.

The *y*-values are obtained from the measurements on the ground plots (LTP and /or STP). The *x*-values are obtained from the photo-plot polygon estimates associated with the ground plot areas projected onto the corresponding approximate location on the photo plot. These *x*-values are subject to the usual photo-interpreter error, as well as the error arising from the interpreter looking at an entire polygon rather than just the location of the ground plot on the photo-plot polygon. We assume that both these sources of error are negligible. The unadjusted estimated population average is obtained using the attribute estimates on all polygons in all the photo plots in the NFI unit.

5.2 Estimator

The estimation methods using the LTP are described below; the methods using combined LTP and STP can be obtained in an analogous manner.

Keeping the notation already defined, let:

t''_{js} = Attribute stand average (e.g., site age) or per-hectare estimate (e.g., volume) for the *s*th photo-plot polygon associated with the *j*th ground plot (LTP) in the NFI unit.

x'_{kj} = Photo-based attribute total estimate associated with the *j*th ground plot for the *k*th classifier class within the NFI unit. (Note: The t''_j value equals zero if, according to photo plot information, the ground plot centre is outside the *k*th classifier class, or is non-forested.)

= $t''_{js} \; b_j$.

Then, the estimation can be implemented in the following steps:

1. Calculate the ground-based weighted estimate of the stand average or per-hectare value over all the ground locations for the *k*th classifier class within the NFI unit:

[56]
$$\overline{y}_k = \frac{\sum_{j=1}^{m} y_{kj}}{\sum_{j=1}^{m} b_j}$$

2. Calculate the photo-based weighted estimate of the stand average or per-hectare value associated with all the ground plot locations for the *k*th classifier class within the NFI unit:

[57]
$$\overline{x}'_k = \frac{\sum_{j=1}^{m} x'_{kj}}{\sum_{j=1}^{m} b_j}$$

3. Calculate the estimated stand average or per-hectare value of all the polygons in all the photo plots for the *k*th classifier class within the NFI unit:

[13] As indicated earlier, the ratio estimators are biased, however, this bias may be more than compensated for by the potentially increased precision. The nature of the relationship between *y* and *x* should be investigated with actual data, to determine if more complex relationships than the ratio would be more appropriate, and if indeed use of this additional information can result in any appreciable gains in precision (for a given cost).

[58]

$$\bar{x}_k = \frac{\sum_{j=1}^{n} x_{ki}}{\sum_{j=1}^{n} a_{Ti}}$$

4. Calculate the estimated stand average or per-hectare value for the *k*th classifier class within the NFI unit:

[59]

$$\bar{y}_k^{\ r} = \frac{\bar{y}_k}{\bar{x}'_k}\bar{x}_k = \frac{\sum_{j=1}^{m} y_{kj}}{\sum_{j=1}^{m} x'_{kj}}\bar{x}_k = \hat{R}'_k \bar{x}_k$$

Note that $\bar{y}_k^{\ r}$ is a product of two ratios, where "r" stands for ratio estimator.

5. Estimate the variance of $\hat{R}'_k$ [Cochran (1977, eq. 6.9-6.13, p. 155)]:

[60]

$$\hat{\text{var}}(\hat{R}'_k) \doteq \frac{1}{(\sum_{j=1}^{m} x'_{kj} / m)^2}\left(\frac{\left[\sum_{j=1}^{m} y_{kj}^{\ 2} + \hat{R}'^{2}_k \sum_{j=1}^{m} x'^{\ 2}_{kj} - 2\hat{R}'_k \sum_{j=1}^{m} y_{kj} x'_{kj}\right]}{m(m-1)}\right)$$

6. Estimate the variance of $\bar{x}_k$ [Cochran (1977, eq. 6.9-6.13, p. 155)]:

[61]

$$\hat{\text{var}}(\bar{x}_k) \doteq \frac{1}{\bar{a}_T^{\ 2}}\left(\frac{\left[\sum_{j=1}^{n} x_{ki}^{\ 2} + \bar{x}_k^{\ 2} \sum_{j=1}^{n} a_{Ti}^{\ 2} - 2\bar{x}_k \sum_{j=1}^{n} x_{ki} a_{Ti}\right]}{n(n-1)}\right)$$

