4.3 Error Sources and Their Influence on the NFI Inventory Results

Michael Köhl

Despite the efforts to achieve high quality data and to use efficient statistical estimators, the results of the NFI are not free of errors. A complete census of the Swiss forests is impossible due to the necessary costs, the available personnel, and the required time from the beginning of the survey up to and including the publication of the results; thus no other alternative to a sample based approach for the NFI exists. In a sample based survey a small portion (sample) is selected from the entire population of the Swiss forests. The selected elements (trees or sample plots) are precisely assessed. The elements included in the sample are then used to draw an inference about the entire population. Inferring about the entire population (e.g., Swiss forests) based on the sample has its roots in theoretical probability assumptions. Simply spoken, the probabilities with which the individual population elements are selected for the sample are taken into account during the derivation of the statistical parameters (e.g., mean values, proportions, and totals). For the inference, the statistical parameters that were calculated using the sample data are applied to the entire population.

Since only some of the population elements are used for the derivation of the statistical parameters, the derived values for the entire population are not the “true values,” but rather “estimates.” These estimates are subject to errors – the so-called estimation error or sampling error. The reason for these errors is the fact that it is possible to select exactly $\binom{N}{n}$ samples of size n out of a population of size N[1], and that the samples show random variation. Inferring about the entire population based on the sample involves this random component and makes it possible to not only calculate the estimates but also their estimation errors.

The calculation of estimation errors is based on the assumption that an observation of a selected element corresponds to its actual (true) value. Deviations of observed and true value can occur due to measurement errors or the wrong assignment of attributes (e.g., tree species). If the population parameters are derived with the help of functions or models, prediction errors can also influence the reliability of the results. These errors are called non-sampling errors, in contrast to the sampling error. This includes errors during the data collection as well.

Sampling and non-sampling errors can influence the results in two different ways: accuracy and precision. Accuracy refers to the systematic deviation between the estimated value and the true value; precision refers to the size of the deviation of the estimate when the sampling procedure is repeatedly applied to the population (Cochran, 1977). The combined effect of both of these components on the reliability of estimates is illustrated in Figure 1.

An estimate can be precise or imprecise, biased or unbiased. With increasing sample size, the sampling error decreases (i.e., the results become more precise). Biases can occur as a result of measurement errors or model errors, but can also occur because of statistical methods. An example for bias could be an interpreter for a Forest Condition Monitoring Program, who systematically overestimates crown transparency by 10%. Independent of the number of assessed trees, the estimated mean crown transparency will always be 10% too high. The bias is not considered for the calculation of the sampling errors, so that even biased results can suggest very reliable results due to a low sampling error.

The objective of a sample based inventory is to obtain a true representation of the target population. Errors that occur during the different steps of an inventory and can have different sources, lead to a different representation of the reality and influence the reliability of the results. For this reason the non-sampling and sampling errors were thoroughly studied during the preparation for the second NFI (GERTNER and KÖHL 1992; 1995; KÖHL 1991; 1993; 1994; KÖHL and KAUFMANN 1993; KÖHL *et al.* 1994; KÖHL and ZINGG 1996).

[1] This holds for random sampling without replacement.

