CHAPTER 3

UNCERTAINTIES

Authors

Christopher Frey (USA), Jim Penman (UK)

Lisa Hanle(USA), Suvi Monni (Finland), and Stephen Ogle (USA)

Contents

Equations

Figures

Tables

Boxes

3 UNCERTAINTIES

3.1 INTRODUCTION

This chapter provides guidance in estimating and reporting uncertainties associated with both annual estimates of emissions and removals, and emission and removal trends over time. It is written from the viewpoint of the inventory compiler and provides, with examples, two approaches for combining category uncertainties into uncertainty estimates for total national net emissions and the trend.

3.1.1 Overview of uncertainty analysis

Uncertainty estimates are an essential element of a complete inventory of greenhouse gas emissions and removals. They should be derived for both the national level and the trend estimate, as well as for the component parts such as emission factors, activity data and other estimation parameters for each category. This guidance therefore develops a structured approach to estimating inventory uncertainty. It includes methods for:

- Determining uncertainties in individual variables used in the inventory (e.g., estimates of emissions from specific categories, emission factors, activity data);
- Aggregating the component uncertainties to the total inventory;
- Determining the uncertainty in the trend; and
- Identifying significant sources of uncertainty in the inventory to help prioritise data collection and efforts to improve the inventory.

While the methods outlined below are intended to estimate uncertainties for the national inventory, it is important to recognize that some uncertainties that are not addressed by statistical means may exist, including those arising from omissions or double counting, or other conceptual errors, or from incomplete understanding of the processes that may lead to inaccuracies in estimates developed from models.

An uncertainty analysis should be seen, first and foremost, as a means to help prioritise national efforts to reduce the uncertainty of inventories in the future, and guide decisions on methodological choice. For this reason, the methods used to attribute uncertainty values must be practical, scientifically defensible, robust enough to be applicable to a range of categories of emissions by source and removals by sinks, methods and national circumstances, and presented in ways comprehensible to inventory users. A reference section is provided for more detailed and more theoretical information on topics discussed in this chapter.

Quantitative uncertainty analysis is performed by estimating the 95 percent confidence interval of the emissions and removals estimates for individual categories and for the total inventory. The definition of the 95 percent confidence interval is given in Section 3.1.3, Key Concepts and Terminology.

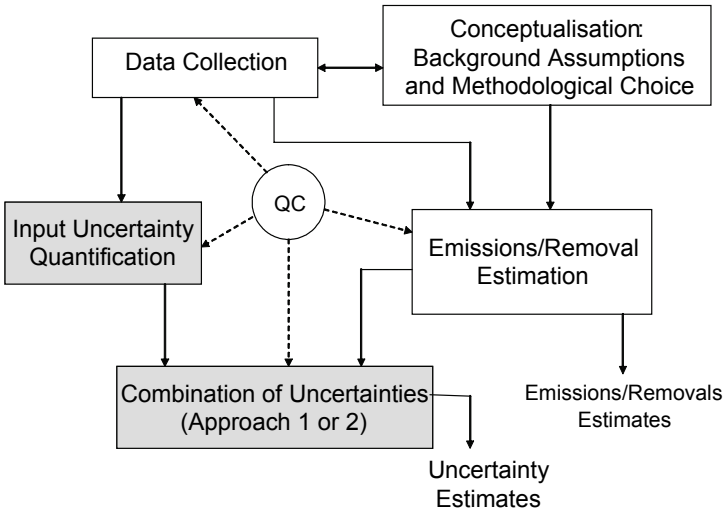
3.1.2 Overall structure of uncertainty analysis

This section provides a brief overview of the overall structure of uncertainty analysis, as illustrated in Figure 3.1. Emissions/removals estimates are based on: (1) conceptualisation; (2) models; and (3) input data and assumptions (e.g., emission factor and activity data). Each of these three can be a source of uncertainty. The analysis begins with a *conceptualisation.* This is a set of assumptions regarding the structure of an inventory or of a sector. These assumptions typically include the scope of geographic area, temporal averaging time, categories, emissions or removal processes, and gases that are included. The assumptions and the methodological choice determine the needs for data and information. There can be some interaction between data and assumptions and methodological choice, indicated by the two-way arrow in the figure. For example, the ability to disaggregate categories, which may be necessary for higher tier methodologies, can depend on the availability of data. Data, whether empirical or based on expert judgment, should undergo appropriate data collection and QC procedures, as detailed in Chapters 2, Approaches to Data Collection, and Chapter 6, Quality Assurance/Quality Control and Verification, respectively.

Models can be as simple as arithmetic multiplication of activity and emission factors for each category and subsequent summation over all categories, but they may also include complex process models specific to particular categories. The data and information obtained from data collection become input to a more specific

knowledge base of data and judgment for uncertainty, as shown in the figure and as discussed in detail in Section 3.2.1, Sources of Data and Information. Specific causes of uncertainty in the conceptualisation, models, and data are discussed in Section 3.2.1 and techniques for quantifying uncertainties in input data are set out in Section 3.2.2. These necessary data include percentage uncertainty estimates and underlining probability density functions (PDFs - discussed in Section 3.1.4) for input to an emission inventory uncertainty analysis. Methods for combining input uncertainties to arrive at uncertainty estimates for single categories and the overall inventory result are detailed in Section 3.2.3. Two Approaches are given for combining uncertainties. Approach 1 is a relatively simple spreadsheet-based calculation procedure based upon some assumptions to simplify the calculations. Approach 2 is based upon Monte Carlo simulation and can be applied more generally. Either approach provides an estimate of the overall uncertainties associated with the total greenhouse gas inventory.

Figure 3.1 Overall structure of a generic uncertainty analysis



Note: Shaded Boxes are the focus of this Chapter.

3.1.3 Key concepts and terminology

Definitions associated with conducting an uncertainty analysis include *uncertainty*, *accuracy*, *precision* and *variability*. These terms are sometimes used loosely and may be misunderstood. They have in fact clear statistical definitions that should be used in order to be clear about what is being quantified and reported. Several definitions are given here, in alphabetical order:

***Accuracy*:** Agreement between the true value and the average of repeated measured observations or estimates of a variable. An accurate measurement or prediction lacks bias or, equivalently, systematic error.

***Bias*:** Lack of accuracy. Bias (systematic error), can occur because of failure to capture all relevant processes involved or because the available data are not representative of all real-world situations, or because of instrument error.

***Confidence Interval*:** The true value of the quantity for which the interval is to be estimated is a fixed but unknown constant, such as the annual total emissions in a given year for a given country. The confidence interval is a range that encloses the true value of this unknown fixed quantity with a specified confidence (probability). Typically, a 95 percent confidence interval is used in greenhouse gas inventories. From a traditional statistical perspective, the 95 percent confidence interval has a 95 percent probability of enclosing the true but unknown value of the quantity. An alternative interpretation is that the confidence interval is a range that may safely be declared to be consistent with observed data or information. The 95 percent confidence interval is enclosed by the 2.5^{th} and 97.5^{th} percentiles of the PDF.

***Precision*:** Agreement among repeated measurements of the same variable. Better precision means less random error. Precision is independent of accuracy.

***Probability density function (PDF)*:** The PDF describes the range and relative likelihood of possible values. The PDF can be used to describe *uncertainty* in the estimate of a quantity that is a fixed constant whose value is not exactly known, or it can be used to describe inherent *variability*. The purpose of the uncertainty analysis for the

emission inventory is to quantify *uncertainty* in the unknown fixed value of total emissions as well as emissions and activity pertaining to specific categories. Thus, throughout this chapter it is presumed that the PDF is used to estimate uncertainty, and not variability, unless otherwise stated.

***Random errors*:** Random variation above or below a mean value. Random error is inversely proportional to precision. Usually, the random error is quantified with respect to a mean value, but the mean could be biased or unbiased. Thus, random error is a distinct concept compared to systematic error.

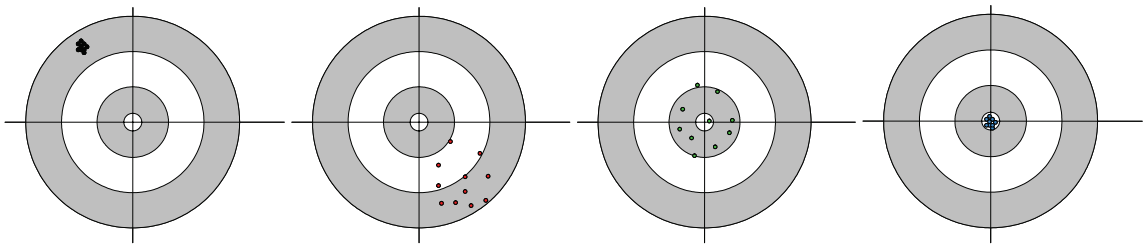
***Systematic error*:** Another term for *bias*, which refers to lack of accuracy.

***Uncertainty*:** Lack of knowledge of the true value of a variable that can be described as a probability density function (PDF) characterising the range and likelihood of possible values. Uncertainty depends on the analyst's state of knowledge, which in turn depends on the quality and quantity of applicable data as well as knowledge of underlying processes and inference methods.

***Variability*:** Heterogeneity of a variable over time, space or members of a population (Morgan and Henrion, 1990; Cullen and Frey, 1999). Variability may arise, for example, due to differences in design from one emitter to another (inter-plant or spatial variability) and in operating conditions from one time to another at a given emitter (intra-plant variability). Variability is an inherent property of the system or of nature, and not of the analyst.

Figure 3.2 Illustration of accuracy and precision

(a) inaccurate but precise; (b) inaccurate and imprecise; (c) accurate but imprecise; and (d) precise and accurate



(a) (b) (c) (d)

Inventories should be accurate in the sense that they are neither over- nor underestimated as far as can be judged, and precise in the sense that uncertainties are reduced as far as practicable. Figure 3.2 provides a conceptual comparison of accuracy and precision. An accurate inventory is one that is free of bias but that could be precise or imprecise. A precise inventory may appear to have low uncertainty but if the inventory is inaccurate, then the inventory systematically over- or under-estimates the true emissions or removals. Inaccuracy, or bias, can occur because of failure to capture all relevant emissions or removal processes or because the available data are not representative of real-world situations. There is no predetermined level of precision, in part because of the inherent variability of some categories.

3.1.4 Basis for uncertainty analysis

The chapter uses two main statistical concepts – the probability density function (PDF) and confidence interval defined in the previous section. While this chapter focuses on aspects of uncertainty that are amenable to quantification, there are typically nonquantifiable uncertainties as well. The quantitative uncertainty analysis tends to deal primarily with random errors based on the inherent variability of a system and the finite sample size of available data, random components of measurement error, or inferences regarding the random component of uncertainty obtained from expert judgement. In contrast, systematic errors that may arise because of imperfections in conceptualisation, models, measurement techniques, or other systems for recording or making inferences from data, can be much more difficult to quantify. As mentioned in Section 3.5, Reporting and Documentation, it is *good practice* for potential sources of uncertainty that have not been quantified to be described, particularly with respect to conceptualisation, models, and data and to make every effort to quantify them in the future.

Good practice requires that bias in conceptualisations, models, and inputs to models be prevented wherever possible, such as by using appropriate QA/QC procedures. Where biases cannot be prevented, it is *good practice* to identify and correct them when developing a mean estimate of the inventory. In particular, the point estimate that is used for reporting the inventory should be free of biases as much as it is practical and possible. Once

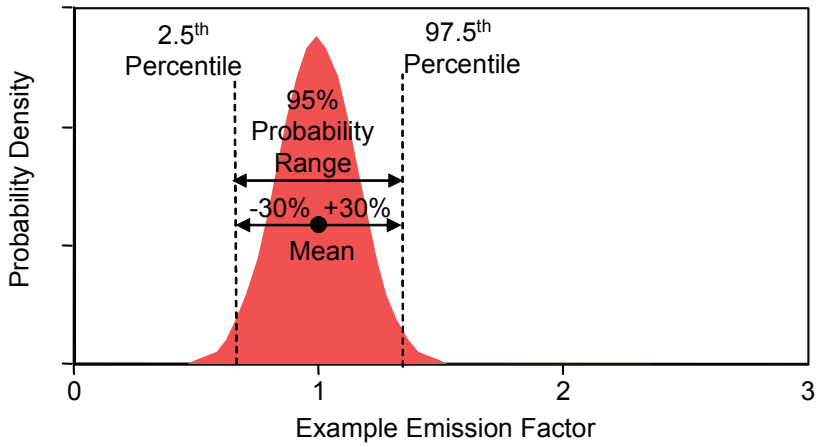
biases are corrected to the extent possible, the uncertainty analysis can then focus on quantification of the random errors with respect to the mean estimate.

Good practice requires the use of a 95 percent confidence interval for quantification of random errors. This may also be expressed as a percentage of the central estimate. Where the PDF is symmetrical the confidence interval can be conveniently expressed as plus or minus half the confidence interval width divided by the estimated value of the variable (e.g., ± 10%). Where the PDF is not symmetrical upper and lower limits of the confidence interval need to be specified separately (e.g., -30%,+50%).

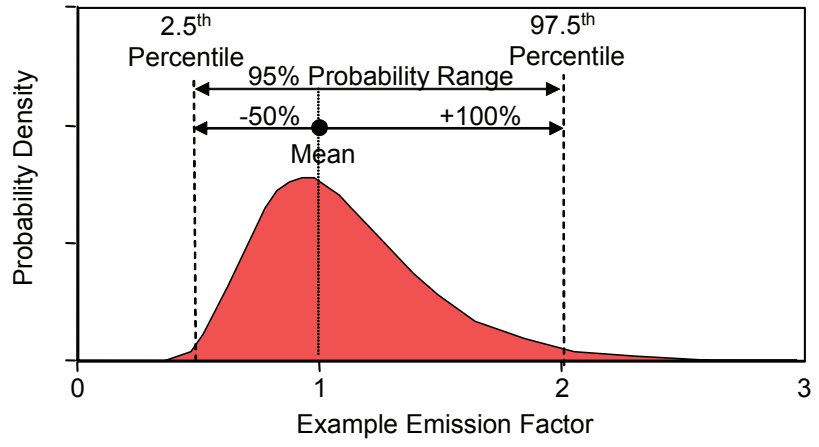
If the range of uncertainty for a non-negative variable is small enough relative to the mean value, then the uncertainty often can be described as a symmetric range with respect to a mean value, as shown in Figure 3.3(a). For example, if the mean emissions are 1.0 units, the 2.5th percentile of uncertainty is 0.7 units, and the 97.5th percentile of uncertainty is 1.3 units, then the uncertainty range could be described as 1.0 units ±30%. However, when the relative range of uncertainty is large, and if the uncertainty is regarding a variable that must be non-negative (such as an emission factor), then the uncertainty range becomes asymmetric with respect to the mean, as shown in Figure 3.3(b). As an example, if the mean emissions are 1.0 units, the 2.5th percentile of uncertainty is 0.5 units, and the 97.5th percentile of uncertainty is 2.0 units, then the range of uncertainty can be described as 1.0 units -50% to +100%. In situations such as the latter, it is often more convenient to summarise uncertainties in a multiplicative, rather than additive, manner. In this particular example, the lower end of the 95 percent probability range is a half the mean, and the upper end is a multiplier of 2 larger than the mean. Such a range is commonly summarised as a "factor of 2." An uncertainty of a "factor of n" refers to a range bounded at the low end by (mean/n) and at the high end by (mean × n). Thus, a factor of 10 uncertainty would have a range of 0.1×mean to 10×mean. The factor 10 uncertainty is also often called "an order of magnitude". Higher powers of 10 are referred to as "orders of magnitude;" for example, a factor of 10^3 would be referred to as three orders-of-magnitude.