7. Estimate the variance of $\bar{y}_k^{\ r}$ [Yates (1971, eq. 7.5.j, p. 198)[14]]:

[62] $$\hat{\text{var}}(\bar{y}_k^{\ r}) \doteq \bar{x}_k^{\ 2}\ \hat{\text{var}}(\hat{R}'_k) + \hat{R}'^{2}_k\ \hat{\text{var}}(\bar{x}_k) + 2\hat{R}'_k \bar{x}_k\ \hat{\text{cov}}(\hat{R}'_k, \bar{x}_k)$$

where, using the result in Cochran (1977, eq. 6.90, p. 181):

[14] Yates, F. 1971. Sampling methods for censuses and surveys (Third edition; 3rd Impression). Charles Griffin & Company Limited, London, UK.

$$\hat{\text{cov}}(\hat{R}_k', \bar{x}_k) \doteq \frac{1}{m(m-1)(\sum_{j=1}^{m} x_{kj}' / m)(\sum_{j=1}^{m} a_{Tj} / m)} \left[\sum_{j=1}^{m} y_{kj} x_{kj} - \hat{R}_k' \sum_{j=1}^{m} x_{kj} x_{kj}' - \bar{x}_k \sum_{j=1}^{m} y_{kj} a_{Tj} + \hat{R}_k' \bar{x}_k \sum_{j=1}^{m} x_{kj}' a_{Tj} \right]$$

8. Estimate the relative standard error of the estimated average:

[63]
$$SE\%(\bar{y}_k{}^r) = \frac{\sqrt{\hat{\text{var}}(\bar{y}_k{}^r)}}{\bar{y}_k{}^r} 100$$

9. Estimate the ground-based total for the *k*th classifier class in the NFI unit:

[64]
$$\hat{Y}_k{}^r = A\bar{y}_k{}^r$$

10. Estimate the variance of the estimated total:

[65]
$$\hat{\text{var}}(\hat{Y}_k{}^r) = A^2 \ \hat{\text{var}}(\bar{y}_k{}^r)$$

11. Estimate the relative standard error of the estimated total in the NFI unit:

[66]
$$SE\%(\hat{Y}_k{}^r) \doteq \frac{\sqrt{\hat{\text{var}}(\hat{Y}_k{}^r)}}{\hat{Y}_k{}^r} 100$$

12. Calculate the approximate (1-α)100% confidence limits of Y_k:

[67]
$$\hat{Y}_k{}^r \pm t_{\alpha/2, m-1} \sqrt{\hat{\text{var}}(\hat{Y}_k{}^r)}$$

where α is a stated probability level of a Type I error (usually 0.05), and $t_{\alpha/2}$ is a value from the *t*-distribution with *m*-1 degrees of freedom.

5.3 Example 6: Estimating tree volume totals using the LTP and photo plots

An earlier dataset that was created from permanent sample data from the province of Nova Scotia was used to illustrate the volume estimation procedures described in section 5.2 (Table 7). A scattergram of the total photo volume versus ground volume corresponding to the LTPs is depicted in Figure 3.

Table 7. Example dataset of photo and ground plot data for the *k*th classifier class.

NFI Plot No.	*Lcc (k)*	*Area of LTP (ha) (b_j)*	*Area of Photo Plot in k (ha) (a_{kj})*	*Area of Photo Plot (ha) (a_{Tj})*	*Total Volume in LTP (m^3) (y_{kj})*	*Total Photo Volume for LTP (m^3) (x_{kj}')*	*Total Volume in Photo Plot (m^3) (x_{kj})*
985121	NLU	0.0404	0.000	276.885	0.00000	0.00	0.0
992006	NLU	0.0404	0.000	55.759	0.00000	0.00	0.0
985126	VNW	0.0404	30.123	400.001	0.00000	0.00	0.0