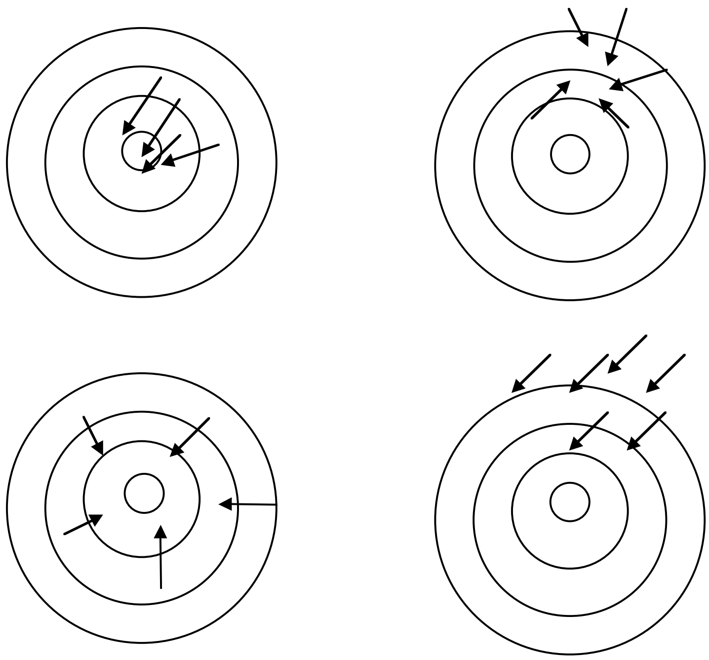

Figure 1. Precision and accuracy of estimators.

An important objective for the method development of the second NFI was to keep the overall error of the estimates as low as possible. Based on the data and check assessments of the first NFI, different potential error components were studied and their proportion of the total error quantified. Methods were finally sought that would reduce the overall error. An important tool for this was the derivation of error budgets, in which several different sources of error could be combined. The analysis of the error budget led, for example, to a revision of the methods for the volume calculation of individual trees (Chapter 3.2). Furthermore, the studies emphasized the necessity of high data quality for the assessment on the forest sample plots and aerial photographs (see also Chapter 2.8 to 2.9 and Chapter 4.4).

The properties of the statistical estimation procedures are especially important in the analysis of the data. In the following, the statistical estimation procedures applied in the second NFI are examined with respect to their bias, efficiency, consistency, and their suitability for employment in the NFI. After that, the example of the volume estimation in the first NFI is used to illustrate the employment of an error budget.

4.3.1 Sampling Error and Bias

A measure to evaluate the quality of an estimator is the mean squared error (MSE, see page 15, COCHRAN 1977).

$$\mathrm{MSE}(\hat{\mu}) = E(\hat{\mu} - \mu)^2 = (\text{variance of } \hat{\mu}) + (\text{bias})^2 \tag{1}$$

where $\hat{\mu}$ is an estimate for the true mean value μ and the bias is the deviation of a calculated mean value m from μ. The MSE consists of two different components which describe the adequacy of an estimate: 1) the precision, given by the variance of $\hat{\mu}$, and 2) the accuracy given by the bias.

The verification of the estimation procedures introduced in Chapter 2.1 was an important step during the development of the sampling designs in the second NFI. Based on a test data set, the MSE of the combined ratio estimators was studied using the Jackknife method. The test data set was compiled from data obtained from the first NFI survey and the National Forest Condition Survey 1990. A total of 723 sample plots were used.

COCHRAN (1977) and SUKHATME et al. (1984) describe Jackknife methods which improve the variance estimation of ratios. The Jackknife estimator was introduced by QUENOUILLE (1956) as a method for bias reduction. TUCKEY (1958) suggested using this method for variance estimates and coined, in unpublished papers, the name of this method (MILLER 1974). MILLER (1974) provides an overview of the Jackknifing.

The basic principle of the Jackknifing is to derive estimates for a sample that is of reduced size. Suppose $X_1,...,X_n$ is a sample of independent and identically distributed random variables. Suppose, furthermore, that $\hat{d}$ is an estimate for the population parameter □ that is derived from the sample of size *n*. If the sample of size *n* is divided into *g* groups of size *h*, so that *n* equals *gh*, it follows that $\hat{d}_{-i}$ is a corresponding estimate, which is based on a sample of size *(g-1)h* in which the i^th^ group was deleted. The estimate $\hat{d}_{Ji}$ is defined as follows:

$$\hat{d}_{Ji} = g\hat{d} - (g-1)\hat{d}_{-i} \tag{2}$$

Tuckey introduced for $\hat{d}_{Ji}$, the term “pseudo value” (MILLER 1974). The Jackknife estimate is:

$$\hat{d}_J = \frac{\sum \hat{d}_{Ji}}{n} \tag{3}$$

with the variance $v(\hat{d}_J)$

$$v\left(\hat{d}_J\right) = \frac{\sum\left(\hat{d}_{Ji} - \hat{d}_J\right)^2}{n(n-1)} \tag{4}$$