Figure 3.3 Examples of symmetric and asymmetric uncertainties in an emission factor

(a) Example of a symmetric uncertainty of ±30% relative to the mean



(b) Example of an asymmetric uncertainty of -50% to +100% relative to the mean, or a factor of two



3.1.5 Causes of uncertainty

Inventory estimates of emissions and removals differ from the true underlying value for many reasons. Some causes of uncertainty (e.g., sampling error or limitations on instrument accuracy) may generate well-defined, easily characterised estimates of the range of potential uncertainty. Other causes of uncertainty (e.g., biases) may be much more difficult to identify and quantify (Rypdal and Winiwarter, 2001). It is *good practice* to account, as far as possible, for all causes of uncertainties in an uncertainty analysis and clearly document if some causes of uncertainties have not been included.

The inventory developer should consider eight broad causes of uncertainty[1]:

- *Lack of completeness*: This is a case where measurement or other data are not available either because the process is not yet recognized or a measurement method does not yet exist. Typically, this cause can lead to incomplete conceptualisation, which results in bias, but can also contribute to random error depending on the situation.
- *Model*: Models can be as simple as a constant multiplier (e.g., an emission factor) and increase in complexity, such as for complicated process models. The use of models to estimate greenhouse gas emissions and removals can introduce uncertainty, including both bias and random error, for a variety of reasons:
  - (i) Models are a simplification of real systems and therefore are not exact. For example, computer programming may involve errors or approximations; model resolution may not be representative, and spatial and temporal coverage may not be fully representative;
  - (ii) Interpolation is application of a model within a range of inputs for which the model is considered to be valid. However, in some cases, a 'hidden extrapolation' can occur when the model is evaluated based on combinations of values of its inputs for which validation has not been done (Cullen and Frey, 1999).
  - (iii) Extrapolation (application of the model beyond the domain for which model predictions are known to be valid) can lead to uncertainty;
  - (iv) Alternative formulations of the model may result in different estimates; and
  - (v) Model inputs including activity data and parameters are generally approximated based on limited information that create additional uncertainties beyond the model formulation.
- *Lack of data*: In some situations, there simply may not yet be data available that would be necessary to characterise a particular emission or removal. In these situations, a common approach is to use proxy (or surrogate) data for analogous or similar categories or to use interpolation or extrapolation as a basis for making estimates.
- *Lack of representativeness of data*: This source of uncertainty is associated with lack of complete correspondence between conditions associated with the available data and the conditions associated with real world emissions/removals or activity. For example, emissions data may be available for situations in which a plant is operating at full load but not for situations involving start-up or load changes. In this case, the data are only partly relevant to the desired emission estimate. Lack of representativeness typically leads to bias.
- *Statistical random sampling error*: This source of uncertainty is associated with data that are a random sample of finite size and typically depends on the variance[2] of the population from which the sample is extracted and the size of the sample itself (number of data points). It can often be reduced by increasing the number of independent samples taken. Here, it is *good practice* to distinguish properly between *variability* and *uncertainty*, as previously defined. For purposes of uncertainty analysis of inventories, one is typically interested in uncertainty in the annual average at the national level, rather than the entire range of variability that might occur over shorter periods of time or small geographic scales. Larger sample sizes will not reduce the inherent variability, but will lead to narrower confidence intervals that are a basis for estimating the random component of uncertainty.

[1] Further discussion can be found in Morgan and Henrion (1990) and Cullen and Frey (1999).

[2] The *variance of a whole population* of values is the average of the square of the difference between individual values in the population and the mean value. The *variance of a sample drawn from a population* is the sum of the squares of differences between values in the sample and the mean of the sample, divided by the number of values in the sample less 1.

- *Measurement error*: Measurement error, which may be random or systematic, results from errors in measuring, recording and transmitting information; finite instrument resolution; inexact values of measurement standards and reference materials; inexact values of constants and other parameters obtained from external sources and used in the data-reduction algorithm (e.g., default values from the *IPCC Guidelines*); approximations and assumptions incorporated in the measurement method and estimation procedure; and/or variations in repeated observations of the emission or removal or associated variable under apparently identical conditions.
- *Misreporting or misclassification*: Uncertainty here may be due to incomplete, unclear, or faulty definition of an emission or removal. This cause of uncertainty typically leads to bias.
- *Missing data*: Uncertainties may result where measurements were attempted but no value was available. An example are measurements that are below a detection limit. This cause of uncertainty can lead to both bias and random error. When measured values are below a detection limit, an upper bound on the uncertainty can be estimated. There are rigorous statistical techniques for dealing with non-detected data as well as other types of missing data, such as data that are missing at random (Cohen and Whitten, 1998; Gelfand, 1996; Zhao and Frey, 2004b). These techniques may involve estimation or imputation in portions of the distribution where data are not available.

Particularly on the issue of extrapolation, uncertainty occurs when extrapolating from recent source and sink data for the purpose of estimating an inventory for a year of interest for which data are not yet available (see also Chapter 5, Time Series Consistency). Usually, the extrapolated estimates are reported as 'provisional' estimates and then later are updated when the relevant data become available. However, until the update occurs, the provisional inventory might be used. The additional uncertainty associated with extrapolation is a type of model uncertainty. Errors associated with extrapolation can be systematic, random, or both. If there is a history of extrapolations and subsequent correction, then it is possible to develop data regarding the distribution of the errors that have been observed in the past. If there are biases in the provisional estimates, then the mean of this distribution will not be zero and the biases can be quantified. This distribution would represent error in the ability to predict actual source and sink fluxes based upon the extrapolation methods used in the past. If the extrapolation methods change, then expert judgement could be used to quantify the uncertainty.

Where a PDF for the mean can be identified, various causes of uncertainty can be quantified by statistical means. As noted in Section 3.2, uncertainties can be quantified by statistical analysis of empirical data, by encoding (quantifying) of expert judgement in the form of PDFs, or by combinations of both. However, there can be structural uncertainties that are not easily incorporated into a quantitative uncertainty analysis in the form of a PDF. Examples of structural uncertainties include possible misidentification or mis-specification of the system to be analysed, as well as possible problems associated with the models that are used, e.g., inappropriateness of the model, or model errors. These latter types of situations are typically outside the scope of statistics (ISO 1993)[3], although probabilistic methods for dealing with model uncertainties have been proposed (e.g., Evans *et al*., 1994). For example, expert judgement can be used to assign weights to alternative models.

Table 3.1 suggests how different causes of uncertainties can be treated in an analysis. Some causes of uncertainties (e.g., misreporting/misclassification) may be reduced or eliminated by implementing QA/QC procedures and improvements in data collection and/or methodologies when identified.

[3] There are some opportunities for addressing these sources of uncertainty. For example, uncertainties associated with model error at least partly can be addressed by comparing modelled output with measured values. Depending on how modelled outputs compare with measurements, one can identify biases associated with the model that may vary depending on the type of system being modelled.

TABLE 3.1
TYPICAL STRATEGIES FOR DEALING WITH DIFFERENT CAUSES OF UNCERTAINTIES

Causes of Uncertainty	Strategy			Other Comments [1]
	Evaluated Conceptualisation and Model Formulation	Empirical and Statistical	Expert judgement	
Lack of completeness	√			Have key components of the system been omitted? If so, what is the quantifiable or nonquantifiable effect on systematic error? Proper QA/QC should help avoid this.
Model (bias and random errors)	√	√	√	Is the model formulation complete and accurate? What is the uncertainty in model predictions based on validation of the model? What is the estimate of model accuracy and precision based on expert judgment if statistical validation data are not available?
Lack of data			√	If data are lacking, can expert judgment be used to make inferences based on analogous (surrogate, proxy) data or theoretical considerations? May be related to lack of completeness and model uncertainty.
Lack of representativeness of data	√	√	√	
Statistical random sampling error		√		E.g., statistical theory for estimating confidence intervals based on variability in the data and sample size.
Measurement error: random component		√	√	
Measurement error: systematic component (bias)	√		√	QA/QC and verification may provide insight.
Misreporting or Misclassification		√	√	Proper QA/QC should help avoid this.
Missing data		√	√	Statistical or judgment-based approaches to estimating uncertainty because of non-detected measurements or other types of missing data.

[1] It is *good practice* to apply procedures for QA/QC and verification prior to or combining with developing uncertainty estimates according to the guidance in Chapter 6. The QA/QC and verification procedures provide a useful basis for preventing mistakes and for identifying (and preferably correcting) biases. Furthermore, QA/QC should prevent or help detect and correct misreporting and misclassification errors, and there should be iteration between uncertainty analysis and QA/QC if application of the uncertainty methods uncovers potential QA/QC problems.

3.1.6 Reducing uncertainty

Uncertainties should be reduced as far as is practicable during the process of compiling an inventory, and it is particularly important to ensure that the model and the data collected are fair representations of the real world. When focusing efforts to reduce uncertainty, priority should be given to those inputs to the inventory that have the most impact on the overall uncertainty of the inventory, as opposed to inputs that are of minor or negligible importance to the assessment as described in Chapter 4, Methodological Choice and Identification of Key Categories. Tools for prioritising where uncertainties should be reduced include *key category* analysis (see Chapter 4) and assessment of the contribution of uncertainties in specific categories to the total uncertainty in the inventory (see Section 3.2.3). Depending on the cause of uncertainty present, uncertainties could be reduced in seven broad ways:

- *Improving conceptualisation*: Improving the inclusiveness of the structural assumptions chosen can reduce uncertainties. An example is better treatment of seasonality effects that leads to more accurate annual estimates of emissions or removals for the AFOLU Sector.
- *Improving models*: Improving the model structure and parameterisation can lead to better understanding and characterisation of the systematic and random errors, as well as reductions in these causes of uncertainty.

- *Improving representativeness*: This may involve stratification or other sampling strategies, as set out in Section 3.2.1.2. This is particularly important for categories in the agriculture, forestry and land use parts of an inventory, but also applies elsewhere, e.g., wherever different technologies are operating within a category. For example, continuous emissions monitoring systems (CEMS) can be used to reduce uncertainty for some sources and gases as long as the representativeness is guaranteed. CEMS produces representative data at the facilities where it is used, but in order to be representative of an entire source category, CEMS data must be available for a random sample or an entire set of individual facilities that comprise the category. When using CEMS both concentration and flow will vary, requiring simultaneous sampling of both attributes.
- *Using more precise measurement methods*: Measurement error can be reduced by using more precise measurement methods, avoiding simplifying assumptions, and ensuring that measurement technologies are appropriately used and calibrated. See Chapter 2, Approaches to Data Collection.
- *Collecting more measured data*: Uncertainty associated with random sampling error can be reduced by increasing the sample size. Both bias and random error can be reduced by filling in data gaps. This applies both to measurements and surveys.
- *Eliminating known risk of bias*: This is achieved by ensuring instrumentation is properly positioned and calibrated (see Section 2.2 in Chapter 2), models or other estimation procedures are appropriate and representative as indicated by the decision trees and other advice on methodological choice in sectoral volumes, and by applying expert judgements in a systematic way.
- *Improving state of knowledge*: Generally, improving the understanding of the categories and the processes leading to emissions and removals can help to discover, and correct for, problems of incompleteness. It is *good practice* to continuously improve emissions and removal estimates based on new knowledge (see Chapter 5, Time Series Consistency).

3.1.7 Implications of methodological choice

Choice of methodological tier for emissions and removals estimation can affect the uncertainty analysis in two different ways. Firstly, moving to higher tier inventory methods should typically reduce uncertainties, provided the higher tier methods are well implemented, because they should reduce bias and better represent the complexity of the system. Secondly, moving to higher tier methods may result in increased estimates of uncertainty in some circumstances. Often, this increase in the estimated uncertainty does not actually represent a decrease in knowledge; rather, it typically reveals a more realistic acknowledgment of the limitations of existing knowledge. This may occur where there was an incomplete accounting of the greenhouse gas emissions in the lower tier method, or where application of higher tier methods reveals additional complexity and uncertainties that were not fully apparent in the lower tier method. This really means that the uncertainty was underestimated previously and moving to the higher tier method in reality produces a more accurate estimate of uncertainty. In some cases, an increase in uncertainty may occur for one inventory development method versus another because each method has different data requirements. For example, sometimes aggregate estimates of emissions are more accurate because they are based upon or can be compared to easily measured values, whereas disaggregated estimates may require additional assumptions for which data or the capability to verify estimates are not as readily available. The appropriate level of disaggregation can differ within and between categories.

3.2 QUANTIFYING UNCERTAINTIES

After identifying the causes of uncertainties associated with inventory estimates, the inventory compiler should collect the appropriate information to develop national, and category-specific estimates of uncertainty at the 95 percent confidence interval. Ideally, emission and removal estimates and uncertainty ranges would be derived from category-specific measured data. Since it may not be practical to measure every emission source or sink category in this way, other methods for quantifying uncertainty may be required. The pragmatic approach for producing quantitative uncertainty estimates is to use the best available estimates, which are often a combination of measured data, published information, model outputs, and expert judgement. The sectoral guidance in Volumes 2 to 5 of *these Guidelines* provides default uncertainty estimates for use with the methods described in this chapter.

Although uncertainties determined from measured data are often perceived to be more rigorous than uncertainty estimates based on models, and similarly, model-based estimates are often perceived as more rigorous than those based on expert judgement, the actual hierarchy depends on the category and/or country-specific circumstances. In particular it is *good practice* to ensure that uncertainties are representative for the application in the inventory and national circumstances and includes all causes of uncertainty listed in Table 3.1.

This section is organised into three major subsections that are interrelated. Section 3.2.1 focuses on sources of data and information that can be used to identify and, where possible, quantify uncertainties. Section 3.2.2 focuses on methods for attempting to prevent or correct for biases and for quantifying the random component of uncertainty in the inputs to models. Section 3.2.3 presents two Approaches for combining uncertainties in inputs in order to arrive at uncertainty estimates for single emission and removal categories and the total emission inventory.

3.2.1 Sources of data and information

This section identifies sources of data and information for acquiring quantitative estimates of uncertainty. There are three broad sources of data and information: information contained in models; empirical data associated with measurements of emissions, and activity data from surveys and censuses; and quantified estimates of uncertainties based upon expert judgement.

3.2.1.1 UNCERTAINTIES ASSOCIATED WITH MODELS

A model is a representation of a real-world system. Modelling typically involves choices regarding what to include versus what to exclude, as well as choices regarding the level of detail (or aggregation) for those phenomena that are included in the model. Thus, the model typically does not exactly mimic the real-world system. The structure of the model is often thought of in terms of the equations used and in terms of inputs and outputs of the model (Kirchner, 1990). More generally, a model may be thought of as a hypothesis regarding how the real-world system behaves. Thus, there are two key considerations in model uncertainty: (1) has the correct, most relevant real-world system been identified, and have conceptualisations been constructed in a way that properly serve as the basis for model development; and (2) is the model an accurate representation of the chosen system. C*onceptualisation uncertainty* describes the lack of proper identification of the system for which a model should be developed and of the conceptualisation(s) of interest. *Model uncertainty* describes the lack of proper model development relative to the intended system and conceptualisation(s).

***Conceptualisation Uncertainty*:** The failure to specify properly appropriate and relevant inventory structural assumptions is known as conceptualisation uncertainty (Cullen and Frey, 1999) and typically results in a bias in estimates. The causes of conceptualisation uncertainty typically include descriptive errors, errors in professional judgement, and incomplete specification of the assumptions (EPA, 1997).