NFI Plot No.	*Lcc (k)*	*Area of LTP (ha) (b_j)*	*Area of Photo Plot in k (ha) (a_{kj})*	*Area of Photo Plot (ha) (a_{Tj})*	*Total Volume in LTP (m^3) (y_{kj})*	*Total Photo Volume for LTP (m^3) (x'_{kj})*	*Total Volume in Photo Plot (m^3) (x_{kj})*
985131	VTU	0.0404	236.243	400.001	12.92266	19.19	112129.0
985136	VTU	0.0404	323.435	399.998	15.46358	19.19	187001.3
991996	VTU	0.0404	140.843	400.004	14.42947	12.12	57245.75
992001	VTU	0.0404	374.863	399.997	24.98102	16.16	168305.5
992011	VTU	0.0404	282.381	400.000	13.0126	24.24	139959.9
992016	VTU	0.0404	206.821	400.005	12.13565	19.19	108154.2
998876	VTU	0.0404	257.627	400.001	13.67285	19.19	125008.4
998881	VTU	0.0404	392.461	399.998	13.30072	16.16	167425.3
998886	VTU	0.0404	287.792	399.998	15.28456	19.19	127176.6
998891	VNU	0.0404	152.908	400.004	0.00000	0.00	0.0
998896	VTU	0.0404	279.410	399.992	20.96259	16.16	132637.1
998901	VTU	0.0404	333.293	400.000	12.14065	19.19	156920.3

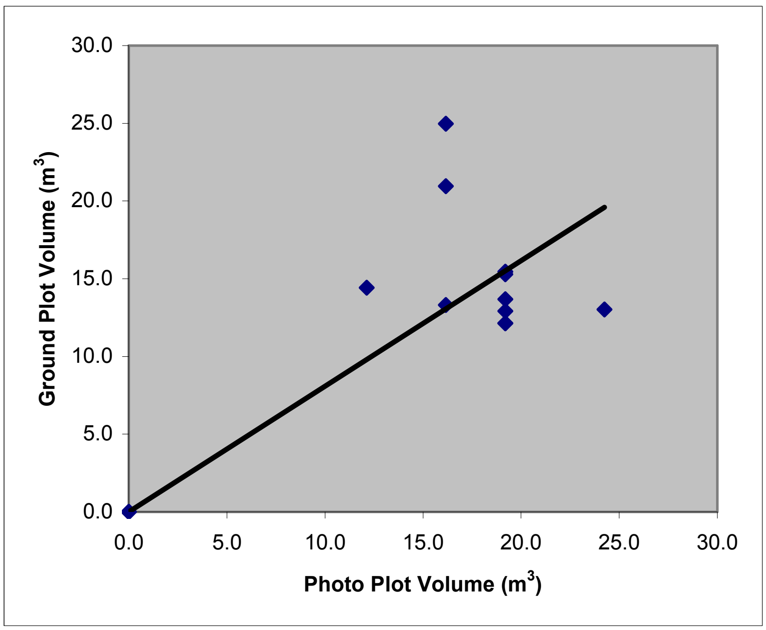

Figure 3. Photo plot volume (x'_{kj}) versus ground plot volume (y_{kj}), with a trend line through the origin.

a) Calculate the following statistics:

Correlation between y_{kj} and x'_{kj} = 0.8104

$$\sum_{j=1}^{m} b_j = 0.6060$$

$$\sum_{j=1}^{m} x_{kj} = 112128.9500 + 187001.2750 + 57245.7500 + ... + 156920.2500 = 1481963.2250$$

$$\sum_{i=1}^{n} a_{Ti} = 276.885 + 55.759 + 400.001 + ... + 400.000 = 5532.6430$$

$$(\sum_{j=1}^{m} a_{Tj}) / m \quad = [276.885+55.759+400.001+\ldots+400.000]/15 = 368.8467$$

$$\bar{a}_T = \frac{5532.643}{15} = 368.84267$$

$$\sum_{j=1}^{m} y_{kj} = 12.922658716 + 15.463579548 + \ldots + 12.140647632 = 168.306349904$$

$$\sum_{j=1}^{m} x'_{kj} = 19.19 + 19.19 + \ldots + 19.19 = 199.98$$

$$\sum_{j=1}^{m} {y_{kj}}^2 = 12.922658716^2 + 15.463579548^2 + \ldots + 12.140647632^2 = 2739.2795056$$