The most common form of the Jackknife estimator uses a group size *h=1*, which implies that *n=g*. Thus, the pseudo values are calculated by leaving out the i^{th} observation for the calculations. This is also assumed in the following presentation:

The Jackknife ratio estimator follows from Equation (3) and is:

$$\hat{R}_J = \frac{\sum \hat{R}_{Ji}}{n} \tag{6}$$

with the pseudo values

$$\hat{R}_{Ji} = n\hat{R} - (n-1)\hat{R}_{-i} \tag{7}$$

and the variance

$$v(\hat{R}_J) = \frac{\sum\left(\hat{R}_{Ji} - \hat{R}_J\right)^2}{n(n-1)} \tag{8}$$

For the calculation of $\hat{R}_{Ji}$, a combined ratio estimator with double sampling $\hat{R}_{ds}$ was used.

$$\hat{R}_{ds} = \frac{\hat{Y}_{ds}}{\hat{X}_{ds}} = \frac{\hat{\bar{Y}}_{ds}}{\hat{\bar{X}}_{ds}} \tag{9}$$

In a Monte-Carlo simulation study $\hat{R}_{ds}$ was compared to the Jackknife estimator (KÖHL 1994). Since the bias of the ratio estimator is of order 1/n, which implies that bias should especially be expected for small sample sizes, the sample size was chosen to be n=50 and n=100. The number of sample units within the production regions was selected to be proportional to the size of the production regions. In order to obtain a data set for the selection of the samples that was as extensive as possible, and to avoid empty strata as much as possible, the estimator was not separately derived for the production regions. Due to the enormous computing time, the simulation study was limited to 100 iterations.

Conducting a simulation study with a double sampling ratio estimator requires determining the sample sizes in both the first and second phase. Since the main focus of the study was on the behavior of the estimator for small sample sizes in the second phase, the sample size in the first phase, with 2,500 samples, was chosen to be relatively high, and the simulated second phase sample sizes were chosen to be constant with n=50 and n=100.

The number of stems and timber volume per hectare, which were determined using the double sampling ratio estimator for the test data set with n'=2,500 observations in phase one and n=723 observations in phase two are presented in Table 1.

Table 1. Double sampling ratio estimator for the test data set (n = 723, n' = 2500).

	N/ha [m^3/ha]	V/ha [n/ha]
R	429.07	310.88
s(R)	9.91	7.01

Table 2 contains the summarized results of the simulation study. The results show that the Jackknife estimator (Jackknife $\hat{R}$) and the double sampling ratio estimator $\hat{R}_{ds}$, with a sample size of n=100, lead to practically the same results for the estimation of stem number as well as for the estimation of timber volume. For a sample size of n=50, differences between both estimators can be observed. The Jackknife estimator leads to lower ratios. The difference between the procedures is, however, relatively low and amounts to approximately 1.5% for the estimation of stem number.

Table 2. Results of the Monte-Carlo simulation study (mean of 100 iterations).

	n=50		n=100	
	N/ha [m^3/ha]	V/ha [n/ha]	N/ha [m^3/ha]	V/ha [n/ha]
$\hat{R}_{ds}$	433.36	314.09	431.01	312.07
Jackknife $\hat{R}$	427.14	312.45	432.06	312.76
s($\hat{R}_{ds}$)	35.90	24.75	25.92	17.81
Jackknife s($\hat{R}$)	42.90	29.73	28.43	19.69

Jackknifing results in a higher standard error in comparison to the double sampling ratio estimator. However, for the 100 iterations that were conducted, the Jackknife procedure has a higher variance of the standard error than the double sampling ratio estimator, which can, therefore, be considered to be more stable. The standard error for the double sampling ratio estimator is for a sample of size n=50 approximately 20% lower than for the Jackknife procedure, and for a sample of size n=100 approximately 10% lower as well. The conservative estimation for the standard error by Jackknifing observed here is known and is, among others, described by EFFRON (1982). The mean bias of the 100 iterations (see Table 3) also varies for the double sampling ratio estimator for both procedures in the same range and is between -0.45% and +1.0%.

