5.3 SAMPLING

5.3.1 Introduction

Data for the LULUCF sector are often obtained from sample surveys and typically are used for estimating changes in land use or in carbon stocks. National forest inventories are important examples of the type of surveys used. This section provides *good practice guidance* for the use of data from sample surveys for the reporting of emissions and removals of greenhouse gases, and for the planning of sample surveys in order to acquire data for this purpose. Sampling also is important for monitoring Kyoto Protocol projects, and Chapter 4 provides specific recommendations consistent with this section. This section provides *good practice guidance* concerning:

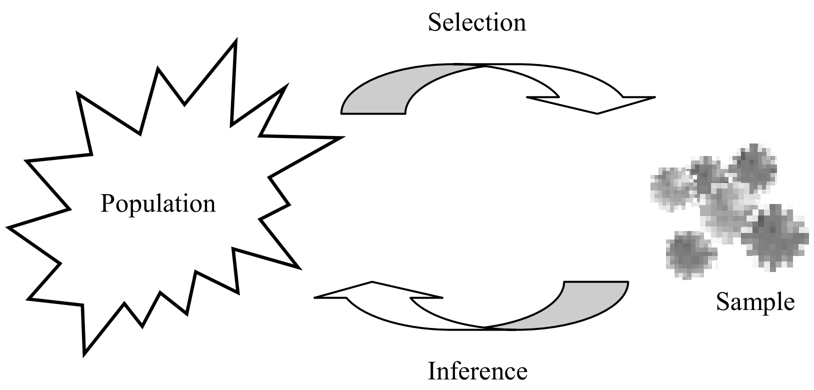
- Overview on sampling principles (Section 5.3.2);
- Sampling design (Section 5.3.3);
- Sampling methods for area estimation (Section 5.3.4);
- Sampling methods for estimating emissions and removals of greenhouse gases (Section 5.3.5);
- Uncertainties in sample based surveys (Section 5.3.6).

Useful general references on sampling include: Raj (1968), Cochran (1977), De Vries (1986), Thompson (1992), Särndal *et al.* (1992), Schreuder *et al.* (1993), Reed and Mroz (1997), and Lund (1998).

5.3.2 Overview on Sampling Principles

Sampling infers information about an entire population by observing a fraction of it: the sample (see Figure 5.3.1). For example, changes of carbon in tree biomass at regional or national levels can be estimated from the growth, mortality and cuttings of trees on a limited number of sample plots. Sampling theory then provides the means for scaling up the information from the sample plots to the selected geographical level. Properly designed sampling can greatly increase efficiency in the use of inventory resources. Furthermore, field sampling is generally needed in developing LULUCF inventories because, even if remote sensing data provide complete territorial coverage, there will be a need for ground-based data from sample sites for interpretation and verification.

Figure 5.3.1 Principle of sampling



Standard sampling theory relies on random selection of a sample from the population; each unit in the population has a specific probability of being included in the sample. This is the case when sample plots have been distributed entirely at random within an area, or when plots have been distributed in a systematic grid system as long as the positioning of the grid is random. Random sampling reduces the risk of bias and allows for an objective assessment of the uncertainty of the estimates. Therefore, randomly sampled data generally should be used where available, or when setting up new surveys.

Samples may also be taken at subjectively chosen locations, which are assumed to be representative for the population. This is called subjective (or purposive) sampling and data from such surveys are often used in greenhouse gas inventories (i.e., when observations from survey sites that were not selected randomly are used to represent an entire land category or subdivision). Under these conditions, observations about, for example, forest type might be extrapolated to areas for which they are not representative. However, due to limited resources greenhouse gas inventories may need to make use of data also from subjectively selected sites or research plots. In this case, it is *good practice* to identify, in consultation with the agencies responsible for the sites or plots, the land areas for which the subjective samples can be regarded as representative.