$$\sum_{j=1}^{m} y_{kj} x'_{kj} = (12.922658716 \text{ x } 19.19) + (15.463579548 \times 19.19)\ldots + (12.140647632 \text{ x} 19.19)$$

$$= 3013.9858295$$

$$\sum_{j=1}^{m} {x'_{kj}}^2 = 19.19^2 + 19.19^2 + \ldots + 19.19^2 = 3727.4454$$

$$\bar{x}_k \sum_{j=1}^{m} y_{kj} a_{Tj} = 21039507.421080$$

$$\sum_{j=1}^{m} y_{kj} x_{kj} = 23070412.633059$$

$$\hat{R}'_k \sum_{j=1}^{m} x_{kj} x'_{kj} = 22994453.011716$$

$$\hat{R}'_k \bar{x}_k \sum_{j=1}^{m} x'_{kj} a_{Tj} = 21039538.1094$$

$$\sum_{i=1}^{n} {a_{Ti}}^2 \quad 276.885^2 + 55.759^2 + 400.001^2 + \ldots + 400.000^2 = 2159773.569451$$

$$\sum_{i=1}^{n} {x_{ki}}^2 \quad = 112128.950^2 + 187001.2750^2 + \ldots + 156920.250^2 = 212481082413.6780$$

b) Calculate the estimated volume totals and their associated variances and confidence limits.

[56] $\bar{y}_k = \dfrac{168.30634990400}{0.6060} = 277.7332507$ m^3/ha

[57] $\bar{x}'_k = \dfrac{199.98}{0.6060} = 330.0000$ m^3/ha

[58] $\bar{x}_k = \dfrac{1729055.45}{5532.643} = 312.5188902$ m^3/ha

$\hat{R}'_k = \dfrac{168.30634990400}{199.98} = 0.841615911$

[59] $\bar{y}_k{}^r = 0.841615911 \times 312.5188902 = 263.02087047969$ m^3/ha

[60] $\hat{\mathrm{var}}(\hat{R}'_k) = 0.00820494975$

[61] $\hat{\mathrm{var}}(\bar{x}_k) = 784.45804190741$ (From Example 3, eq. [20])

$\hat{\mathrm{cov}}(\hat{R}'_k, \bar{x}_k) = 0.073587191$

[62] $\hat{\mathrm{var}}(\bar{y}_k{}^r) = 1395.7166$

[63] $SE\%(\bar{y}_k{}^r) = 14.20\%$

[64] $\hat{Y}_k{}^r = 263.02087047969 \times 5549000 = 1459502810.292$ m^3

[65] $\hat{\mathrm{var}}(\hat{Y}_k{}^r) = 42976071552748500$

[66] $SE\%(\hat{Y}_k{}^r) = 14.20\%$

[67] Approximate confidence limits for Y_k (95%, t_{15-1}) = 1014873745.26 m^3 to 1904131875.32 m^3

c) Summarize the volume statistics as follows:

Lcc (k)	$\bar{y}_k^r$ *(m^3/ha)*	$\hat{\mathrm{var}}(\bar{y}_k^r)$	$\hat{Y}_k{}^r$ *(m^3)*	$\hat{\mathrm{var}}(\hat{Y}_k{}^r)$	*Confidence limits of Y_k (m^3)*	
					Lower	*Upper*
VTU	263.0	1395.72	1459502810	42976071552748500	1014873745	1904131875

As before, we can also estimate the classifier-class average. This can be done as follows:

1. Estimate the average "per-hectare" attributes (e.g. volume/ha) within the *k*th classifier-class area in the NFI unit:

[68]
$$\hat{\bar{y}}_k{}^r = \frac{\hat{Y}_k}{\hat{A}_k} = \frac{A\bar{y}_k{}^r}{A\hat{p}_k} = \frac{\bar{y}_k{}^r}{\hat{p}_k} = \frac{263.02087047969}{0.6645428} = 395.7922282 \text{ m}^3\text{/ha}$$