Table 3. Mean bias of the double sampling ratio estimator (values in parentheses: deviation from the value for the total population in percentage).

	n=50		n=100	
	N/ha [m^3/ha]	V/ha [n/ha]	N/ha [m^3/ha]	V/ha [n/ha]
$\hat{R}_{ds}$	4.29 (1.0%)	2.21 (0.71%)	1.94 (0.45%)	1.19 (0.38%)
Jackknife $\hat{R}$	-1.93 (-0.45%)	1.57 (0.51%)	2.98 (0.69%)	1.88 (0.60%)

The sample size of the simulation study was kept low on purpose and corresponds with samples that each represent 100 hectare to a reference unit of 5,000 ha and 10,000 ha. No results had to be derived in the NFI for such small areas (i.e., the sample size is usually significantly larger for the analysis of subunits). The combined ratio estimator and the Jackknife estimator had similar results in the simulation study presented here. Thus, it can be assumed that the estimation method used in the second NFI does not lead to a systematic error (bias), and that the sampling error is low even for small sample sizes. The method is therefore appropriate for the use in the NFI and can replace the Jackknife estimator which requires considerably higher computing time.

4.3.2 Nonsampling Error and Error Budget

During the preparation for the second NFI GERTNER and KÖHL (1992) studied the influence of non-sampling errors on the reliability of the inventory results. The objective was to identify error sources that have a substantial influence on the reliability of the results and to minimize these error sources by the revision of the inventory manuals and models.

The sensitivity of the first NFI results were investigated with the help of so-called error budgets. A model was constructed initially for each individual error source in order to quantify the error (Figure 2). For this, random and systematic errors had to be distinguished. The construction and application of the error budget is illustrated in the following for volume estimation of spruce from the first NFI. This shows how, based on the data of the first NFI, the necessity to revise the volume prediction of individual trees was recognized (see Chapter 3.2).

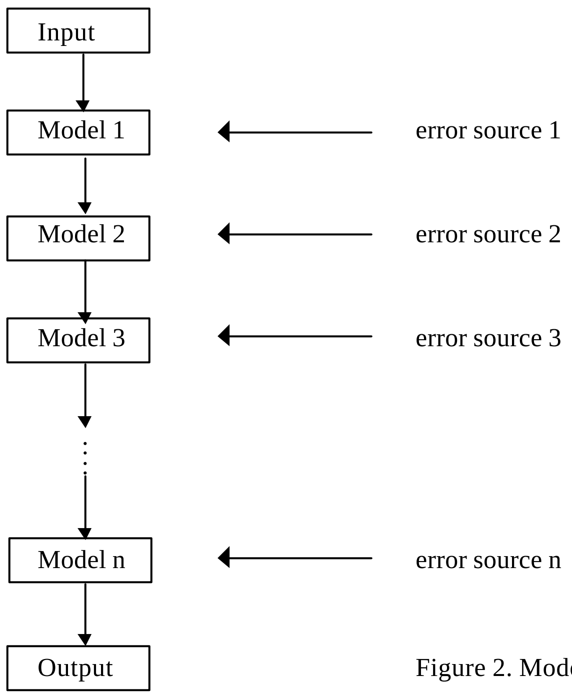

Figure 2. Model used to derive the error budget.

In the first NFI the individual tree volumes were derived using tariff functions. At first, the individual tree volume was determined with the help of volume functions for all trees on which the diameters in height 1.3 meters and 7 meters and total tree height were measured. The individual tree volumes, which were only available for a subsample (the so-called tariff sample trees), were used for the construction of tariff functions. For the tree species spruce, 19 different tariff functions were derived that depended on stand classification attributes, development stage, stand structure, and stand density (KAUFMANN 1991). Independent variables for the calibration of the tariff function included, for example, the DBH, slope of the plots, site index, tree crown position, and the presence of stem forks. For the construction of the error models of the tariff functions the following different error sources were considered:

– The variance of the tariff functions, assuming that the tariff functions are free of error
– Measurement errors of the variables used for the calibration of the tariff function (e.g., measurement errors in the DBH)
– Errors of the discrete variables used for the calibration of the tariff function (e.g., wrong assignment of the crown position)
– Errors that occurred during the assessment of attributes used to select the tariff function (stage of development, stand structure, and stand density)