5.3.3 Sampling Design

Sampling design determines how the sampling units (the sites or plots) are selected from the population and thus what statistical estimation procedures should be applied to make inferences from the sample. Random sampling designs can be divided into two main groups, depending on whether or not the population is *stratified* (i.e., subdivided before sampling) using auxiliary information. Stratified surveys will generally be more efficient in terms of what accuracy can be achieved at a certain cost. On the other hand, they tend to be slightly more complex, which increases the risk of non-sampling errors due to incorrect use of the collected data. Sampling designs should aim for a good compromise between simplicity and efficiency, and this can be promoted by following three aspects of *good practice* as set out below:

- Use of auxiliary data and stratification;
- Systematic sampling;
- Permanent sample plots and time series data.

5.3.3.1 USE OF AUXILIARY DATA AND STRATIFICATION

One of the most important sampling designs which incorporate auxiliary information is *stratification*, whereby the population is divided into subpopulations on the basis of *auxiliary data*. These data may consist of knowledge of legal, administrative boundaries or boundaries of forest administrations which will be efficient to sample separately, or maps or remote sensing data distinguishing between upland and lowland areas or between different ecosystem types. Since stratification is intended to increase efficiency, it is *good practice* to use auxiliary data when such data are available or can be made available at low additional cost.

Stratification increases efficiency in two main ways: (i) by improving the accuracy of the estimate for the entire population; and (ii) by ensuring that adequate results are obtained for certain subpopulations, e.g., for certain administrative regions.

On the first issue, stratification increases sampling efficiency if a subdivision of the population is made so that the variability between units within a stratum is reduced as compared to the variability within the entire population. For example, a country may be divided into a lowland region (with certain features of the land-use categories of interest) and an upland region (with different features of the corresponding categories). If each stratum is homogeneous a precise overall estimate can be obtained using only a limited sample from each stratum. The second issue is important for purposes of providing results at a specific degree of accuracy for all administrative regions of interest, but also in case sampled data are to be used together with other existing datasets, which have been collected using different protocols with the same administrative or legal boundaries.

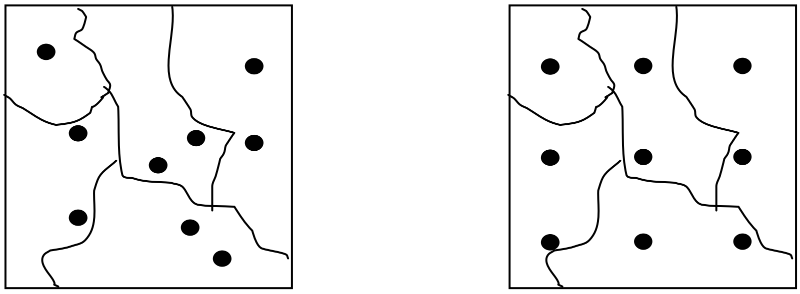
Use of remote sensing or map data for identifying the boundaries of the strata (the land-use class subdivisions to be included in a sample survey) can introduce errors where some areas may be incorrectly classified as belonging to the stratum whilst other areas that do belong to the specific class are missed. Errors of this kind can lead to substantial bias in the final estimates, since the area identified for sampling will then not correspond to the target population. Whenever there is an obvious risk that errors of this kind may occur, it is *good practice* to make an assessment of the potential impact of such errors using ground truth data.

When data for the reporting of greenhouse gas emissions or removals are taken from existing large-scale inventories, such as national forest inventories, it is convenient to apply the standard estimation procedures of that inventory, as long as they are based on sound statistical principles. In addition, *post-stratification* (i.e., defining strata based on remote sensing or map auxiliary data after the field survey has been conducted) means that it may be possible to use new auxiliary data to increase efficiency without changing the basic field design (Dees *et al.* 1998). Using this estimation principle, the risk for bias pointed out in the previous paragraph also can be avoided.