***Model uncertainty*:** Uncertainty arises from imperfections in how the chosen conceptualisations are modelled. Sometimes these imperfections occur because of limitations of available data. A model may have other sources of structural errors, such as failure to properly take into account the sensitivity of emissions to ambient conditions or other factors. Modelling can be a basis for estimating emissions or removals for specific categories as well as for managing data in the entire inventory. In some cases, model uncertainty can be significant. It is typically poorly characterised and may not be characterised at all.

3.2.1.2 EMPIRICAL DATA FOR SOURCES AND SINKS AND ACTIVITY

This section describes sources of empirical data, and their implications for uncertainty, and is relevant to measured emissions data, data obtained from literature, and activity data.

UNCERTAINTY ESTIMATES OBTAINED FROM MEASURED EMISSIONS/REMOVALS DATA

This section assumes that *good practices* are used to obtain the data, as outlined in Chapters 2 and Chapter 6, Quality Assurance/Quality Control and Verification. When estimating uncertainty from measured emissions data, considerations include: (a) representativeness of the data and potential for bias; (b) precision and accuracy of the measurements; (c) sample size and inter-individual variability in measurements, and their implications for uncertainty in mean annual emissions/removals; (d) inter-annual variability in emissions/removals and whether estimates are based upon an average of several years or on the basis of a particular year.

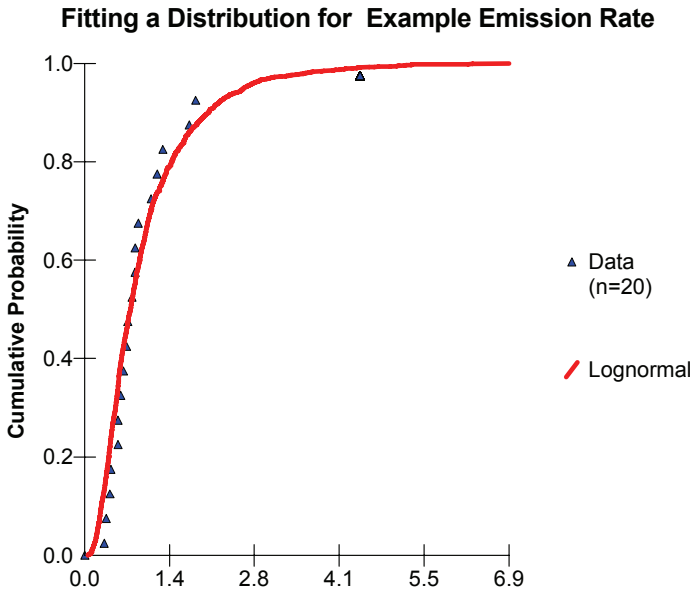
Representative sampling (or sampling design) implies that measurements are made for typical system characteristics, operating conditions, time periods, and/or geographic areas of interest. The precision and accuracy of individual measurements will depend upon the equipment and protocols used to make the measurements. The sample size will often be a trade-off between the desirability for more data and the cost of making measurements. In some cases, such as for continuous monitoring, the sample size may be large enough to effectively serve as a census, rather than a partial sample, of data. In general the variability in the data from

one short-term time period (e.g., hour, day, week) to another will depend upon the characteristics of the category. If the goal is to develop an estimate of annual average emissions or removals, then judgement may be required as to whether measurements conducted over a short term are representative of rates over a longer time period and, if not, whether the measurement programme can be expanded to additional time periods. For example, flux measurements (data on emission factors) should represent the entire year. In the AFOLU Sector this is crucial, since emissions are highly dependent on climatic conditions which typically are not the same for the growing period and winter period.

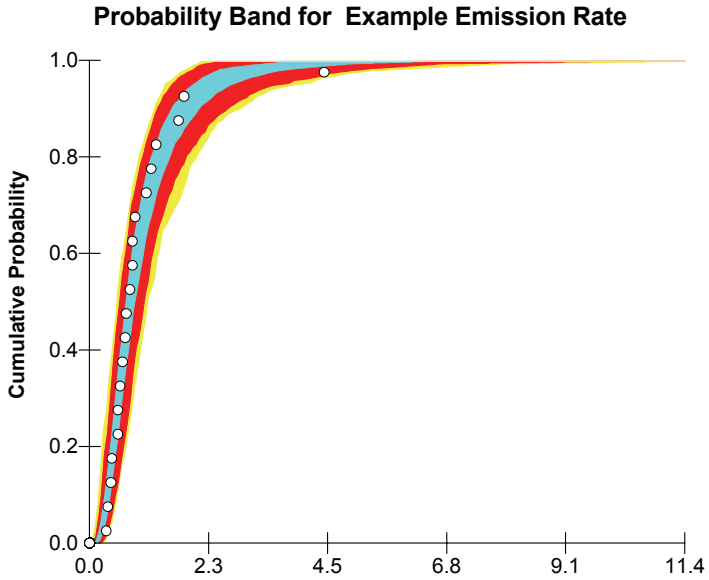
Figure 3.4 Example of uncertainty in emission measurements and mean emission rate

(a) Fitted distribution for inter-unit variability in emissions;
(b) Uncertainty in fitted distribution because of small sample size (n=20);
(c) Uncertainty in mean emission rate.

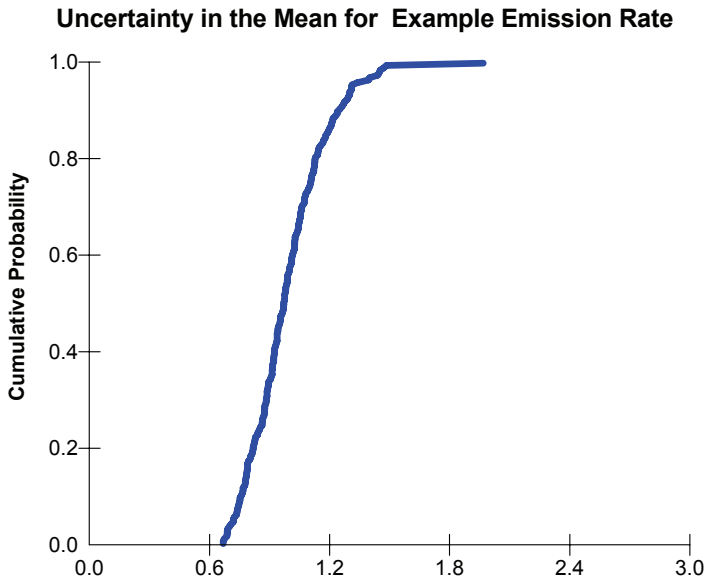
(a) Inter-Unit Variability



(b) Uncertainty in Distribution of Variability



(c) Uncertainty in Mean



For a second example, suppose that one wishes to estimate the uncertainty in national annual emissions for a particular category, such as emissions from gasoline-fuelled passenger automobiles. The rate of emission varies from one individual vehicle to another, illustrated by the inter-unit variability shown in Figure 3.4(a). Because the distribution for inter-vehicle variability is estimated from a small, finite sample of data that could be subject to random sampling error, there is uncertainty regarding what the true but unknown population distribution for inter-vehicle variability might be, as suggested in Figure 3.4(b). There is also intra-unit variability in emissions over time for any particular vehicle. However, for purposes of the national annual estimate, the focus is on the combined contribution of all such vehicles to total emissions during a year long time frame. In this case, we are not interested in the range of inter-vehicle variability, but rather in the range of uncertainty for the average emission rate among all such vehicles (e.g., Figure 3.4(c)). Often, the range of uncertainty is substantially less than that for inter-vehicle (or, more generally, inter-unit) variability (e.g., Frey and Zheng, 2002). Therefore, when the objective of an analysis requires that the assessment be based upon uncertainty in the mean, rather than variability among individual units, it is important to properly focus the analysis on the former. Failure to do so can lead to a misleading over-estimate of the range of uncertainty.

In the case of continuous monitoring of point emissions, or a periodic sampling scheme that captures typical activity patterns, there may be adequate and representative empirical data upon which to base an estimate of uncertainty in mean annual emissions. For example, if there are several years of such data, then the average annual emissions over several years can be quantified, and the distribution of annual emissions from year-to-year can be used to assess a 95 percent confidence interval in the annual average. Provided that the annual average is based upon data from many individual categories, it is unlikely that there will be correlation of errors between years. This has implications for estimation of uncertainty in trends, as discussed in Section 3.3, Uncertainty and Temporal Autocorrelation. However, for diffuse categories, such as agricultural crops, there could be high autocorrelations if they are determined by climate, and this could affect the representativeness of the data for a particular assessment purpose.

Where continuous emission measurements are not available, there may be periodic emission measurements available from which to estimate uncertainty. If these measurements can be linked to representative activity data, which of course is crucial, then it is possible to determine a site-specific emission factor, together with an associated PDF to represent annual emissions. This can be a complex task. To achieve representativeness it may be necessary to partition (or stratify) the data to reflect typical operating conditions. For example:

- Start-up and shut down can give different emission rates relative to activity data. In this case, the data should be partitioned, with separate emission factors and probability density functions derived for steady state, start-up and shut down conditions.
- Emission factors can depend on load. In this case, the total emissions estimation and uncertainty analysis may need to be stratified to take account of load, expressed, for example, as percentage of full capacity. This could be done by regression analysis and scatter plots of the emission rate against likely controlling variables (e.g., emissions versus load) with load becoming part of the activity data needed.

- Measurements taken for another purpose may not be representative. For example, methane measurements made for safety reasons at coal mines and landfills may not necessarily reflect total emissions because they may have been made only when methane emissions were suspected of being high, as a compliance check. In such cases, the ratio between the measured data and total emissions should be estimated for the uncertainty analysis.
- Systematic short-term measurements might not adequately sample episodic events (such as rainfall) that initiate large fluxes of short duration that may nevertheless account for a major fraction of annual emissions. If the sampling strategy misses a significant proportion of these events, then the annual average emission estimate could be substantially biased. Nitrous oxide emissions from agricultural soils can fall into this class.

If the data sample size is large enough, standard statistical goodness-of-fit tests can be used, in combination with expert judgement, to help in deciding which PDF to use for describing variability in the data (partitioned if necessary) and how to parameterise it. However, in many cases, the number of measurements from which to make an inference regarding uncertainty will be small. Theoretically, as long as there are three or more data points, and they are a random representative sample of the variable of interest, it is possible to apply statistical techniques to estimate the values of the parameters of many two-parameter distributions (e.g., normal, lognormal) that can be used to describe variability in the data set (Cullen and Frey, 1999, pp. 116-117). While it is commonly perceived that one must have approximately 8 or 9 data points, and preferably more, as the basis for fitting a distribution to data, the more fundamental and key assumption that must be made in order to fit a distribution to data is that the data are a random, representative sample. If this assumption is valid, then the sample size influences the width of the confidence intervals for any statistic estimated from the sample. As a matter of preference, many analysts may prefer to have a minimum sample size, but this preference is not related to the key issue of representativeness. Data do not become more representative only because of an increase in sample size.

With small sample sizes, there will be large uncertainties regarding the parameter estimates that should be reflected in the quantification of uncertainty for use in the inventory. Furthermore, it is typically not possible to rely on statistical methods to differentiate goodness-of-fit of alternative parametric distributions when sample sizes are very small (Cullen and Frey, 1999, pp. 158-159). Therefore judgement will be required in selecting an appropriate parametric distribution to fit to a very small data set. In situations where the coefficient of variation (standard deviation divided by the mean) is less than approximately 0.3 and is known with reasonable confidence, a normal distribution may be a reasonable assumption (Robinson, 1989). When the coefficient of variation is large and the variable is non-negative, then a positively skewed distribution such as a lognormal may be appropriate. Guidance on the selection of distributions is elaborated in Sections 3.2.2.2 and 3.2.2.4 below.

In cases with large data sets, the uncertainty in the mean can often be estimated as plus or minus 1.96 (or approximately 2) multiples of the standard error, where the standard error is the sample standard deviation divided by the square root of the sample size. This calculation is based on an assumption of a normal distribution. However, in cases of a small number of samples/measurements that will often be the case in determining emission factors, the multiple of 1.96 is replaced with a "coverage factor," referred to as k, that is obtained from the student's t-distribution. For small sample sizes, k is greater than 1.96 for a 95 percent interval, but asymptotically approaches 1.96 as the sample size increases to approximately 30 or more. However, in cases where the uncertainty in the mean is not a symmetric distribution, then numerical methods such as bootstrap simulation can be used instead to obtain the confidence interval for the mean.

Where an annual estimate is based on an average over several years, the uncertainty in the average represents the uncertainty in an average year and not the inter-annual variability. If the objective is to estimate uncertainty in source or sink fluxes for a specific year, then *good practice* is to make a best estimate of the annual total and to quantify uncertainty associated with the models and data used consistent with the one year time period. If, instead, an averaged annual estimate is used, then the uncertainty in the estimate when applied to a specific year would be described by the inter-annual variability (including measurement errors) relative to the mean, whereas when applied to an average year it would be the confidence interval of the average.

UNCERTAINTY ESTIMATES FOR EMISSION FACTORS AND OTHER PARAMETERS OBTAINED FROM PUBLISHED REFERENCES

When site-specific data are unavailable, inventories should where possible be based on emission factors derived from published studies specific to the conditions in that country. Where sufficient country-specific information is unavailable, information may be derived from other published studies if those studies are reflective of conditions in the country, or emission factors or other estimation parameters may be drawn from sectoral Volumes 2 to 5 of these *Guidelines*. Factors provided in the sectoral volumes have been derived for circumstances that are judged to be typical. There are uncertainties associated with the original measurements, as well as with the use of the factors in circumstances other than those associated with the original measurements.

Where published emission factors or other estimation parameters are used, the associated uncertainties should be estimated from:

- *Original research including country-specific data*: For measurement-based emission factors, the data from the original measurement programme or experiments may enable an assessment of the uncertainty and possibly the PDF. Well-designed measurement programmes and experiments will provide sample data that cover the range of types of plants and their maintenance, size and age, so that the factors and their uncertainties can be used directly. In other cases, expert judgement, taking into account the causes of uncertainty identified in Table 3.1, will be needed to extrapolate from the measurements to the full population of plants in that particular category (detail on how to elicit expert judgement is elaborated in Section 3.2.1.3).
- *Default Values from Guidelines*: For most emission factors and other estimation parameters, sectoral guidelines provide default uncertainty estimates that should be used in the absence of other information. Unless clear evidence to the contrary is available, the PDFs are assumed to be normal. However, the inventory compiler should evaluate the representativeness of the default for its own national circumstances. If the default is judged to be not representative and the category is important to the inventory, improved assumptions based upon expert judgement should be developed, assuming sufficient original research is unavailable to derive country-specific emission factors or other estimation parameters.

Default methods represent a compromise between the level of detail that would be needed to create the most accurate estimates for each country, and the input data likely to be available or readily obtainable in most countries. Default methods are often simplifications, and may introduce large uncertainties into a national estimate. Within many of the default methods different optional levels of detail are provided to reflect whether users have detailed data for their national situation or have to rely strictly on general default values. There may be considerable variation in how well the general default values represent conditions of the actual population of activities in a particular country. For example, the uncertainty relating to default carbon emission factors for the global population of fossil fuel combustion sources may be characterised as quite low (5-10 percent) in the IPCC methodology; but national experts for a particular country may know that the characteristics of such fuels in their country vary from global average values. In such a country, use of default values would introduce a larger uncertainty, and thus it is *good practice* to use country-specific estimates where possible. Thus the applicability of default uncertainty values should always be considered.

Another example is the use of default values to estimate country-specific emissions and removals of the AFOLU Sector. Uncertainty could be high unless the suitability of the available default parameters to a country's circumstances is known. The application of default data in a country or region that has very different characteristics from those of the category data can lead to large systematic (bias) errors in estimates of emissions or removals.