The errors of these individual sources increase the variance of the volume predicted with the tariff function $var_{tariff}(v)$. Instead of the true volume of individual trees the values predicted by volume functions were used for the construction of the tariff functions. These predicted volumes are smoothed estimates and do not contain the variability of the true individual tree volumes. If the true tree volumes had been used for the derivations of the tariff functions, the variance of the residuals would have been larger than the ones used for the individual volumes estimated with the volume functions. Since the volume functions were developed independently of the NFI survey (HOFFMANN 1984), a general model which was introduced by KISH (1965) for summarizing several different sources of error could be applied and the variance of the volume functions $var_{function}(v)$ could be added to the variance of the tariff function.

The individual tree volumes of the trees that were selected in the NFI, which were estimated using the volume functions, include potentially a bias which depends on systematic errors that arise from measurements of the independent variable in the volume function (e.g., DBH, height). In contrast to the bias of the independent variables within the tariff function, this bias is not adjusted through the calibration. Thus, the bias of the volume function $bias_{function}(v)$ leads to a bias in the tariff function $bias_{tariff}(v)$.

The total error of the individual tree volume that was derived using the tariff functions $var_{tariff}(v)$, for which the measurement errors of the independent variable as well as the errors in the volume function were accounted for is:

$$var_{tariff}(v) \cong var_{function}(v) + var_{tariff\ error}(v) \qquad (10)$$

where $var_{tariff\ errorr}(v)$ the variance of the tariff function and the variance of the different above mentioned errors of the calibration variables, as well as the attributes used to select the tariff function, are all included. The bias of the tariff function is:

$$bias_{tariff}(v) \cong bias_{function}(v) \qquad (11)$$

The individual tree volumes, which were determined by using the tariff functions, are expanded in the NFI to a per hectare basis. The bias $bias_{tariff}(v)$ and the variance $var_{tariff}(v)$ were similarly expanded to a per hectare basis.

For the derivation of the estimates, the individual tree volumes are accumulated for each sample plot to a sample plot volume Y_i (see Chapter 2.1.4). The variance of the sample plot volume $var_{sample}(\overline{Y})$ was calculated in the first NFI using the equations of the random selection (see page 18 and following, COCHRAN 1977). Since the smoothed individual tree volumes from the tariff functions were used, the errors of the tariff function given up to this point are not implicitly accounted for. The variance of the sample plot volume can, however, be traced back to one important source of error. A large portion of the NFI sample plots is situated in inclined terrain and requires correcting for the slope. In the NFI, sample plots on slopes were sampled with concentric circles, which when projected on the horizontal plane were elliptical with a certain defined area (200 m^2 and 500 m^2). The radius of the sample plots on a slope depends on the defined area and the angle of the slope on the plots. An error in the measured angle of the slope will cause an error in actual area of a plot on the horizontal plane and can be either random or systematic. The radius of the sample plots on slopes depends on the defined horizontal plane and the angle of the slope. An error in the measured slope will cause an error of the actual area on the horizontal projection and can be either random or systematic. Thus, the error is described with $var_{slope}(\overline{Y})$ and $bias_{slope}(\overline{Y})$. These errors were accounted for in the construction of the error budget.

The estimated bias, the variance, the mean square error, as well as the percent root mean square error, $bias_{total}(\overline{Y})$, $var_{total}(\overline{Y})$, $mse_{total}(\overline{Y})$ and $prmse_{total}(\overline{Y})$ respectively, are calculated as:

$$bias_{total}(\overline{Y}) \cong bias_{tariff}(\overline{Y}) + bias_{slope}(\overline{Y}) \tag{12}$$

$$var_{total}(\overline{Y}) \cong var_{tariff}(\overline{Y}) + var_{sample}(\overline{Y}) + var_{slope}(\overline{Y}) \tag{13}$$

$$mse_{total}(\overline{Y}) \cong var_{total}(\overline{Y}) + (bias_{total}(\overline{Y}))^2 \tag{14}$$

$$prmse_{total}(\overline{Y}) \cong 100\frac{\sqrt{mse_{total}(\overline{Y})}}{\overline{Y}} \tag{15}$$