5.3.3.2 SYSTEMATIC SAMPLING

Sample based forest or land-use surveys generally make use of sample points or plots on which the characteristics of interest are recorded. One important issue here regards the layout of these points or plots. It is often appropriate to allocate the plots in small clusters in order to minimise travel costs when covering large areas with a sample based survey. With cluster sampling, the distance between plots should be large enough to avoid major between-plot correlation, taking (for forest sampling) stand size into account. An important issue is whether plots (or clusters of plots) should be laid out entirely at random or systematically using a regular grid, which is randomly located over the area of interest (see Figure 5.3.2). In general, it is efficient to use systematic sampling, since in most cases this will increase the precision of the estimates. Systematic sampling also simplifies the fieldwork.

Figure 5.3.2 Simple random layout of plots (left) and systematic layout (right)



Somewhat simplified, the reason why systematic random sampling generally is superior to simple random sampling is that sample plots will be distributed evenly to all parts of the target area.[3] With simple random sampling, some parts of an area may have many plots while other parts will not have any plots at all.

5.3.3.3 PERMANENT SAMPLE PLOTS AND TIME SERIES DATA

Greenhouse gas inventories must assess both current state and changes over time (e.g., in areas of land-use types and carbon stocks). Assessment of changes is most important and it involves repeated sampling over time. The time interval between measurements should be determined based on the frequency of the events that cause changes, and also on the reporting requirements. Generally, sampling intervals of 5-10 years are adequate in the LULUCF sector, and in many countries data from well designed surveys are already available for many decades, especially in the forest sector. Nevertheless, since estimates for the reporting are required on an annual basis, interpolation and extrapolation methods of the kind described in Section 5.6 will need to be applied. Where sufficiently long time series are not available, it may be necessary to extrapolate backwards in time to capture the dynamics of carbon stock changes, using the *good practice guidance* in Section 5.6 in conjunction with *good practice guidance* in Chapters 3 and 4 about the periods required and assumptions to be made.

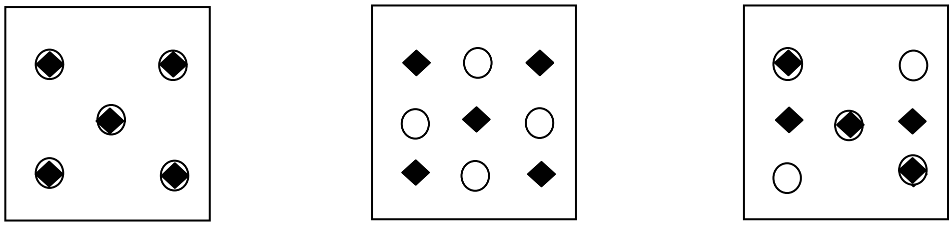
When undertaking repeated sampling, the required data regarding the current state of areas or carbon stocks are assessed on each occasion. Changes are then estimated by calculating the difference between the state at time $t + 1$ from the state at time t. Three common sampling designs can be used for change estimation:

- The same sampling units are used on both occasions (permanent sampling units);
- Different, independent sets of sampling units are used on both occasions (temporary sampling units);
- Some sampling units can be replaced between occasions while others remain the same (sampling with partial replacement).

Figure 5.3.3 shows these three approaches.

[3] In unusual cases when there is a regular pattern in the terrain that may coincide with the systematic grid system, systematic sampling may lead to less precise estimates than simple random sampling. However, such potential problems generally can be handled by orienting the grid system in another direction.

Figure 5.3.3 Use of different configurations of permanent and temporary sampling units for estimating changes



Identical set (permanent plots) | Independent sets (temporary plots) | Sampling with partial replacement (permanent and temporary plots)

○ Sampling unit measured at occasion 1

◆ Sampling unit measured at occasion 2

Permanent sample plots generally are more efficient in estimating changes than temporary plots because it is easier to distinguish actual trends from differences that are only due to changed plot selection. However, there are also some risks in the use of permanent sample plots. If the locations of permanent sample plots become known to land managers (e.g., by visibly marking the plots), there is a risk that management of the permanent plots will differ from the management of other areas. If this occurs, the plots will no longer be representative and there is an obvious risk that the results will be biased. If it is perceived that there might be a risk of the above kind, it is *good practice* to assess some temporary plots as a control sample in order to determine if the conditions on these plots deviate from the conditions on the permanent plots.