UNCERTAINTIES ASSOCIATED WITH ACTIVITY DATA

Activity data are often more closely linked to economic activity than are emission factors are. However, unlike emission factor data, there is typically no statistical sample of alternative activity data estimates readily available to fit distributions and estimate uncertainty. There are often well established price incentives and fiscal requirements for accurate accounting of economic activity. Activity data therefore tend to have lower uncertainties and a lower correlation between years than emission factor data. Activity data are often collected and published regularly by national statistical agencies, which may have already assessed the uncertainties associated with their data as part of their data collection procedures. These previously developed uncertainty estimates can be used to construct PDFs. This information will not necessarily have been published, so it is recommended to contact the statistical agencies directly. Since economic activity data are not usually collected for the purpose of estimating greenhouse gas emissions and removals, it is *good practice* to assess the applicability of the uncertainty estimates before using them.

There are several approaches that may be helpful in assessing the uncertainty of activity data in particular circumstances:

***Activity data based on complete samples* (*censuses*):** Census data are activity data that are based, in principle, on counting every instance of a particular activity. Census typically includes both systematic and random errors. Systematic errors arise through systematic undercounting or double counting. Random errors are typically the sum of a range of commonplace errors. Random errors usually can be expected to be normally distributed and serially uncorrelated. Because activity data are usually collected by the same people, using the same processes, for each observation, systematic errors are likely to take approximately the same value each year. There are several approaches to identifying the potential uncertainty of activity data for complete samples. These approaches are often an integrated part of a QA/QC plan:

- To check for the size of random errors, look for fluctuations over time, and differential fluctuations in series that ought to be highly correlated with the data of interest.

- To check for bias errors, cross-check the data of interest with other, related information. One might, for instance, look up and down the supply chain for fuels, comparing coal production, coal import/export, and reported consumption. Or, one might study activities for which data are collected independently but which ought to be highly correlated with the data of interest, for instance reported fuel input vs. electricity output. One might also look at activity data of different frequencies (e.g., monthly, annual), if they are collected using different approaches.
- Interpretation of statistical differences, within, for instance, national energy data are an example of cross-checking. The comparison between energy-related carbon dioxide emissions derived from the IPCC reference approach is a formal cross-check with emissions estimates derived from other sources.

Census-based activity data are often 'precise but inaccurate' in the taxonomy shown in Figure 3.2, the random errors are small, but there may be larger bias errors. Cross-checking can suggest upper and lower bounds for possible bias errors, and sometimes will permit an actual estimate of the bias error. A possible bias error lurking within these bounds may often be characterised as a truncated uniform distribution: cross-checking shows that the unobservable true value must lie within a particular range, but there may be no reason to think any point within that range is more or less likely. However, because the bias errors in activity data are likely to be highly correlated, the difference between the reported value and the unknown true value is likely to be about the same every year, and this characteristic should be taken into account when estimating trend uncertainty.

***Activity data based on random samples*:** Some kinds of activity data are derived from sample surveys, for instance consumer surveys, land use surveys, or forest cover surveys. In these cases, the data will be subject to sampling errors, which will be normally distributed and uncorrelated over time. The agency conducting the sample will normally be able to advise on sampling error. If this information is unavailable, it may be possible to identify or infer the sample and population sizes and calculate sampling error directly.

3.2.1.3 EXPERT JUDGEMENT AS A SOURCE OF INFORMATION

In many situations, directly relevant empirical data are not available for sources, sinks, or activity inputs to an inventory. In such situations, a practical solution to dealing with the absence of adequate data is to obtain well-informed judgements from domain experts regarding best estimates and uncertainties of inputs to the inventory. Chapter 2, Approaches to Data Collection, discusses the basis of formal expert elicitation protocols. In particular, Section 2.2 and Annex 2A.1 provides a general treatment of expert judgment and elicitation. Annex 2A.1 provides details regarding expert elicitation protocol. In this chapter, methods for encoding uncertainties based upon expert judgement are recommended in Section 3.2.2.3.

3.2.2 Techniques for quantifying uncertainties

This section discusses key techniques for quantifying uncertainties, building upon the sources of data and information described in the previous section. This section focuses on uncertainty in models, statistical analysis of empirical data, identifying and selecting PDFs, and methods for encoding expert judgement regarding uncertainties.

3.2.2.1 UNCERTAINTY IN MODELS

Conceptualisation and model uncertainty can be more difficult to address than the uncertainties in the inputs to a model. The most significant concern with conceptualisation and model uncertainties is that they have the potential to produce substantial bias in emissions and removal estimates. Approaches to deal with these causes of uncertainties should aim therefore to evaluate and correct for known or suspected biases.

It is clear that proper specification of a conceptualisation is set by the *2006 Guidelines*, the interpretation of which depends on input from experts and stakeholders who are familiar with the systems for which emissions or removals are to be estimated. A conceptualisation should be complete, within the scope of *these Guidelines*, in enumerating all key components without producing redundancy or overlap, and it should be applicable to the geographic scope, time period, and agreed set of greenhouse gases covered.

Model uncertainty is typically dealt with in several ways. One approach is to simply acknowledge the limitations of models that are used and to qualitatively discuss the implications for uncertainty in estimates obtained using the model. However, qualitative caveats are not useful in providing quantitative insight regarding the possible magnitude of uncertainty, and by themselves are not considered *good practice*. There are at least three major approaches for estimating uncertainty: (1) comparison of model results with independent data for purposes of

verification; (2) comparison of the predictions of alternative models; and (3) expert judgement regarding the magnitude of model uncertainty. These approaches can be used in combination.

Comparison of model predictions with independent data can be used to assess the precision and accuracy of the model, and is an important aspect of verification, as discussed in Chapter 6. Such comparisons can reveal whether the model systematically over- or under-predicts the quantities of interest. However, it can be difficult to obtain data for direct verification of a model. Nonetheless, sometimes these types of comparisons are the best or only available ones, and might help in identifying unexplained inconsistencies that, in turn, might imply model bias that could be corrected for by parameter choice.

In other cases, there may be alternative models that could be used to make predictions for the same quantities of interest. To the extent that the alternative models are based upon different data or theoretical assumptions, a comparison of model predictions may provide useful insight regarding the magnitude of disagreement. The fact that two or more models disagree is not conclusive proof that either of the models is wrong, since both or all of the models could be wrong.

Based on the results of the comparison of the model used for inventory development with independent data and/or alternative models, it may be desirable to revise model assumptions or parameters to reduce the bias. The remaining uncertainty can then be quantitatively assessed by expert judgement about how uncertainties in the data used to drive the model and the model parameters combine, or more formally by Monte Carlo analysis.

3.2.2.2 STATISTICAL ANALYSIS OF EMPIRICAL DATA

Statistical analysis of empirical data is an approach that can be employed to quantify uncertainty in inventories, emission factors and other estimation parameters, and it can be summarised as the following major steps (e.g., Frey and Zheng, 2002):

Step 1: Compilation and evaluation of a database for emission factors, activity data and other estimation parameters. Such data typically represent variability.

Step 2: Visualisation of data by developing empirical distribution functions (in which the data are plotted vertical according to their rank order and are plotted horizontally according to their numerical value – see Cullen and Frey, 1999, for details) for individual activity and emission factors.

Step 3: Fitting, evaluation, and selection of alternative PDF models for representing variability in activity data and emission factor data.

Step 4: Characterisation of uncertainty in the mean of the distributions for variability. If the standard error of the mean is small enough (as discussed in Section 3.2.1.2), a normality assumption can be made regardless of the sample size or skewness of the data. If the standard error of the mean is large, then either a lognormality assumption can be made, or other methods can be employed (e.g., bootstrap simulation) to estimate uncertainty in the mean. Publicly available software tools could be used to assist with the latter.

Step 5: Once uncertainties have been appropriately specified, these can be used as input to a probabilistic analysis for purposes of estimating uncertainty in total emissions.

Step 6: Sensitivity analysis is recommended to determine which of the input uncertainties to an inventory contributes most substantially to the overall uncertainty, and to prioritise efforts to develop good estimates of these key uncertainties (see Chapter 4, Methodological Choice and Identification of Key categories).

Step 3 typically involves; identification of candidate parametric PDFs to fit to the data, estimation of the parameters of such distributions, and evaluation of goodness-of-fit (e.g., Cullen and Frey, 1999). Rigorous methods can be applied to data sets that contain values below the detection limit of a measurement method, called *non-detects* (e.g., Zhao and Frey, 2004a). Distributions can be used in combination even when the data contain two or more subgroups that cannot otherwise be separated (e.g., Zheng and Frey, 2004).

3.2.2.3 METHODS FOR ENCODING EXPERT JUDGEMENTS

When empirical data are lacking or are not considered fully representative for all causes of uncertainty (Table 3.1), expert judgement may be necessary for estimating uncertainty. This section focuses on methods for encoding (quantifying) expert judgement regarding uncertainty in the form of PDFs. Encoding is the process of converting an expert's judgement regarding uncertainty into a quantitative PDF. Chapter 2 provides guidance on the definition of an expert, considerations in choosing expert(s), sources of possible bias in expert judgement and how to avoid them, and a recommended protocol for expert elicitation. In the context of uncertainties, a key goal of expert elicitation is to characterise the state of knowledge regarding possible values of a particular variable. Therefore, it is neither

necessary nor desirable to attempt to force consensus among experts; rather, it is more useful to take into account the full range of values when obtaining judgements from two or more experts for the same variable.

The goal of the process of eliciting (obtaining) expert judgement is to develop a PDF taking into account relevant information such as:

- Is the category similar to others? How is the uncertainty likely to compare?
- How well is the emission or removal process understood? Have all possible sources or sinks been identified?
- Are there physical limits on how much the emission factor or other estimation parameter can vary? Mass balance considerations or other process data may place an upper limit on emissions or removal rates.
- Are the emissions and removal estimates consistent with independent data that might be used to help verify the inventory?

A key concern with expert elicitation is to overcome the typical heuristic biases of *availability*, *representativeness*, and *anchoring and adjustment* (as described in Chapter 2, Annex 2A.1, Protocol for Expert Elicitation) to avoid the potential problem of obtaining an 'overconfident' estimate of uncertainty. 'Overconfidence' refers to a situation in which the estimated range of uncertainty is too narrow. It is desirable to avoid overconfidence so as not to underestimate the true uncertainty. It is *good practice* to use a formal elicitation protocol, such as the Stanford/SRI protocol that is detailed in Chapter 2, Annex 2A.1. In particular, these protocols include several steps prior to the actual encoding step for the purpose of familiarising the expert with the purpose and methods of the elicitation and encouraging the expert to think about all relevant data, models, theories, and other inference approaches. With this background, the expert is in a better position to make an unbiased estimate of uncertainty.

The method to be used for encoding should depend upon the expert's familiarity with PDFs. Some commonly used methods are:

- Fixed Value: Estimate the probability of being higher (or lower) than an arbitrary value and repeat, typically three or five times. For example, what is the probability that an emission factor would be less than 100?
- Fixed Probability: Estimate the value associated with a specified probability of being higher (or lower). For example, what is the emission factor such that there is only a 2.5 percent probability (or 1 in 40 chance) that the emission factor could be lower (or higher) than that value.
- Interval Methods: This method focuses on the median and the quartiles. For example, the expert would be asked to choose a value of the emission factor such that it is equally likely that the true emission factor would be higher or lower than that value. This yields the median. Then the expert would divide the lower range into two bins such that he or she felt it to be equally likely (25 percent probability) that the emission factor could be in either bin, and this would be repeated for the other end of the distribution. Finally, either fixed probability or fixed value methods could be used to get judgements for extreme values.
- Graphing: The expert draws his/her own distributions. This should be used cautiously because some experts are overconfident about their knowledge of PDFs.

An example of an expert elicitation that results in encoding (quantification) of a PDF is given in Box 3.1.

Sometimes the only available expert judgement will consist of a range, perhaps quoted together with a most likely value. Under these circumstances the following rules are considered *good practice*:

- Where experts only provide an upper and a lower value, assume that the probability density function is uniform and that the range corresponds to the 95 percent confidence interval.
- Where experts also provide a most likely value (which is often likely to be the same as the point estimate used in developing the inventory), assume a triangular probability density function using the most likely values as the mode and assume that the upper and lower values each exclude 2.5 percent of the population. The distribution needs not to be symmetrical. Other reasonable distribution choices, such as a normal or lognormal distribution, can be made given appropriate justifications.

Some other sources of information on expert elicitation include Spetzler and von Holstein (1975), Morgan and Henrion (1990), Merkhofer (1987), Hora and Iman (1989), and NCRP (1996).

The subjective nature of expert judgement increases the need for quality assurance and quality control procedures to improve comparability of uncertainty estimates between countries. Therefore expert judgements should be documented as part of the national archiving process, and inventory compilers are encouraged to apply QA/QC procedures to expert judgements, particularly for *key categories* (see Chapter 6).

Documentation requirements for expert judgement are discussed in Annex 2A.1 of Chapter 2.

BOX 3.1
A BRIEF EXAMPLE OF DETAILED EXPERT JUDGEMENT

Suppose that the inventory compiler has identified an expert for emissions of CH_4 from power plants and wishes to obtain his/her judgement regarding the uncertainty in annual average emissions for this category. As part of the motivation step, the elicitor has explained to the expert the general purpose of the analysis and the expert elicitation protocol to be used. In the structuring step, the elicitor works with the expert to set up the specific elicitation protocol. For example, although all the inventory compiler may want is an annual average uncertainty estimate, the expert may tell the elicitor that he/she prefers to provide judgements separately for start-up, part load, and full load operation of the plant, and that these three judgements should be weighted in order to come up with the combined uncertainty for an annual average. After structuring the problem, the elicitor then reviews the expert information relevant to the assessment, such as measurements that may have been made on similar types of power plants or other combustion sources. In the elicitation step, the elicitor might ask the expert for an upper value such that there is only a one in 40 chance (2.5 percent probability) of obtaining a higher value. After getting the value, the elicitor asks the expert to explain the logical basis for this estimate, such as the scenario of operation at the plant that might lead to such a high emission rate. Then the process might be repeated for the lower end of the range, and perhaps for the median, 25th percentile, and 75th percentile. A mixture of fixed value and fixed probability questions might be used. The elicitor should plot these on a graph so that any inconsistencies can be identified and corrected during the time available with the expert. In the verification step, the elicitor would make sure that the expert is comfortable that their judgement has been well represented. The elicitor might also see how the expert would react to the possibility of values outside of the interval for which judgements were provided, so as to ensure that the expert is not being overconfident.

3.2.2.4 GOOD PRACTICE GUIDANCE FOR SELECTING PROBABILITY DENSITY FUNCTIONS

Prior to selecting a PDF, it is *good practice* to account for biases in the data to the extent possible. As noted previously, data collection and QA/QC procedures can assist in preventing or correcting biases. For example, if national statistics on timber harvest exist, but it is also suggested that these statistics have a bias of 5 percent, then the mean estimate can be adjusted by 5 percent prior to estimating the random component of the uncertainty. It is *good practice* that adjustments for bias should be done in developing the point estimate emission inventory. Another consideration is that the amount of bias can change over time as data measurement or collection procedures change, or as the geographic and temporal scope of data collection changes. Thus, the bias corrections may be different for different years.

However, to the extent that biases are believed or known to exist in data even after QA/QC procedures have been applied, then either empirical or judgment based techniques can be applied to account for the bias. Apparent biases can arise in probabilistic analysis for at least two reasons: (1) a fitted distribution may have a mean that is different from the most likely value used in the point estimate of the inventory (e.g., a skewed triangular distribution based on expert judgment); and (2) the mean value of a prediction from a nonlinear model that has uncertain inputs can be different from the point estimate obtained from the same model if only point estimates of the mean values of the inputs are used. Thus, there are some types of biases that may be revealed only after an uncertainty analysis has been done.