Based on the data of the first NFI, it was possible to derive an error budget for the different attributes with this approach. The example of Norway spruce (*Picea abies* (L.) Karst.) is used here to illustrate the error budget for volume estimation. An error budget displays the effects of measurement errors of individual attributes and groups of attributes on the reliability of overall estimates. The amount of measurement errors entered into the error budget was either taken from the analysis of the check assessments of the first NFI (WINZELER 1989) or was determined based on expert opinion. The coefficient of variation due to the measurement errors was 2% for the DBH, 4% for the d_7, 7% for the tree height, 20% for the site index, and 5% for the slope. All independent variables were considered to be unbiased. The classification error, which occurred by selecting the wrong tariff function, was determined using a simulation study (GERTNER and KÖHL 1992).

Table 4 shows the error budget when these measurement and classification errors were used. The mean total stem volume was 267.35 m^3/ha, which amounts to a percent root mean square error of 1.28%. The sampling error is the most important source of error and the tariff functions are the second most important source of error. Random measurement and classification errors were of minor importance. Even if random measurement errors were doubled, in terms of the continuous variables, the coefficient of variation would not change significantly.

Control surveys have shown that the probability of systematic measurement errors in the NFI is very low because every precaution has been taken to avoid systematic errors. In order to understand how sensitive the sampling design of the first NFI was in regard to systematic errors, an error budget was prepared assuming that the measurable variables DBH, d_7, tree height, and slope were biased by one percent. The relatively small bias causes a very drastic increase of the mean squared error (see Table 5). This is particularly true for the d_7. The coefficient of variation increased by about four percent. The main reason for this strong increase is the very large sample size. An increasing sample size reduces the variance due to the sampling error, the random measurement errors, and the prediction errors; however, the increasing sample size does not effect the bias of the volume estimation. Thus, the weight given to the bias increases with increasing sample size.

Error budgets were used to develop the inventory methods of the second NFI in order to demonstrate the influence of different attributes on the reliability of the inventory results. The design of the first NFI is very sensitive to systematic biases. As a consequence, the methods for the volume prediction (i.e., the volume and tariff functions) and the assortment tariffs were completely revised (see Chapter 2.1, Standing Timber, Increments and Utilization).

Table 4. Error budget for spruce in Switzerland (according to Gertner and Köhl, 1992). Measurement error from the check assessment. Values in parentheses are the percentage change of the MSE with respects to the corresponding error source.

	Variance of the stem volume $[m^3/ha]^2$		Bias of the stem volume $[m^3/ha]$
Sampling error		11.5196 (98.416%)	0
Function error			
Volume function	0.0025 (0.0213%)		0
Tariff function	0.1219 (1.1033%)		0
Subtotal		0.1316 (1.1246%)	0
Measurement error			
DBH	0.0000 (0.0000%)		0
Height	0.0017 (0.0145%)		0
d_7	0.0276 (0.2361%)		0
Site class	0.0003 (0.0028%)		0
Slope (%)	0.0008 (0.0066%)		0
Subtotal		0.0304 (0.2600%)	0
Assignment error			
Crown class	0.0090 (0.0768%)		0
Tariff function	0.0143 (0.1225%)		0
Subtotal		0.0233 (0.1993%)	0
Total		11.704 (100%)	0

Table 5. Error budget for spruce in Switzerland with bias of the measurable attribute 1 (according to Gertner and Köhl, 1992). Measurement error from the check assessment. Values in parentheses are the percentage change of the MSE with respects to the corresponding error source.