The use of sampling with partial replacement can address some of the potential problems with relying on permanent plots, because it is possible to replace sites that are believed to have been treated differently. Sampling with partial replacement may be used, although the estimation procedures are complicated (Scott and Köhl 1994; Köhl *et al.* 1995).

When only temporary plots are used, overall changes still can be estimated but it will no longer be possible to study land-use transfers between different classes unless a time dimension can be introduced into the sample. This can be done by drawing on auxiliary data, for example maps, remote sensing or administrative records about the state of land in the past. This will introduce additional uncertainty into the assessment which it may be difficult to quantify other than by expert judgement.

5.3.4 Sampling Methods for Area Estimation

Chapter 2 presents different approaches for assessing areas or changes in areas of land-use classes. Many of these approaches rely on sampling. Areas and changes in areas can be estimated in two different ways using sampling:

- Estimation via proportions;
- Direct estimation of area.

The first approach requires that the total area of the survey region is known, and that the sample survey provides only the proportions of different land-use classes. The second approach does not require the total area to be known.

Both approaches require assessment of a given number of sampling units located in the inventory area. Selection of sampling units may be performed using simple random sampling or systematic sampling (see Figure 5.3.2). Systematic sampling generally improves the precision of the area estimates, especially when the different land-use classes occur in large patches. Stratification, which is discussed in Section 5.3.3.1, also may be applied to improve the efficiency of the area estimates; in this case it is *good practice* to perform the procedures described below independently in each stratum.

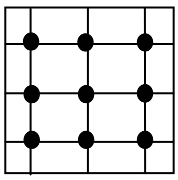
In estimating proportions it is assumed that the sampling units are dimensionless points, although a small area around each point must be considered when the land-use class is determined. Sample plots may also be used for area estimation, although this principle is not further elaborated here.

5.3.4.1 ESTIMATION OF AREAS VIA PROPORTIONS

The total area of an inventory region is generally known. In this case the estimation of the areas of different land-use classes can be based on assessments of area proportions. When applying this approach, the inventory area is covered by a certain number of sample points, and land-use is determined for each point. The proportion of each land-use class then is calculated by dividing the number of points located in the specific class by the total number of points. Area estimates for each land use class are obtained by multiplying the proportion of each class by the total area.

Table 5.3.1 provides an example of this procedure. The standard error of an area estimate is obtained as $A\sqrt{(p_i \bullet (1-p_i))/(n-1)}$, where p_i is the proportion of points in the particular land-use class, A the known total area, and *n* the total number of sample points.[4] The 95% confidence interval for A_i, the estimated area of land use class *i*, will be given approximately by ±2 times the standard error.

TABLE 5.3.1
EXAMPLE OF AREA ESTIMATION VIA PROPORTIONS

Sampling procedure	Estimation of proportions	Estimated areas of land use classes	Standard error
	$p_i = n_i / n$	$A_i = p_i \cdot A$	$s(A_i)$
	$p_1 = 3/9 \cong 0.333$	$A_1 = 300$ ha	$s(A_1)$= 150.0 ha
	$p_2 = 2/9 \cong 0.222$	$A_2 = 200$ ha	$s(A_2)$= 132.2 ha
	$p_3 = 4/9 \cong 0.444$	$A_3 = 400$ ha	$s(A_3)$= 158.1 ha
	Sum = 1.0	Total = 900 ha	

Where:

A = total area (= 900 ha in the example)

A_i = estimated area of land use class *i*

n_i = number of points located in land-use class *i*

n = total number of points

Estimates of areas involved in land-use change can be made by introducing classes of the type A_{ij} where land use changes from class *i* to class *j* between successive surveys.