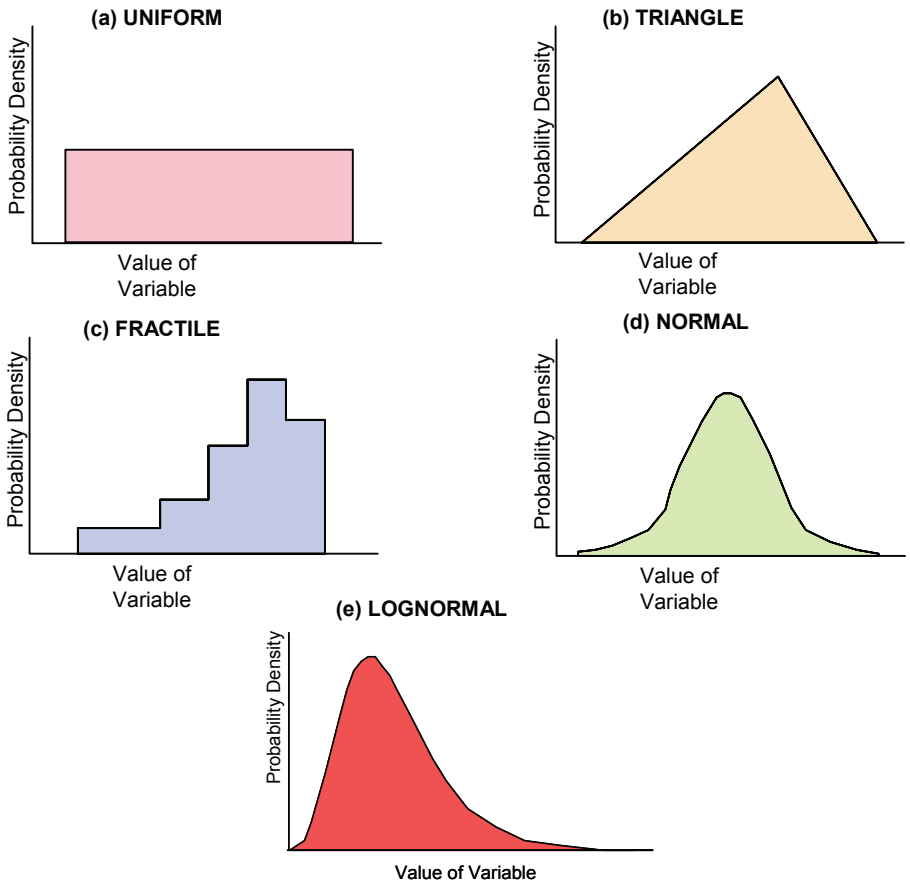
TYPES OF PROBABILITY DENSITY FUNCTIONS

There are many PDFs outlined in the statistical literature that often represent particular real situations. The choice of a particular type of PDF depends, at least in part, on the domain of the function (e.g., can it have both positive or negative values, or only non-negative values), the range of the function (e.g., is the range narrow or does it cover orders-of-magnitude), the shape (e.g., symmetry), and processes that generated the data (e.g., additive, multiplicative). These considerations are elaborated below in a brief discussion of many commonly used distributions of practical importance. Examples of such functions and the situations they represent are[4]:

[4] Further information on methods for developing distributions based upon statistical analysis of data are described and illustrated by Cullen and Frey (1999). Other useful references include Hahn and Shapiro (1967), Ang and Tang (1975) D'Agostino and Stephens (1986), Morgan and Henrion (1990), and U.S.EPA (1996, 1997, 1999). Some examples of probabilistic analyses applied to emission inventories are given by Frey and Zheng (2002) and Frey and Zhao (2004).

- *The normal distribution* is most appropriate when the range of uncertainty is small, and symmetric relative to the mean. The normal distribution arises in situations where many individual inputs contribute to an overall uncertainty, and in which none of the individual uncertainties dominates the total uncertainty. Similarly, if an inventory is the sum of uncertainties of many individual categories, however, none of which dominates the total uncertainty, then the overall uncertainty is likely to be normal. A normality assumption is often appropriate for many categories for which the relative range of uncertainty is small, e.g., fossil fuel emission factors and activity data.
- *The lognormal distribution* may be appropriate when uncertainties are large for a non-negative variable and known to be positively skewed. The emission factor for nitrous oxide from fertiliser applied to soil provides a typical inventory example. If many uncertain variables are multiplied, the product asymptotically approaches lognormality. Because concentrations are the result of mixing processes, which are in turn multiplicative, concentration data tend to be distributed similar to a lognormal. However, real-world data may not be as tail-heavy as a lognormal distribution. The Weibull and Gamma distributions have approximately similar properties to the lognormal but are less tail-heavy and, therefore, are sometimes a better fit to data than the lognormal.
- *Uniform distribution* describes an equal likelihood of obtaining any value within a range. Sometimes the uniform distribution is useful for representing physically-bounded quantities (e.g., a fraction that must vary between 0 and 1) or for representing expert judgement when an expert is able to specify an upper and lower bound. The uniform distribution is a special case of the Beta distribution.
- *The triangular distribution* is appropriate where upper and lower limits and a preferred value are provided by experts but there is no other information about the PDF. The triangular distribution can be asymmetrical.
- *Fractile distribution* is a type of empirical distribution in which judgements are made regarding the relative likelihood of different ranges of values for a variable, such as illustrated in Figure 3.5. This type of distribution is sometimes useful in representing expert judgement regarding uncertainty.

Figure 3.5 Examples of some commonly used probability density function models (e.g., based on Frey and Rubin, 1991)



ISSUES TO CONSIDER WHEN DEVELOPING THE PROBABILITY DENSITY FUNCTION

The following describes how inventory compilers can satisfy the principles of comparability, consistency and transparency in emissions inventories when selecting a PDF:

- Where empirical data are available, the first consideration should be whether a normal distribution would be appropriate as a representation of uncertainty. If the variable must be non-negative, then the standard deviation of the normal distribution should not exceed 30 percent of the mean value to avoid an unacceptably high probability of erroneously predicting negative values. Truncation of the lower tail of the normal distribution should generally be avoided, because it changes the mean and other statistics of the distribution. Typically, a better alternative to truncation is to find a more appropriate distribution that is a better fit to the data. For example, for positively skewed data that must be non-negative, lognormal, Weibull, or Gamma distributions often can provide an acceptable fit; however, an empirical distribution of the data can also be used;
- Where expert judgement is used, the distribution function adopted might typically be normal or lognormal, supplemented by uniform, triangular, or fractile distributions as appropriate;
- Other distributions may be used where there are compelling reasons, either from empirical observations or from expert judgement supported by theoretical argument.

The issue of identifying which function best fits a set of data can be difficult. One approach is to use the square of the skewness and the kurtosis to look for functional forms that can fit the data (Cullen and Frey, 1999). Kurtosis and skewness should only be applied if there are sufficient data from which to estimate these values. The function is then fitted to the data by least squares fit or other means. Tests are available to assess the goodness-of-fit, including the chi-squared test and others (Cullen and Frey, 1999). In many cases, several functions will fit the data satisfactorily within a given probability limit. These different functions can have radically different distributions at the extremes where there are few or no data to constrain them, and the choice of one function over another can systematically change the outcome of an uncertainty analysis. Cullen and Frey (1999) reiterate the advice of previous authors in these cases that *it must be knowledge of the underlying physical processes that governs the choice of a probability function*. What the tests provide, in the light of this physical knowledge, is guidance on whether this function does or does not satisfactorily fit the data.

In order to use empirical data as a basis for developing PDFs, the first critical step is to determine if the data are a random, representative sample, in the case of a sample from a population. Some key questions to ask regarding the data include:

- Are the data representative of the conditions pertaining to the emission or activity factors specific to national circumstances? For example, in AFOLU, are the data representative of management practices and other national circumstances?
- Are the data a random sample?
- What is the averaging time associated with the data set, and is it the same as for the assessment (which will be for annual emissions in a given year)? For example, emissions data might be measured during a short time period and not for an entire year. Thus, expert judgment may be required in order to extrapolate short term data to a longer term basis.

If the data are a random, representative sample, then the distribution can be established directly using classical statistical techniques, even if the sample size is small. Ideally the available data will represent an annual average but may be necessary to convert data using an appropriate averaging time. For normal distributions the 95 percent confidence interval would be plus or minus twice the estimated standard deviation of the population. In other cases, the data may represent an exhaustive census of the sum of all activity (e.g., total energy use for a particular fuel). In this case, information regarding errors in the measurements or survey instruments would form a basis for assessing uncertainty. The range of uncertainty of activity data might be bounded by using independent methods or consistency checks. For example, fuel consumption data can be compared with estimates of production, including estimates of production via different methods.

There is a distinction between uncertainty in the mean and variability in the data for situations in which the data represent intra-country variability within a category. Since the goal is to estimate annual average emissions at the level of an individual country, data that represent intra-country variability should be averaged over the entire geographic area of the country, and uncertainty in this average should be assessed and used as the basis for the inventory. Conversely, if international data are available at an aggregate level, without supporting details as to how such data can be disaggregated by country, there is a mismatch in scale that is more difficult to correct. Typically, in this case, the uncertainty will tend to increase as the geographic scope decreases, i.e., if the number of categories included decreases and if site-specific emissions data are not available. Thus uncertainty ranges

that are developed for aggregated international data may have to be widened for applicability to individual countries. In the absence of any empirical basis for estimating the relative range of uncertainty at the country level versus the aggregated international level, expert judgement can be used.

For a sample of an underlying population, the need is to evaluate whether the data are random and representative of the population. If so, classical statistical methods can be used to define the distribution. If not, then some combination of data analysis and elicitation of expert judgement regarding distributions will be required. In the former case, Cullen and Frey (1999) suggest exploration of the data set using summary statistics and graphics to evaluate essential features (e.g., central tendency, range of variation, skewness). The insights obtained by examining the data, combined with knowledge of the processes that generated the data, should be considered when selecting a mathematical or numerical representation of the distribution for input into Approaches 1 or 2. (See Section 3.2.3.)

If a parametric distribution is selected as a candidate for fitting to the data set, techniques such as 'maximum likelihood estimation[5]' or the 'method of matching moments[6]' can be used to estimate the parameters of the distribution. The goodness-of-fit of the distribution can be evaluated in numerous ways, including comparison of the fitted cumulative distribution function (CDF) with the original data set, probability plots, and goodness-of-fit tests (e.g., Cullen and Frey, 1999). It is important that the selection of a parametric distribution to represent a data set should be based not solely upon goodness-of-fit tests, but upon similarities in processes that generated the data versus the theoretical basis for a distribution (e.g., Hahn and Shapiro, 1967).

If the data are averaged over less than one year, it may be necessary to extrapolate the uncertainty over the year. Consider an example in which the data set represents variability in daily average emissions measurements for a particular category. One approach, described in detail by Frey and Rhodes (1996), is to fit a parametric distribution to the data set for daily variability, use a numerical technique known as bootstrap simulation to estimate uncertainty in the parameters of the distribution, and use Monte Carlo simulation to simulate randomised annual averages of the emission factor. Using bootstrap simulation, the uncertainty in the sampling distribution for the parameters of the fitted distribution can be simulated (e.g., Efron and Tibshirani, 1993; Frey and Rhodes, 1996; Frey and Bammi, 2002).

DEPENDENCE AND CORRELATION AMONG INPUTS

This section provides a brief overview of issues pertaining to dependence and correlation among inputs. More details on this topic can be found in Morgan and Henrion (1990), Cullen and Frey (1999), and Smith *et al*. (1992).

When setting up a probabilistic analysis it is preferable to define the model so that the inputs are as statistically independent as possible. For example, rather than to try to estimate activity data for many subcategories for which data are derived at least in part by differences, it may be better to assign uncertainties to better known aggregate measures of activity. For example, residential fuel use might be estimated as the difference between total consumption and usage in the transportation, industrial, and commercial sectors. In this case, the estimate of uncertainty in residential fuel use is negatively correlated with the uncertainties in fuel use in the other subcategories, and may even be very large compared to the uncertainty in total consumption. Therefore, rather than trying to estimate uncertainties separately for each subcategory, it may be more practical to estimate uncertainty for aggregated categories, for which good estimates and cross-checks may be available.

Dependencies, if they exist, may not always be important to the assessment of uncertainties. Dependencies among inputs will matter only if the dependencies exist between two inputs to which the uncertainty in the inventory is sensitive and if the dependencies are sufficiently strong. In contrast, weak dependencies among inputs, or strong dependencies among inputs to which the uncertainty in the inventory is insensitive, will be of relatively little consequence to the analysis. Of course, some interdependencies are important and failure to account for those relationships can lead to misleading results. Positive correlations between inputs tend to increase the range of uncertainty in the output, whereas negative correlations tend to decrease the range of uncertainty in the output. However, positive correlations in uncertainties when comparing two years as part of trend analysis will decrease uncertainty in the trend.

Techniques can be considered for incorporating dependencies into the analysis including:

- stratifying or aggregating the categories to minimise the effect of the dependencies;

[5] The method of maximum likelihood selects as estimates the values of the parameters that maximise the likelihood of the observed sample (e.g., Holland and Fitz-Simons, 1982).

[6] The method of moments finds estimators of unknown parameters by equating corresponding sample and population moments. The method is easy to employ and provides consistent estimators. In many cases, the method of moments estimators are biased (Wackerly, Mendenhall III and Scheaffer, 1996; pp. 395-397).

- modelling the dependence explicitly;
- simulating correlation using restricted pairing methods (that are included in many software packages);
- use of resampling techniques in cases where multivariate datasets are available;
- considering bounding or sensitivity cases (e.g., one case assuming independence and another case assuming complete positive correlation); and
- time series techniques can be used to analyse or simulate temporal autocorrelation.

As a simple example, Zhao and Frey (2004a) evaluated the implications of whether or not emission factor uncertainty estimates for different categories obtained from the same data source should be considered as dependent or independent among the categories, and found that it did not matter to the overall inventory uncertainty. Of course, this result is specific to the particular case studies and should be tested in other applications. As a more complex example, given in Box 3.2, Ogle *et al*. (2003) accounted for dependencies in tillage management factors, which were estimated from a common set of data in a single regression-type model, by determining the covariance[7] between factors for reduced tillage and no-till management, and then using that information to generate tillage factor values with appropriate correlation during a Monte Carlo simulation[8]. One should consider the potential for correlations among input variables and focus on those that would be likely to have the largest dependencies (e.g., applying management factors for the same practice in different years of an inventory, or correlations among management activities from one year to the next).