2. To estimate the variance of the estimated average value within the classifier class area, re-write $\hat{\bar{y}}_k{}^r$ as:

[69]
$$\hat{\bar{y}}_k{}^r = \frac{\sum_{j=1}^{m} y_{kj} / \sum_{j=1}^{m} x'_{kj}}{\sum_{i=1}^{n} a_{ki} / \sum_{i=1}^{n} x_{ki}} = \frac{\hat{R}_k{}'}{\hat{R}''_k} = \frac{0.841615911}{0.002126408} = 395.7922282$$

Then, using the result in Cochran (1977, eq. 6.95, p. 184):

[70]
$$\hat{\text{var}}(\hat{\bar{y}}_k{}^r) \doteq (\frac{\hat{R}'_k}{\hat{R}''_k})^2 \left[\frac{\hat{\text{var}}(\hat{R}'_k)}{\hat{R}'^2_k} + \frac{\hat{\text{var}}(\hat{R}''_k)}{\hat{R}''^2_k} - \frac{2\,\hat{\text{cov}}(\hat{R}''_k, \hat{R}'_k)}{\hat{R}''_k \hat{R}'_k} \right]$$

$$= 1604.55878$$

where, using the result in Cochran [(1977, eq. 6.90, p. 181)]:

$$\hat{\text{cov}}(\hat{R}''_k, \hat{R}'_k) \doteq \frac{1}{(m)(m-1)(\sum_{j=1}^{m} x'_{kj}/m)(\sum_{j=1}^{m} x_{kj}/m)} \left[\sum_{j=1}^{m} y_{kj} a_{kj} - \hat{R}'_k \sum_{j=1}^{m} a_{kj} x'_{kj} - \hat{R}''_k \sum_{j=1}^{m} y_{kj} x_{kj} + \hat{R}''_k \hat{R}'_k \sum_{j=1}^{m} x'_{kj} x_{kj} \right]$$

$$= 0.000001858637375$$

and, using the result in Cochran [(1977, eq. 6.9-6.13, p. 155]:

$$\hat{\text{var}}(\hat{R}''_k) \doteq \frac{1}{n(n-1)(\sum_{i=1}^{n} x_{ki}/n)^2} \left[\sum_{i=1}^{n} a_{ki}{}^2 + \hat{R}''^2_k \sum_{i=1}^{n} x_{ki}{}^2 - 2\hat{R}''_k \sum_{i=1}^{n} a_{ki} x_{ki} \right] = 0.00000000332912$$

3. Estimate the relative standard error of the estimated average in the NFI unit:

[71]
$$SE\%(\hat{\bar{y}}_k{}^r) \doteq \frac{\sqrt{\hat{\text{var}}(\hat{\bar{y}}_k{}^r)}}{\hat{\bar{y}}_k{}^r} 100 = 10.12\%$$

4. Calculate the approximate (1-α)100% confidence limits of Y_k / A_k :

[72] $$\hat{\bar{y}}_k^{\,r} \pm t_{\alpha/2,m-1}\sqrt{\hat{\text{var}}(\hat{\bar{y}}_k^{\,r})} = 309.88 \text{ m}^3/\text{ha to } 481.71 \text{ m}^3/\text{ha}$$

where α is a stated probability level of a Type I error (here α = 0.05), and $t_{\alpha/2}$ is a value from the *t*-distribution with *m*-1 (15-1) degrees of freedom.

6. Ecozone and national totals

As indicated in section 1.4, the ecozone and national area and attribute totals are obtained by summing the individual NFI-unit totals and their associated variances. This can be achieved using the stratified sampling formulae [e.g., Cochran (1977, ch. 5)]. The method is identical for both area and attribute totals, and for ecozone and national totals; we present below the approach for ecozone area totals.

1. Estimate the total area A_{kE} (ha) for the *k*th classifier class in the ecozone:

[73] $$\hat{A}_{kE} = \sum_{h=1}^{L} \hat{A}_{kh}$$

where $\hat{A}_{kh}$ is the area estimate for the *k*th classifier class in the *h*th NFI unit as determined in section 3.1, and *L* is the total number of NFI units in ecozone *E*.