5.3.4.2 DIRECT ESTIMATION OF AREA

Whenever the total inventory area is known, it is efficient to estimate areas, and area changes, via assessment of proportions, since that procedure will result in the highest accuracy. In cases where the total inventory area is not known or is subject to unacceptable uncertainty, an alternative procedure that involves a direct assessment of areas of different land-use classes can be applied. This approach can only be used when systematic sampling is applied; each sample point will represent an area corresponding to the size of the grid cell of the sample layout.

For example, when sample points are selected from a square systematic grid with 1000 metres distance between the points, each sample point will represent an area of 1km ● 1km = 100 ha. Thus, if 15 plots fall within a specific land-use class of interest the area estimate will be 15 ● 100 ha = 1500 ha.

5.3.5 Sampling Methods for Estimating Greenhouse Gases Emissions and Removals

Sampling is needed not only for area estimation, but also for estimating the state of carbon stocks and emissions and removals of greenhouse gases. As a basis for this, assessment of variables such as tree biomass and soil

[4] Note that this formula is only approximate when systematic sampling is applied.

carbon content is made on the plots. Measurements of these quantities can be made directly on site, or by laboratory analysis of samples, or deduced using models based on correlated variables (such as standard measurements of tree height and diameter) to obtain actual stock, or emissions and removals, of greenhouse gases at the plot level.

Only general guidelines can be given regarding the use of sampling for direct estimation of greenhouse gas emissions or removals. Compared to traditional forest or land-use inventories, the assessments on the plots tend to be slightly more complicated, particularly for the soil carbon pool. An important issue in random sampling surveys is the layout of plots e.g., tree measurements or soil sampling. It is important that this layout is conducted according to strict procedures rather than leaving it to the surveyors to choose appropriate spots for measurements or selecting samples.

Often, inventories of greenhouse gases will be incorporated into on-going national forest or land-use monitoring programmes. In this case it is generally *good practice* to use the established procedures of those inventories, both for purposes of estimating the quantities of interest and the corresponding uncertainties. However, the effects of model conversion errors in final conversion steps (e.g., when applying biomass expansion factors) in this case need to be taken into account. This is further discussed in the next section.

5.3.6 Uncertainties in Sample Based Surveys

The methods described in Chapters 3 and 4 are linked with default uncertainty ranges for the default values presented, and Section 5.2 of this chapter describes how to combine uncertainties in order to estimate the overall uncertainty of an inventory. If an inventory agency uses default values, they can refer to the uncertainty ranges provided in Chapters 3 and 4. When implementing higher tier methods, however, the inventory agency often will use country-specific values and data obtained through research, literature review, field sampling, or remote sensing. Where country-specific data are used, inventory agencies need to develop their own uncertainty estimates, based on expert judgment or – if sampling has been used – based on direct assessment of the precision of the derived data or estimates.

The possibility to derive uncertainty estimates based on formal statistical procedures is a very important advantage of applying sampling procedures in comparison to other methods; the reliability of the information can be assessed based on the data acquired.

Thus, when data from random sampling are used for purposes of greenhouse gas inventory reporting, it is *good practice* to base the assessment of uncertainties on sampling principles, rather than using default values or expert judgement. These uncertainties can then be combined with the uncertainties of other data or models used according to the guidance in Section 5.2 of this chapter.

This section describes the different sources of errors in sample surveys and their effects on overall uncertainty in estimates. *Good practice guidance* is given on how to assess uncertainties in sample based surveys. The discussion on causes of errors is general, and is valid also when data are derived using non-random sampling schemes (e.g., data from research plots) and then scaled up on the basis of area estimates to obtain results on national level. The discussion of the sources of errors first describes errors in assessments at the sample unit level, and then discusses issues in scaling up to estimates for some larger area.