BOX 3.2
EXAMPLE OF MONTE CARLO UNCERTAINTY ASSESSMENT DEALING WITH CORRELATIONS

Ogle *et al*. (2003) performed a Monte Carlo analysis to assess uncertainty in a Tier 2 inventory that addressed changes in soil C attributed to land use and management of agricultural lands in the United States. Management factors were estimated from about 75 published studies using linear mixed effect models. PDFs were derived for the management effect at a depth of 30 cm following 20 years after its implementation. Reference stocks were estimated using a National Soil Survey Characterisation Database, which contained pedon data collected by United States Department of Agriculture (USDA). PDFs were based on the mean and variance from about 3700 pedons, taking into account the spatial autocorrelation of pedon locations due to clumped distribution patterns. The land use and management activity data were recorded in the USDA National Resources Inventory, which tracks agricultural land management at more than 400,000 point locations in the United States, along with supplemental data on tillage practices provided by the Conservation Technology and Information Center (CTIC). The Monte Carlo analysis was implemented using a commercially available software package and code developed by U.S. analysts. Their analysis accounted for dependencies between estimation parameters that were derived from common datasets. For example factors for set-aside lands and land use change between cultivated and uncultivated conditions were derived from a single regression analysis using an indicator variable for set-asides, and hence were interdependent. Their analysis also accounted for dependencies in the land use and management activity data. When simulating input values, factors were considered completely dependent from the base and current year in the inventory because the relative influence of management on soil C was assumed to be the same regardless of the year when a practice was implemented. As such, factors were simulated with identical random seed values. In contrast, reference carbon stocks for the various soil types in each climate region were simulated independently, with different random seeds, because stocks for each region were constructed from separate independent sets of data. U.S. analysts chose to use 50,000 iterations for their Monte Carlo analysis. This was satisfactory because they were only reporting one digit after the decimal, and simulation results were considered relatively stable at that level of significance. Ogle *et al*. (2003) estimated that mineral soils gained an average of 10.8 Tg C yr^{-1} between 1982 and 1997, with a 95 percent confidence interval ranging from 6.5 to 15.3 Tg C yr^{-1}. In contrast, managed organic soils lost an average of 9.4 Tg C yr^{-1}, ranging from 6.4 to 13.3 Tg C yr^{-1}. Further, Ogle *et al*. (2003) found that the variability in management factors contributed 90 percent of the overall uncertainty for the final estimates of soil carbon change.

[7] The covariance between two variables (x and y) measures the mutual dependence between them. The covariance of a sample consisting of n pairs of values is the total of the products of the deviation of individual x values from the mean x value times the deviation of the corresponding individual y value from the mean of the y values, divided by (n-1).

[8] More discussion and examples of these types of methods are given in Cullen and Frey (1999), Morgan and Henrion (1990), and USEPA (1996). These documents also contain reference lists with citations to relevant literature.

3.2.3 Methods to combine uncertainties

Once the uncertainties in activity data, emission factor or emissions for a category have been determined, they may be combined to provide uncertainty estimates for the entire inventory in any year and the uncertainty in the overall inventory trend over time. Results from sampling theory, as described in Section 2.5.1, Measurement-Based Tier 3 Inventories, of Chapter 2 in Volume 4 for the AFOLU Sector, may be used in cases where sampling is applied for direct measurement of, e.g., carbon stock changes. In these situations, sampling theory provides an estimate of the uncertainty in emissions/removals for a given category, without need to separately characterise an activity and emission factor.

Two approaches for the estimation of combined uncertainties are presented in the following sections: Approach 1 uses simple error propagation equations, while Approach 2 uses Monte Carlo or similar techniques. Either Approach may be used for emission sources or sinks, subject to the assumptions and limitations of each Approach and availability of resources. Complementary step by step explanation of the statistical calculation methods of Approaches is given in Sections 3.7.1 and 3.7.2.

Biases should be addressed prior to applying either Approach 1 or 2, as set out in Section 3.2.2.1. For example, as discussed in Section 3.2.2.1, an assessment of bias, and potential disagreements among modelling approaches, should be conducted, and any action identified to improve the inventory estimate should be taken. Approaches 1 and 2 focus on quantifying the random component of the uncertainty of the inventory results where known sources of bias have been removed. The inventory estimates may still include unknown bias and in the analysis all errors are assumed behaving as random (Winiwarter and Rypdal, 2001).

3.2.3.1 APPROACH 1: PROPAGATION OF ERROR

Approach 1 is based upon error propagation and is used to estimate uncertainty in individual categories, in the inventory as a whole, and in trends between a year of interest and a base year. The key assumptions, requirements, and procedures are described here.

Approach 1 should be implemented using Table 3.2, Approach 1 Uncertainty Calculation, that can be set up on commercial spreadsheet software. The table is completed at the category level using uncertainty ranges for activity data and emission factors consistent with the sectoral *good practice guidance*[9]. Different gases should be entered separately as CO_2 equivalents.

KEY ASSUMPTIONS OF APPROACH 1

In Approach 1 uncertainty in emissions or removals can be propagated from uncertainties in the activity data, emission factor and other estimation parameters through the error propagation equation (Mandel, 1984, Bevington and Robinson, 1992). If correlations exist, then either the correlation can be included explicitly or data can be aggregated to an appropriate level such that correlations become less important. Approach 1 also theoretically requires that the standard deviation divided by the mean value is less than 0.3. In practice, however, the approach will give informative results even if this criterion is not strictly met and some correlations remain. Approach 1 assumes that the relative ranges of uncertainty in the emission and activity factors are the same in the base year and in year *t*. This assumption is often correct or approximately correct. If any of the key assumptions of Approach 1 do not apply, then either an alternative version of Approach 1 can be developed (e.g., see Section 3.4) or Approach 2 can be used instead.

Where the standard deviation divided by the mean is greater than 0.3 the reliability of Approach 1 can be improved. The section 'Dealing with Large and Asymmetric Uncertainties in the Results of Approach 1' in this section describes how to do this.

KEY REQUIREMENTS OF APPROACH 1

In order to quantify uncertainty using Approach 1, estimates of the mean and the standard deviation for each input are required, as well as the equation through which all inputs are combined to estimate an output. The simplest equations include statistically independent (uncorrelated) inputs.

[9] Where estimates are derived from models, enter the uncertainty associated with the activity data used to drive the model, and enter the uncertainty associated with the model parameters instead of the emission factor uncertainty. It may be necessary to use expert judgement, or error propagation calculations associated with the model structure. If it is impractical to separate the uncertainty estimate obtained from a model for a category into separate activity and emission factor components, then enter the total uncertainty for the category in the emission factor column and assign zero uncertainty to the activity factor column.

Once the uncertainties in the categories have been determined, they may be combined to provide uncertainty estimates for the entire inventory in any year and the uncertainty in the overall inventory trend over time. As discussed further below, these uncertainty estimates can be combined using two convenient rules for combining uncorrelated uncertainties under addition and multiplication.

PROCEDURE OF APPROACH 1

The Approach 1 analysis estimates uncertainties by using the error propagation equation in two steps. First, the Equation 3.1 approximation is used to combine emission factor, activity data and other estimation parameter ranges by category and greenhouse gas. Second, the Equation 3.2 approximation is used to arrive at the overall uncertainty in national emissions and the trend in national emissions between the base year and the current year.

Uncertainty of an Annual Estimate

The error propagation equation[10] yields two convenient rules for combining uncorrelated uncertainties under addition and multiplication:

- Where uncertain quantities are to be combined by multiplication, the standard deviation of the sum will be the square root of the sum of the squares of the standard deviations of the quantities that are added, with the standard deviations all expressed as coefficients of variation, which are the ratios of the standard deviations to the appropriate mean values. This rule is approximate for all random variables. Under typical circumstances this rule is reasonably accurate as long as the coefficient of variation is less than approximately 0.3. This rule is not applicable to division.

  A simple equation (Equation 3.1) can then be derived for the uncertainty of the product, expressed in percentage terms:

EQUATION 3.1
COMBINING UNCERTAINTIES – APPROACH 1 – MULTIPLICATION

$$U_{total} = \sqrt{U_1^2 + U_2^2 + \ldots + U_n^2}$$

Where:

U_{total} = the percentage uncertainty in the product of the quantities (half the 95 percent confidence interval divided by the total and expressed as a percentage);

U_i = the percentage uncertainties associated with each of the quantities.

- Where uncertain quantities are to be combined by addition or subtraction, the standard deviation of the sum will be the square root of the sum of the squares of the standard deviations of the quantities that are added with the standard deviations all expressed in absolute terms (this rule is exact for uncorrelated variables).

  Using this interpretation, a simple equation (Equation 3.2) can be derived for the uncertainty of the sum, expressed in percentage terms:

EQUATION 3.2
COMBINING UNCERTAINTIES – APPROACH 1 – ADDITION AND SUBTRACTION

$$U_{total} = \frac{\sqrt{(U_1 \bullet x_1)^2 + (U_2 \bullet x_2)^2 + \ldots + (U_n \bullet x_n)^2}}{\left|x_1 + x_2 + \ldots + x_n\right|}$$

Where:

U_{total} = the percentage uncertainty in the sum of the quantities (half the 95 percent confidence interval divided by the total (i.e., mean) and expressed as a percentage). This term 'uncertainty' is thus based upon the 95 percent confidence interval;

x_i and U_i = the uncertain quantities and the percentage uncertainties associated with them, respectively.

The greenhouse gas inventory is principally the sum of products of emission factors, activity data and other estimation parameters. Therefore, Equations 3.1 and 3.2 can be used repeatedly to estimate the uncertainty of the

[10] As discussed more extensively in Annex 1 of the *Good Practice Guidance and Uncertainty Management* (*GPG2000*, IPCC, 2000), and in Annex I of the *Revised 1996 IPCC Guidelines* (Reporting Instructions) (*1996 IPCC Guidelines*, IPCC, 1997).

total inventory. In practice, uncertainties found in inventory categories vary from a few percent to orders of magnitude, and may be correlated. This is not consistent with the assumptions of Equations 3.1 and 3.2 that the variables are uncorrelated, and with the assumption of Equation 3.2 that the coefficient of variation is less than about 30 percent, but under these circumstances, Equations 3.1 and 3.2 may still be used to obtain an approximate result.

Uncertainty in the Trend

Trend uncertainties are estimated using two sensitivities:

- *Type A sensitivity*: the change in the difference in overall emissions between the base year and the current year, expressed as a percentage, resulting from a 1 percent increase in emissions or removals of a given category and gas in both the base year and the current year.
- *Type B sensitivity*: the change in the difference in overall emissions between the base year and the current year, expressed as a percentage, resulting from a 1 percent increase in emissions or removals of a given category and gas in the current year only.

The Type A and Type B sensitivities are merely intermediate variables that simplify the calculation procedure. The results of the analysis are not constrained to a change of only one percent, but instead depend upon the range of uncertainty for each category.

Conceptually, Type A sensitivity arises from uncertainties that affect emissions or removals in the base year and the current year equally, and Type B sensitivity arises from uncertainties that affect emissions or removals in the current year only. Uncertainties that are fully correlated between years will be associated with Type A sensitivities, and uncertainties that are not correlated between years will be associated with Type B sensitivities. Emission factor (and other estimation parameters) uncertainties will tend to have Type A sensitivities, and activity data uncertainties will tend to have Type B. However, this association will not always hold and it is possible to apply Type A sensitivities to activity data, and Type B sensitivities to emission factors to reflect particular national circumstances. Type A and Type B sensitivities are simplifications introduced for the approximate analysis of correlation.

Once the uncertainties introduced into the national inventory by Type A and Type B sensitivities have been calculated, they can be summed using the error propagation equation (Equation 3.1) to give the overall uncertainty in the trend.

Worksheet for Approach 1 Uncertainty Calculation

The columns of Table 3.2, Approach 1 Uncertainty Calculation, are labelled A to M and contain the following information, of which the derivation of key equations is given in Section 3.7.1 in Section 3.7, Technical Background Information.

- A and B show the IPCC category and greenhouse gas.
- C and D are the inventory estimates in the base year and the current year[11] respectively, for the category and gas specified in Columns A and B, expressed in CO_2 equivalents.
- E and F contain the uncertainties for the activity data and emission factors respectively, derived from a mixture of empirical data and expert judgement as previously described in this chapter, entered as half the 95 percent confidence interval divided by the mean and expressed as a percentage. The reason for halving the 95 percent confidence interval is that the value entered in Columns E and F corresponds to the familiar plus or minus value when uncertainties are loosely quoted as 'plus or minus *x* percent', so expert judgements of this type can be directly entered in the spreadsheet. If uncertainty is known to be highly asymmetrical, enter the larger percentage difference between the mean and the confidence limit.
- G is the combined uncertainty by category derived from the data in Columns E and F using the error propagation equation (Equation 3.2). The entry in Column G is therefore the square root of the sum of the squares of the entries in Columns E and F.
- H shows the uncertainty in Column G as a percentage of total national emissions in the current year. The entry in each row of Column H is the square of the entry in Column G multiplied by the square of the entry in Column D, divided by the square of total at the foot of Column D. The value at the foot of Column H is an estimate of the percentage uncertainty in total national net emissions in the current year, calculated from the entries above using Equation 3.1. This total is obtained by summing the entries in Column H and taking the square root.

[11] The current year is the most recent year for which inventory data are available.

- I shows how the percentage difference in emissions between the base year and the current year changes in response to a one percent increase in category emissions/removals for both the base year and the current year. This shows the sensitivity of the trend in emissions to a systematic uncertainty in the estimate (i.e., one that is correlated between the base year and the current year). This is the Type A sensitivity as defined above.
- J shows how the percentage difference in emissions between the base year and the current year changes in response to a one percent increase in category emissions/removals in the current year only. This shows the sensitivity of the trend in emissions to random error in the estimate (i.e., one, that is not correlated, between the base year and the current year). This is the Type B sensitivity as described above.
- K uses the information in Columns I and F to show the uncertainty introduced into the trend in emissions by emission factor uncertainty, under the assumption that uncertainty in emission factors is correlated between years. If the user decides that the emission factor uncertainties are not correlated between years then the entry in Column J should be used in place of that in Column I and the result multiplied by $\sqrt{2}$.
- L uses the information in Columns J and E to show the uncertainty introduced into the trend in emissions by activity data uncertainty, under the assumption that uncertainty in activity data is not correlated between years. If the user decides that the activity data uncertainties are correlated between years then the entry in Column I should be used in place of that in Column J and the $\sqrt{2}$ factor does not then apply.
- M is an estimate of the uncertainty introduced into the trend in national emissions by the category in question. Under Approach 1, this is derived from the data in Columns K and L using Equation 3.2. The entry in Column M is therefore the sum of the squares of the entries in Columns K and L. The total at the foot of this column is an estimate of the total uncertainty in the trend, calculated from the entries above using the error propagation equation. This total is obtained by summing the entries in Column M and taking the square root. The uncertainty in the trend is a *percentage point* range relative to the inventory trend. For example, if the current year emissions are 10 percent greater than the base year emissions, and if the trend uncertainty at the foot of Column M is reported as 5 percent, then the trend uncertainty is 10%±5% (or from 5% to 15% increase) for the current year emissions relative to the base year emissions.
- Explanatory footnotes go at the bottom of the table and give documentary references of uncertainty data (including measured data) or other relevant comments.

An example of the spreadsheet with all the numerical data completed is provided in Section 3.6, Approach 1 uncertainty calculation example. Details of this approach are given in Section 3.7.1 and derivation of the uncertainty in the trend is in Section 3.7.2.