2. Estimate the variance of the total area in the *k*th classifier class in the ecozone:

[74] $$\hat{\text{var}}(\hat{A}_{kE}) = \sum_{h=1}^{L} \hat{\text{var}}(\hat{A}_{kh})$$

3. Estimate the relative standard error of the estimated total area for the *k*th classifier class in the ecozone:

[75] $$SE\%(\hat{A}_{kE}) = \frac{\sqrt{\hat{\text{var}}(\hat{A}_{kE})}}{\hat{A}_{kE}} 100$$

4. Calculate the approximate (1-α)100% confidence limits of A_{kE} as:

[76] $$\hat{A}_{kE} \pm t_{\alpha/2}\sqrt{\hat{\text{var}}(\hat{A}_{kE})}$$

where $t_{\alpha/2}$ is assumed to be approximately 2.0, a value from the *t*-distribution with α = 0.05 and degrees of freedom ≥ 30.

7. Discussion

We have outlined estimation procedures based on the NFI design and ratio-of-means estimators for area totals and tree and other attribute totals and their associated variances. These ratio estimators would be appropriate under the assumptions 1 to 4 stated in section 2.2. These assumptions should be checked during implementation of these procedures, e.g, using the methods used in Example 2 (section 3.1.5). If these assumptions are not met, then other more complex methods should be considered for the estimation of the variance of the population totals. Other non-classical variance estimators could also be investigated and compared to the design-based ones we have proposed.

In the approximation of variance estimates for the attribute totals, the single systematic sample of photo plots was assumed to be a simple random sample. Alternative variance estimation methods that take into account the spatial correlation associated with the systematic sample, such as model-based estimation systems, or design-based systems[15], could also be investigated. However, the nominal distance of 20 km between NFI sampling locations casts doubt on the existence of spatial correlation between photo plots within some NFI units, such as in coastal British Columbia.

Zarnoch and Bechtold (2000)[16] discussed the ratio-of-means estimator and other potential alternative estimators, such as the mean-of-ratios and the Horwitz-Thompson estimators. They recommend the ratio-of-means for large-scale inventories. They concluded in their paper that, among their class of estimators, the ratio-of-means "… appears qualified for the best linear unbiased estimator for inventory attributes involving relationships between trees and plot size …". The alternative estimators could be investigated empirically with actual or simulated NFI data.

A sample size of at least 30 photo plots or 30 ground plots has been assumed as a minimum requirement for reliable variance estimation and to minimize potential bias associated with ratio estimators. However, several NFI units or subunits will have very few plots or no plots at all. Small under-sampled NFI units could be pooled prior to analysis. The observations from ground plots must, however, be weighted during the estimation, to account for differences in sample selection probabilities among NFI units or among subunits within an NFI unit. The sampling weights that reflect sampling intensity may be used: $w'_h = \frac{A_h}{m_h}$, where w'_h is the sampling weight attached to each ground plot in the hth NFI unit, m_h is the number of ground plots in the hth NFI unit or sub-unit, and A_h is the total area of the hth NFI unit or sub-unit.

The suggested pooling may, however, be undesirable (e.g., different ecozones) or impossible (e.g., missing data); this creates a "small area estimation" problem [Rao (2003)]. [17] This problem is also likely to occur in user-defined domains within ecozones or across ecozones, since these user-defined domains are likely to be small and contain very few plots (or none at all). Several techniques have been proposed in the statistical literature to address this small area estimation problem (e.g., Rao, 2003); however, research is needed to evaluate and adapt these estimation techniques for the NFI.

[15] D'Orazio, M. 2003. Estimating the variance of the sample mean in two-dimensional systematic sampling. *J. Agric., Biol., and Env. Stat*. 8 (3): 280-295.
[16] Zarnoch, S.J. and W.A. Bechtold. 2000. Estimating mapped-plot forest attributes with ratios of means. *Can. J. For. Res*. 30: 688-697.
[17] Rao, J.N.K. 2003. Small area estimation. John Wiley & Sons, Inc., Hoboken, New Jersey.