5.3.6.1 TYPES OF ERRORS

Typically for LULUCF inventories, sampling data are acquired from sample plots in the field. To obtain estimates for some larger area (e.g., a country), measurements made at the plot level need to be scaled up. Several kinds of errors may occur in these steps:

- First, whenever measurements are carried out measurement errors due to various imperfections in technique or instrumentation often occur. Measurement errors often are systematic, always deviating in a certain direction from the true value. Such errors then will be propagated during the process of scaling up. Measurement errors also may be random. In this case the average error is zero and the deviations are just as likely to be positive as negative. The latter kinds of errors are less harmful than the systematic ones, although they may lead to systematic errors when basic measurements are applied in models for deriving the quantity of interest (e.g., the volume of a tree).
- Second, the quantities of interest are not always measured directly, but models are applied to derive them. For example, the amount of carbon in a tree usually is calculated by first deriving the tree volume based on models that use parameters such as tree species, diameter, and height as input variables, and then using other models or static expansion factors to convert volume to biomass and biomass to carbon. When applying

models, *model errors* will occur since models seldom are able to predict target quantities exactly. Model errors may be both random and systematic. The sizes are likely to vary depending on the values of the input variables. As shown by Gertner and Köhl (1992), systematic model errors sometimes contribute significantly to the overall uncertainty.

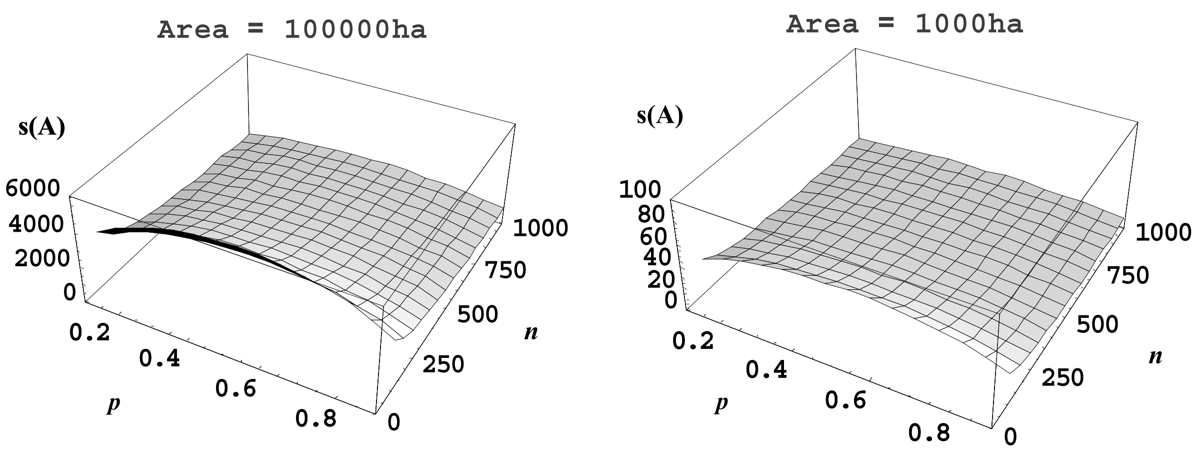
- When plot level measurements are scaled up to some larger area, *sampling errors* occur due to the fact that conditions across the larger area vary and measurements have only been made at the sample locations. The average conditions within the selected sample plots seldom coincide exactly with the average conditions within the entire area of interest. Sampling errors (using random sampling designs and unbiased estimators) are only random, and these effects can be reduced by increasing the sample size, as discussed below and shown in Figure 5.3.4.
- If upscaling is based on complete cover information (e.g., from remote sensing) rather than a sample based survey, uncertainty will be introduced due to land areas being incorrectly classified. Classification errors can be identified and corrected if a sample survey is conducted for studying the extent of such errors. In this case, surveys should be based on random sampling in order to avoid the likely systematic errors of a subjectively selected sample.
- Data registration and calculation errors are the final types of error that may occur. These errors are less technical yet potentially important sources of uncertainty in connection with sample-based surveys. Data registry should be made directly to field computers or different people should independently register data from field forms to computer media in order to avoid registration errors. Calculations need to be checked according to the basic principles of Quality Assurance in Section 5.5. The effects of registration and calculation errors are difficult to assess. Often they are detected and can be corrected for when they cause major deviations from plausible values. When they only cause minor deviations, they are likely to remain undetected.