TABLE 3.2
APPROACH 1 UNCERTAINTY CALCULATION

A	B	C	D	E	F	G	H	I	J	K	L	M
IPCC category	Gas	Base year emissions or removals	Year *t* emissions or removals	Activity data uncertainty	Emission factor / estimation parameter uncertainty	Combined uncertainty	Contribution to Variance by Category in Year *t*	Type A sensitivity	Type B sensitivity	Uncertainty in trend in national emissions introduced by emission factor / estimation parameter uncertainty	Uncertainty in trend in national emissions introduced by activity data uncertainty	Uncertainty introduced into the trend in total national emissions
		Input data	Input data	Input data Note A	Input data Note A	$\sqrt{E^2 + F^2}$	$\frac{(G \bullet D)^2}{(\sum D)^2}$	Note B	$\left\lvert \frac{D}{\sum C} \right\rvert$	$I \bullet F$ Note C	$J \bullet E \bullet \sqrt{2}$ Note D	$K^2 + L^2$
		Gg CO_2 equivalent	Gg CO_2 equivalent	%	%	%		%	%	%	%	%
E.g., 1.A.1. Energy Industries Fuel 1	CO_2											
E.g., 1.A.1. Energy Industries Fuel 2	CO_2											
Etc...	…											
Total		$\sum C$	$\sum D$				$\sum H$					$\sum M$
					Percentage uncertainty in total inventory:		$\sqrt{\sum H}$				Trend uncertainty:	$\sqrt{\sum M}$

Note A: If only total uncertainty is known for a category (not for emission factor and activity data separately), then:

- If uncertainty is correlated across years, enter the uncertainty into Column F, and enter 0 in Column E;
- If uncertainty is not correlated across years, enter the uncertainty into Column E, and enter 0 in Column F

Note B:

$$\text{Absolute value of: } \frac{0.01 \bullet D_x + \sum D_i - \left(0.01 \bullet C_x + \sum C_i\right)}{\left(0.01 \bullet C_x + \sum C_i\right)} \bullet 100 - \frac{\sum D_i - \sum C_i}{\sum C_i} \bullet 100$$

Where:

C_x, D_x = entry from row x of the table from the corresponding column, representing a specific category

$\sum C_i$, $\sum D_i$ = Sum over all categories (rows) of the inventory of the corresponding column

Note C: In the case where no correlation between emission factors is assumed, sensitivity B should be used and the result multiplied by √2:

$$K_x = J_x \bullet F_x \bullet \sqrt{2}$$

Note D: In the case where correlation between activity data is assumed, sensitivity A should be used and the √2 is not required:

$$L_x = I_x \bullet E_x$$

DEALING WITH LARGE AND ASSYMMETRIC UNCERTAINTIES

Section 3.7.3 provides details on how the results from Approach 1 can be interpreted if the relative range of uncertainty is large for a quantity that must be non-negative. The error propagation method that is the basis for Approach 1 works well if the uncertainties are relatively small, meaning that the standard deviation divided by the mean is less than 0.3. If the uncertainties are larger, Approach 1 may continue to be used, providing informative results. However without any corrections, this approach will tend to underestimate uncertainty of the multiplicative (or quotient) terms. Furthermore, if the relative uncertainties are large for non-negative quantities, then the uncertainty ranges are typically asymmetric, and Approach 1 does not quantify such asymmetry. A second option is to use Approach 2, however this may not be always feasible. A third option is to use Approach 1 with corrections. For example, as discussed in more detail later in Section 3.7.3, an uncertainty of -65% to +126% relative to the mean might be estimated to be simply plus or minus 100 percent. This example can be properly addressed with some corrections to the results of Approach 1. The advantage of using the correction applied to Approach 1 (where applicable), rather than Approach 2, is that relatively simple spreadsheet-based calculation methods can be used and it is not necessary to use specialised Monte Carlo simulation software.

3.2.3.2 APPROACH 2: MONTE CARLO SIMULATION

The Monte Carlo analysis is suitable for detailed category-by-category assessment of uncertainty, particularly where uncertainties are large, distribution is non-normal, the algorithms are complex functions and/or there are correlations between some of the activity sets, emissions factors, or both.

In Monte Carlo simulation, pseudo-random samples of model inputs are generated according to the PDFs specified for each input. The samples are referred to as 'pseudo-random' because they are generated by an algorithm, referred to as a pseudo-random number generator (PRNG), that can provide a reproducible series of numbers (according to the random seeds assigned as input to the PRNG) but for which any series has properties of randomness. Details are available elsewhere (e.g., Barry, 1996). If the model has two or more inputs, then random samples are generated from the PDFs for each of the inputs, and one random value for each input is entered into the model to arrive at one estimate of the model output. This process is repeated over a desired number of iterations to arrive at multiple estimates of the model output. The multiple estimates are sample values of the PDF of the model output. By analyzing the samples of the PDF for the model output, the mean, standard deviation, 95 percent confidence interval, and other properties of the output PDF can be inferred. Because Monte

Carlo simulation is a numerical method, the precision of the results typically improves as the number of iterations is increased. More details regarding the methodology of Monte Carlo simulation, as well as regarding similar techniques such as Latin Hypercube sampling (LHS), are given by Hahn and Shapiro (1967); Ang and Tang (1984); and Morgan and Henrion (1990).

KEY ASSUMPTIONS OF APPROACH 2

Under Approach 2, the simplifying assumptions required for Approach 1 can be relaxed. Thus, numerical statistical techniques, particularly the Monte Carlo technique, as they can be generally applied, are more appropriate than Approach 1 for estimating uncertainty in emissions/removals (from uncertainties in activity measures and emission factors/estimation parameters) when:

- uncertainties are large;
- their distribution are non-Gaussian;
- algorithms are complex functions;
- correlations occur between some of the activity data sets, emission factors, or both;
- uncertainties are different for different years of the inventory.

KEY REQUIREMENTS OF APPROACH 2

Monte Carlo simulation requires the analyst to specify PDFs (see Fishman, 1996) that reasonably represent each model input for which the uncertainty is quantified. The PDFs may be obtained by a variety of methods, as described in Section 3.2.2.4 including statistical analysis of data or expert elicitation. A key consideration is to develop the distributions for the input variables to the emission/removal calculation model so that they are based upon consistent underlying assumptions regarding averaging time, location, and other conditioning factors relevant to the particular assessment (e.g., climatic conditions influencing agricultural greenhouse gas emissions).

Monte Carlo analysis can deal with probability density functions of any physically possible shape and width, as well as handling varying degrees of correlation (both in time and between source/sink categories). Monte Carlo analysis can deal with simple models (e.g., emission inventories that are the sum of sources and sinks, each of which is estimated using multiplicative factors) as well as more complex models (e.g., the first order decay for CH_4 from landfills).

PROCEDURES OF APPROACH 2

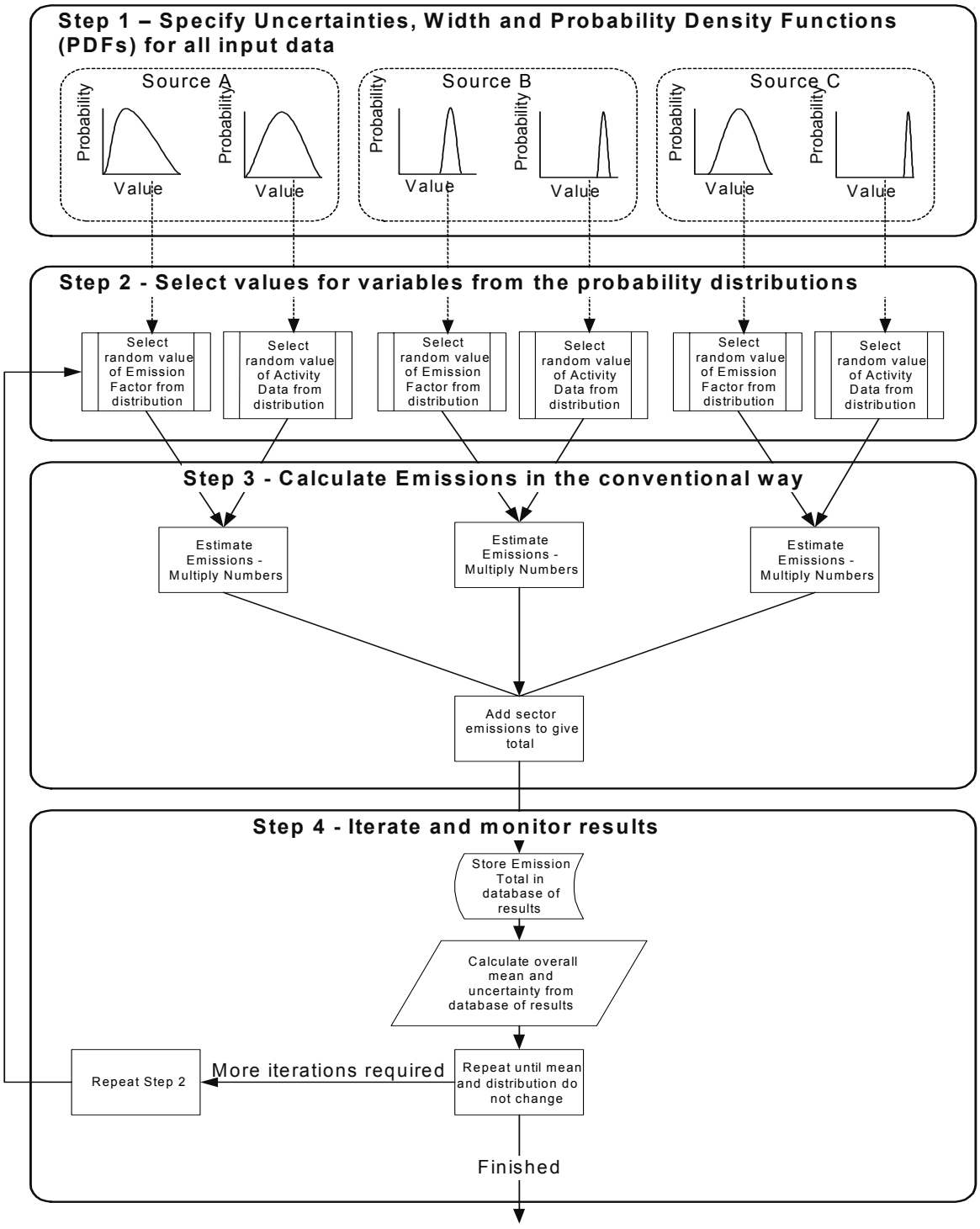
The principle of Monte Carlo analysis is to select random values of emission factor, activity data and other estimation parameters from within their individual probability density functions, and to calculate the corresponding emission values. This procedure is repeated many times, using a computer, and the results of each calculation run build up the overall emission probability density function. Monte Carlo analysis can be performed at the category level, for aggregations of categories or for the inventory as a whole. Statistical software packages are readily available – some of which include Monte Carlo algorithms that are very user-friendly[12].

Like all methods, Monte Carlo analysis only provides satisfactory results if it is properly implemented. This requires the analyst to have scientific and technical understanding of the inventory. Of course, the results will only be valid to the extent that the input data, including any expert judgements, are sound.

The Monte Carlo approach consists of four clearly defined steps shown in Figure 3.7. Only the first of these requires effort from the user, the remainder being handled by the software package. The emission inventory calculation, the PDFs, and the correlation values should be set up in the Monte Carlo package. The software performs the subsequent steps. In some cases, the inventory compiler may decide to set up its own programme to run a Monte Carlo simulation; this can be done using statistical software. The Section, 'Choosing a simulation technique and sample size' below contains a short discussion of various software packages.

[12] Winiwarter and Rypdal (2001), Eggleston *et al*. (1998) and Monni et al. (2004) provide examples of Monte Carlo analysis applied to national greenhouse gas inventories to estimate uncertainties both in overall emissions and emissions trends. Another example of the use of Monte Carlo analysis is given in McCann *et al*. (1994). More detailed descriptions and applications of this method are presented in Bevington and Robinson (1992), Manly (1997), Morgan and Henrion (1990), and Cullen and Frey (1999). A brief example of the application of Monte Carlo analysis is provided in Box 3.2 based on Ogle *et al*. (2003).

Figure 3.6 Illustration of Monte Carlo method



Step 1: Specify category uncertainties. This includes estimation parameters and activity data, their associated means and PDFs, and any correlations. The uncertainties can be assessed following the guidance in Sections 3.2.1 and 3.2.2. For guidance on assessment of correlations, see 'Dependence and correlation among inputs' in this section and Box 3.2.

Step 2: Select random variables. Select input values. Input values are the estimates applied in the inventory calculation. This is the start of the iterations. For each input data item, a number is randomly selected from the PDF of that variable.

Step 3: Estimate emissions and removals. The variables selected in Step 2 are used to estimate annual emissions and removals based on input values. Correlations of 100 percent are easy to incorporate, and good Monte Carlo packages allow other correlations to be included. Since the emission calculations should be the

same as those used to estimate the national inventory, the Monte Carlo process could be fully integrated into the annual emission estimates.

Step 4: Iterate and monitor results. Iterate and monitor results. The calculated total from Step 3 is stored, and the process then repeats from Step 2. The results from the repetitions are used to calculate the mean and the PDF.

APPROACH 2 UNCERTAINTIES IN TRENDS

The Approach 2 Monte Carlo method can be used to estimate uncertainties in the trend as well as in the absolute emission value in a given year. The procedure is a simple extension of that described in the previous section.

The trend is defined here as the percentage difference[13] between the base year and the year of interest (year *t*). Therefore, the Monte Carlo analysis needs to be set up to estimate both years simultaneously. The following steps show the procedure.

Step 1: Specify source/sink category uncertainties. Determine the probability density functions for emission factors, activity data and other estimation parameters. This is the same process as described above except that it needs to be done for both the base year and the current year, and relationships between the data need to be considered. For many categories, the same emission factor will be used for each year (i.e., the emission factors for both years are 100 percent correlated). In these cases, one distribution is described and the value selected from it is used for each year in step 3. Changes in the technologies or practices will alter the emission factor over time. In this case, two emission factors should be used, that have a lower or zero correlation. If the emission factors contain a random element or vary unpredictably from year to year, then separate emission factors should also be used (e.g., with fossil fuel carbon content that can change according to the market supply of the fuel and also contains its own uncertainty). Generally, uncertainty in activity data are assumed to be uncorrelated between years, and so two distributions should be input, even if their parameters are the same, so that two different random selections from these distributions will be generated in step 3. The computer package used may well enable other correlations to be set up and these capabilities could be used if sufficient information is available. However, this will probably be necessary in only a few cases.