	Variance of the stem volume $[m^3/ha]^2$		Bias of the stem volume $[m^3/ha]$	
Sampling error		11.5196 (9.427%)		0.00 (0.000%)
Function error				
Volume function	0.0025 (0.0020%)		0.00 (0.000%)	
Tariff function	0.1219 (0.1057%)		0.00 (0.000%)	
Subtotal		0.1316 (0.1077%)		0.00 (0.000%)
Measurement error				
DBH	0.0000 (0.0000%)		2.72 (40.796%)	
Height	0.0017 (0.0014%)		1.75 (27.616%)	
d_7	0.0276 (0.0233%)		5.53 (70.120%)	
Site class	0.0003 (0.0003%)		0.00 (0.000%)	
Slope (%)	0.0008 (0.0006%)		0.50 (8.484%)	
Subtotal		0.0304 (0.0256%)		10.51 (147.016%)
Assignment error				
Crown class	0.0090 (0.0074%)		0.00 (0.000%)	
Tariff function	0.0143 (0.0117%)		0.00 (0.000%)	
Subtotal		0.0233 (0.0191%)		0.00 (0.000%)
Total		11.7049		10.51

4.3.3 Literature

COCHRAN, W.G. 1977. Sampling Techniques. New York: J. Wiley & Sons. 428 p.

EFFRON, B. 1982: The jackknife, the bootstrap and other resampling plans: (ten lectures given at a regional conference on resampling methods; Bowling Green State University, June 1980). CBMS-NSF Reg. Conf. Ser. Appl. Math. 38:92 p.

GERTNER, G.; KÖHL, M. 1992: An Assessment of Some Nonsampling Errors in a National Survey Using an Error Budget. For. Sci. 38 (3):525-538.

GERTNER, G.Z.; KÖHL, M. 1995: Correlated Observer Errors and their Effects on Survey Estimates of Needle-Leaf Loss. For. Sci. 41 (4):758-776.

HOFFMANN, CH. 1984. Volumenfunktionen: WSL. internal report.

KAUFMANN, E. 1991. Volumenermittlung im zweiten Schweizerischen Landesforstinventar. Birmensdorf: Eidgenöss. Forsch.anst. Wald Schnee Landsch. interner Forschungsplan.

KISH, L. 1965. Survey Sampling. New York: J. Wiley & Sons. 643 p.

KÖHL, M. 1991: Waldschadensinventuren: mögliche Ursachen der Variation zwischen Beobachtern bei der Nadel-/Blattverlustschätzung und Folgerungen für Kontrollaufnahmen. Allg. Forst- Jagdztg. 162 (11/12):210-221.

KÖHL, M. 1993: Quantifizierung der Beobachterfehler bei der Nadel-/Blattverlustschätzung. Allgem. Forst- u. Jagdzeitung 164 (5):83-92.

KÖHL, M. 1994: Statistisches Design für das zweite Schweizerische Landesforstinventar: Ein Folgeinventurkonzept unter Verwendung von Luftbildern und terrestrischen Aufnahmen. Mitt. Eidgenöss. Forsch.anst. Wald Schnee Landsch. 69 (1):141.

KÖHL, M.; KAUFMANN, E. 1993: Zur Berechnung der Stichprobenfehler bei Waldschadensinventuren. Schweiz. Z. Forstwes. 144 (4):297-311.

KÖHL, M.; KAUFMANN, E.; INNES, J.L. 1994: Reliability of Differing Densities of Sample Grids Used for the Monitoring of Forest Condition in Europe, Environmental Monitoring and Assessment. 29:201-220.

KÖHL, M.; ZINGG, A. 1996: Eignung von Diversitätsindizes bei Langzeituntersuchungen zur Biodiversität in Waldbeständen. Allgem. Forst- u. Jagdzeitung 167 (4):76-85.

MILLER, R.G. 1974: The jackknife - a review. Biometrika 61:1-15.

QUENOUILLE, M.H. 1956: Notes on bias in estimation. Biometrika 43:353-360.

SUKHATME, P.V.; SUKHATME, B.V.; SUKHATME, S.; ASOK, C. 1984. Sampling Theory of Surveys with Applications. Ames, Iowa: Iowa State University Press. 526 p.

TUCKEY, J.W. 1958: Bias and confidence in not-quite large samples. Ann. Math. Stat. 29:614-622.

WINZELER, C. 1989. LFI-Kontrollaufnahmen. Birmensdorf, Eidgenössische Forschungsanstalt für Wald, Schnee und Landschaft.