5.3.6.2 SAMPLE SIZE AND SAMPLING ERROR

The relation between sampling errors, population variance, and sample size is commonly understood; increasing sample size results in higher precision and heterogeneous populations (i.e., those with large within population variation) require larger sample sizes to reach a certain precision. Where area proportions are to be estimated, sampling errors do not only depend on sample size but on the proportion itself. For a given sample size, the sampling error is largest for land-use class proportions $p = 0.5$; it decreases for p approaching 0 or 1.

The effect of different land-use class proportions (from $p = 0.1$ to $p = 0.9$) and sample sizes (from $n = 100$ to $n = 1,000$) on the sampling error of the area estimate is shown in Figure 5.3.4 for two different area sizes (1,000 ha and 100,000 ha).

Figure 5.3.4 Relationship between the standard error of the area estimate s(A), the proportion of the land-use class *p*, and the sample size *n*



5.3.6.3 QUANTIFYING ERRORS IN SAMPLE BASED SURVEYS

In basic sampling theory, the quantities connected to the population units are assumed to be observed without errors. Moreover, the variables of interest (e.g., removals of greenhouse gases) are assumed to be directly recorded at the sampling units; thus no errors due to model conversions need to be considered. In this case, provided adequate statistical estimators have been used, the sample-based estimates of totals (e.g., removals of greenhouse gases at the national level) are unbiased and the corresponding precision can be assessed based on the data acquired.

In many cases (e.g., sampling for area estimation) the above assumptions can be considered valid, and then it is *good practice* to assess the uncertainty of the estimates strictly according to the principles of sampling theory, taking into account what sampling design and estimator were used. The details of such calculations are provided in sampling textbooks such as the references that are introduced in Section 5.3.1. Model errors may enter into the overall uncertainty estimates in different ways. One important case is when the models only give rise to random errors at the level of individual sampling units (e.g., if biomass models have been applied to plot-level tree data). In such cases, the random model errors will inflate the between-plot variability, which will lead to an increased uncertainty of the overall estimates. In this case the standard methods of estimating uncertainties according to sampling theory still can be used, with good approximation, without modifications. Thus, under these conditions it is *good practice* to apply standard sampling theory for deriving the uncertainty estimates, rather than the approaches of Section 5.2.

When models are likely to give rise to (unknown) systematic errors or when they have been used only at some final conversion step (like biomass expansion factors applied to estimates of total volume) the uncertainties introduced should be accounted for. In this case it is *good practice* to use the Tier 1 – or Tier 2 – approach of Section 5.2 for deriving overall uncertainty.

In general, it is *good practice* to assess the applicability of core models for the target population through pilot studies. When models are applied on datasets representing conditions and measurement procedures far different from the ones they were derived upon, there is an obvious risk that the models will incur systematic errors.

Measurement errors can lead to substantial systematic errors, especially in case changes are estimated based on repeated measurements and the systematic error levels vary over time. The size of measurement errors can only be estimated by careful control measurements – on a subsample of the plots – although such check assessments are in some cases difficult to implement (e.g., in soil surveys). In case greenhouse gas inventory reporting is based on sampling, it is *good practice* to conduct careful check assessments on a (small) fraction of the plots, in order to assess the size of the measurement errors. This fraction may be in the order of 1% to 10% depending on the actual sample size and the cost of the control survey, as well as the level of training and experience of the surveyors.

For some variables it is possible to obtain true measurement values through very accurate control procedures, and in such cases the goal should be to estimate the size of the systematic measurement errors. In other cases it may be impossible to measure/assess a true value, and in such cases only the variability between surveyors should be reported.

If major measurement errors are found in a carefully conducted control survey, it is *good practice* to correct for these errors before the final estimates of greenhouse gas emissions/removals are calculated.