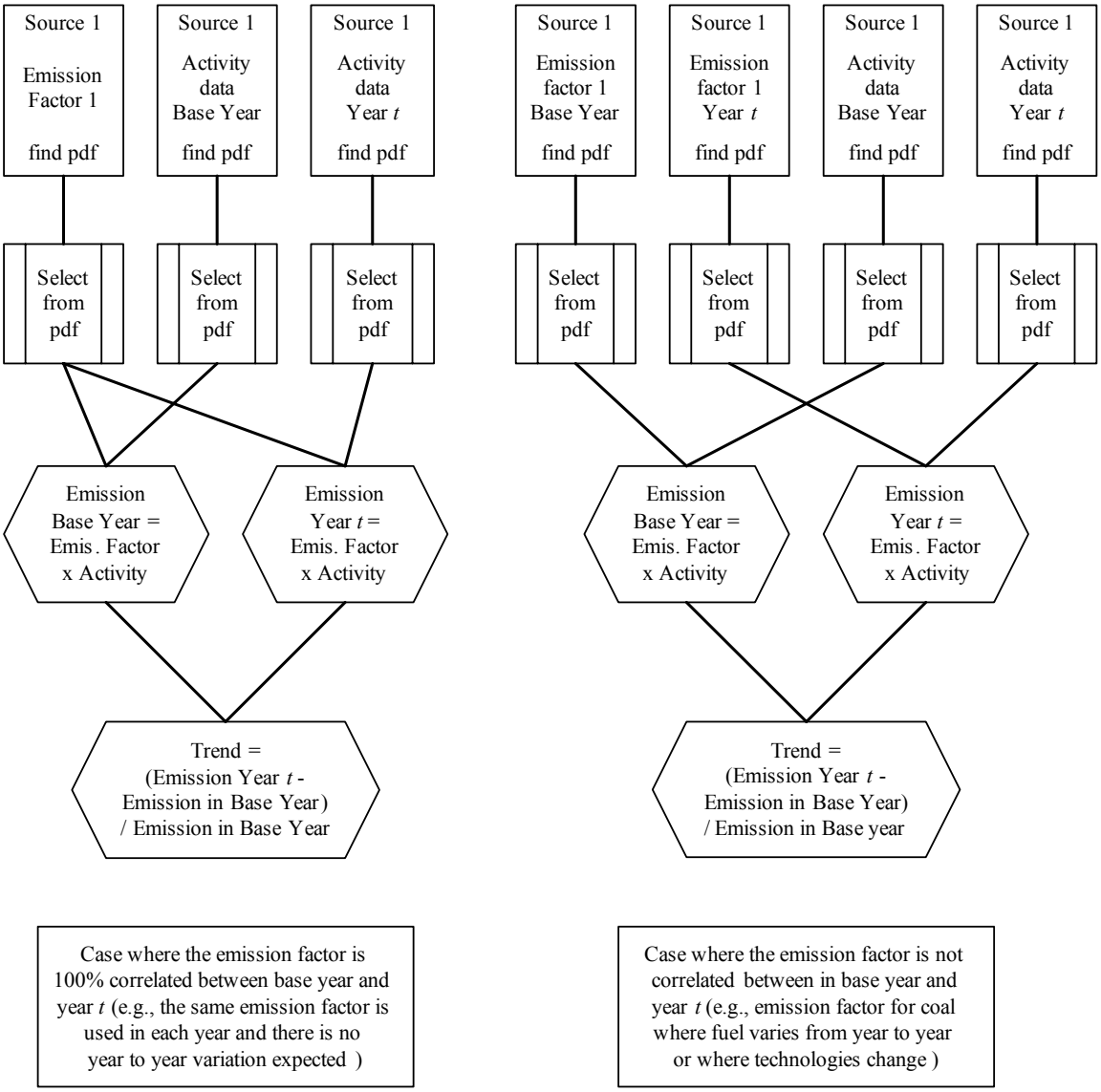
Step 2: Select random variables. The computer program will proceed as previously described, taking into account of any correlation between probability density functions (PDF). Figure 3.7, below, shows the calculation scheme for trend analysis.

Step 3: Estimate Emissions. As in the previous description, the variables selected in Step 2 will be used to estimate the total emissions.

Step 4: Results. The emissions total calculated in Step 3 is stored in a data file. The process then repeats from Step 2 until there is adequate convergence of the results. Considerations for this are the same as described above. A range of results is estimated at the same time including total and sectoral emissions/removals for the base year, total and sectoral emissions/removals for year *t*, and the percentage differences (trends) between these for the total and any sectors of interest.

[13] percentage difference = (value in year *t* - value in base year) / value in base year

Figure 3.7 **Calculation scheme for Monte Carlo analysis of the absolute emissions and the trend of a single category, estimated as emission factor times an activity rate**



CHOOSING A SIMULATION TECHNIQUE AND SAMPLE SIZE

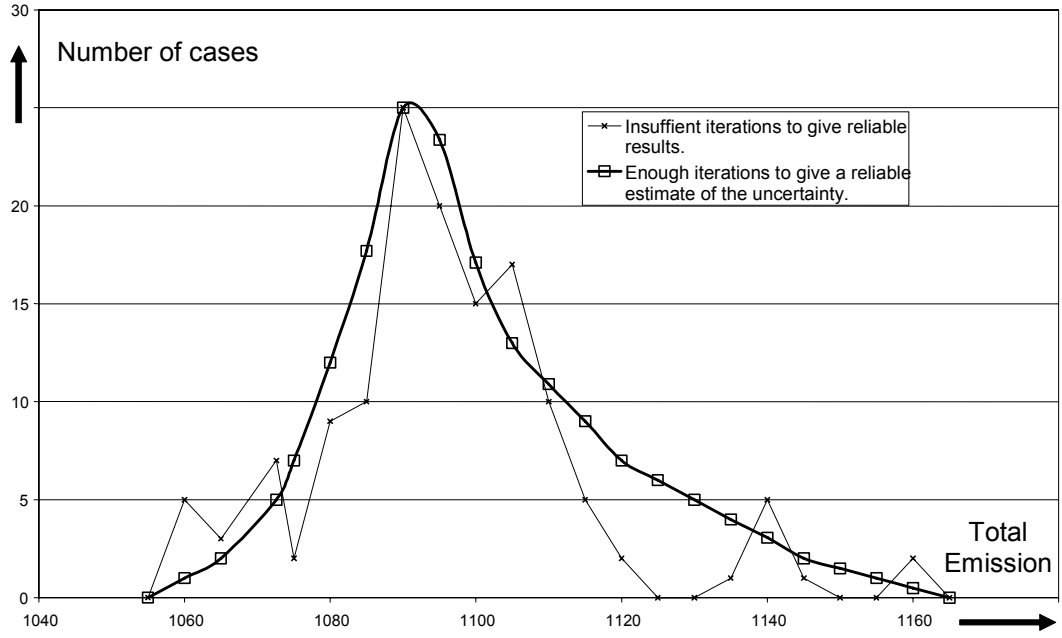
Several commercially available software tools can be used to perform Monte Carlo simulation. These tools can be stand-alone or used as add-ins to commonly used spreadsheet programs. Many software tools offer an option of different sampling methods, including random Monte Carlo simulation and variations of Latin Hypercube Sampling (LHS), which can produce 'smoother' looking model output distributions for sample sizes of only a few hundred samples. The disadvantage of using LHS is that one must decide ahead of time how many iterations to use. This is because two or more LHS simulations cannot be combined since they will use overlapping strata, leading to difficulties in interpreting results. In some cases, LHS can yield underestimates of the higher moments of PDFs, since the stratification method also can preclude clustering of very high or low values as can occur in random data sets. The overall suggestion is to use random Monte Carlo simulation as the default method, because it will give flexibility to continue a random simulation to larger and larger simulation sample sizes if necessary until the model output distribution converges[14] .

[14] Cullen and Frey (1999) provide more information on the comparison of LHS and Monte Carlo simulation (pp. 207-213).

The number of iterations can be determined either by setting the number of model runs, a priori, such as 10,000 and allowing the simulation to continue until reaching the set number, or by allowing the mean to reach a relatively stable point before terminating the simulation. For example, when the estimate for the 95 percent confidence range is determined to within ± 1%, then an adequately stable result has been found. This can be checked by plotting a frequency plot of the estimates of the emission. This plot should be reasonably smooth (see Figure 3.8).

Another alternative is to assess the precision of the current number of replicates based on the standard errors of the percentiles that were used to construct 95 percent confidence intervals. If the range of the confidence intervals for each percentile (2.5 and 97.5) is less than the reported precision, then the number of iterations should be adequate (e.g., emissions are reported values to a single digit after the decimal and the percentile confidence intervals are less than 0.1, such as 0.005). Therefore, the Monte Carlo percentile estimates are unlikely to change in the reported digits for other simulations with the same number of iterations.

Figure 3.8 Example frequency plots of the results of a Monte Carlo simulation



3.2.3.3 HYBRID COMBINATIONS OF APPROACHES 1 AND 2

For some inventories, it may be possible to use Approach 1 for most but not all of the source and sink categories. For example, many sources and sinks can be quantified using emission factors and activity data, but for some it is necessary to use a model or a more complex calculation procedure. Furthermore, dependencies may be important for some categories but not others, or the range of uncertainties may be large for some categories and not others. For these cases, a Monte Carlo-based method is more flexible and would typically produce better results.

If an inventory compiler has completed Approach 2 for only a subset of categories, the results can be combined with Approach 1 to produce an estimate of the uncertainty in national total emissions and the trend. This can be achieved by entering information at a disaggregated level, if correlations do not prevent doing so, into Approach 1. If there are significant correlations among a subset of categories, then the subset could be treated individually in Approach 2 but as an aggregation of categories in Approach 1. In the latter case, total emissions for the aggregation of the subset in the base year and in year *t* is entered in Columns C and D of the Approach 1 table. The results of the Approach 2 analysis for the uncertainty in total emissions in year *t* will be entered in Column G, and the results of the Approach 2 analysis for the contribution to the trend in national total emissions will be entered in Column M. The uncertainty contributions of the affected categories can be combined with those of other categories using the error propagation rules of Approach 1.

In some cases, most of the category uncertainties in an inventory might be estimated using Approach 2, with relatively few estimated using Approach 1. It is possible to incorporate Approach 1 estimates of uncertainty for

some categories into an Approach 2 methodology for combining uncertainties for the total inventory. This is done by using the uncertainty half-range obtained from Approach 1 to specify an appropriate PDF model to represent uncertainty for each category as part of the Monte Carlo simulation. Typically, a normal distribution would be a reasonable choice if the uncertainty range is small enough and a lognormal distribution is often appropriate if the uncertainty range is large. See also the section of 'Dealing with Large and Asymmetric Uncertainties in the Results of Approach 1' in Section 3.2.3.1 for more discussion of normal versus lognormal distribution assumptions.

3.2.3.4 COMPARISON BETWEEN APPROACHES

Two Approaches for uncertainty analysis have been presented:

- *Approach 1*: Estimation of uncertainties by category using Equations 3.1 and 3.2, and simple combination of uncertainties by category to estimate overall uncertainty for one year and the uncertainty in the trend.
- *Approach 2*: Estimation of uncertainties by category using Monte Carlo analysis, followed by the use of Monte Carlo techniques to estimate overall uncertainty for one year and the uncertainty in the trend.

Monte Carlo analysis can also be used in a restricted way within Approach 1 to combine activity data and emission factor uncertainties that have very wide or non-normal PDFs or both. This approach can also help deal with categories within Approach 1 that are estimated by process models, rather than by the classical 'emission factor times activity data' calculation. The choice between Approaches is discussed in Section 3.2.3.5 below.

Use of either Approach will provide insight into how individual categories and greenhouse gases contribute to uncertainty in total emissions in any year and to the emissions trend over time.

Application of Approach 2 to the UK inventory (Baggott *et al*., 2005) suggests that the 95 percent confidence interval is asymmetric and lies between about 6 percent below and 17 percent above the mean estimated value in 2003. The result for the UK takes into account the large relative range of uncertainty for N_2O flux from soils as well as the large contribution to total emissions from fossil fuel combustion. Application of Approach 1 to the same inventory suggests an uncertainty of about ±17%. On the trend, between 1990 and 2003, UK total net emissions in CO_2 equivalent are estimated to have fallen by 13 percent. Application of the Approach 2 suggests that 95 percent confidence interval is roughly symmetrical and lies between -11% and -16%. The corresponding Approach 1 result gives a range of about ±2% (i.e., -11% to -15%). So both methods give similar magnitudes in the trend uncertainty.

In the case of Finland, as shown in Section 3.6, year 2003 uncertainty (including both sources and sinks of greenhouse gases) was -14 to +15% according to Approach 2, and ±16% with Approach 1. For Finland, carbon stock changes in the AFOLU Sector are the dominant sources of uncertainty, while fossil fuels contribute the largest share to total emissions. Since the approximations inherent in Approach 1 mean that it cannot deal with asymmetry, this comparison is encouraging. Physically, the reason for the asymmetry identified under Approach 2 is that the uncertainty range of some very uncertain categories is constrained by the knowledge that the emissions cannot be less than zero. The Approach 2 method can make use of this extra knowledge, but the Approach 1 method cannot. In case of trend from 1990 to 2003 the uncertainty in Finland was -18 to +23% (percentage points) with Approach 2 and ±19% (percentage points) with Approach 1.

A separate evaluation of Approach 1 and Approach 2 for case studies based on synthetic inventory data revealed excellent agreement when the same set of input assumptions were used and when uncertainties were relatively small (Frey, 2005). For example, in a case study for which Approach 1 produced an estimate of ±6% in the current year inventory, and ±10% in the trend (in terms of percentage points relative to the mean percentage change), results from Approach 2 for the same input assumptions produced essentially the same result. When the uncertainty ranges were doubled for the emission factors and activity data, the uncertainty in the base estimates continued to agree well for Approach 1 and Approach 2, at approximately ±13% of the mean total emissions. The uncertainty in the trend was approximately ±20% (percentage points) in both cases. However, the uncertainty in the trend was slightly asymmetric in the Approach 2 result, at -19% to +22%. Thus, as the ranges of uncertainty increase, it is expected that Approach 2 will more appropriately characterise the range and skewness of uncertainties than Approach 1.

Although Approaches 1 and 2 focus on propagation of the random component of uncertainty through a model, it is *good practice* to combine methods for dealing with model uncertainty with either of the Approaches. An example of how to deal with model uncertainty in the context of Tier 3 is given in Box 3.3.

Furthermore, although Approach 1 is based upon key simplifying assumptions, it is possible to increase the flexibility of this approach by increasing the complexity of the error propagation equations. For example, error

propagation equations that contain additional terms can more accurately propagate uncertainty for multiplicative and quotient models and when uncertainties are skewed.

BOX 3.3
DEALING WITH MODEL UNCERTAINTY IN A PROBABILISTIC ANALYSIS

A Tier 3 modelling approach is designed for flexibility so that a national inventory can be conducted using a more highly refined model representing national circumstances than in Tier 1 or 2. In particular, it is *good practice* to address uncertainties attributed to model inputs and structure. Input uncertainty deals with activity data and possibly other ancillary information that is needed to describe the environmental setting, such as climate and soil characteristics in an inventory for the AFOLU Sector. Uncertainty in model structure is attributed to imperfect algorithms and parameterisation. Empirically-based approaches are commonly used for assessing structural uncertainties (Monte *et al.* 1996). This approach involves comparing modelled emissions estimates with measurements from experiments or a national monitoring network, which was designed for validation of model-based inventories, addressing both the bias and variance in modelled values (Falloon and Smith 2003).

A statistically-derived relationship may be used to quantify uncertainties in model structural error for a Tier 3 inventory, addressing imprecision based on the estimated variance, or a similar measure such as the Root Mean Squared Error, while also dealing with biases based on statistically significant differences between modelled and measured values (Falloon and Smith 2003). In practice, modelled emissions would be adjusted for biases to more accurately represent emissions for reporting purposes. Further, a statistically-derived relationship would yield a measure of variance for each condition that would be associated with the modelled values, similar to uncertainties attributed to emission factors in Tier 1 and 2 approaches. To complete the assessment, uncertainties in model inputs, such as activity data, would be combined with model structural uncertainty using error propagation equations or a Monte Carlo approach.

3.2.3.5 GUIDANCE ON CHOICE OF APPROACH

Where the conditions for applicability are met (relatively low uncertainty, no correlation between sources except those dealt with explicitly by Approach 1), Approach 1 and Approach 2 will give identical results. However, and perhaps paradoxically, these conditions are most likely to be satisfied where Tier 2 and Tier 3 methods are widely used and properly applied in the inventory, because these methods should give the most accurate and perhaps also the most precise results. There is therefore no direct theoretical connection between choice of Approach and choice of Tier. In practice, when Tier 1 methods are applied, Approach 1 will usually be used while the ability to apply Approach 2 is more likely where Tier 2 and 3 methods are being used, moreover for quantifying the uncertainty of emissions/removal estimates of complex systems such as in the AFOLU Sector.

When Approach 2 is selected, as part of QA/QC activities inventory agencies also are encouraged to apply Approach 1 because of the insights it provides and because it will not require a significant amount of additional work. Where Approach 2 is used, its estimates of overall uncertainty are to be preferred when reporting uncertainties (see Section 3.2.3.3).

3.3 UNCERTAINTY AND TEMPORAL AUTOCORRELATION

Where emission factors, sources of activity data or estimation methods vary within a time series the associated sources of uncertainty may also change. Approach 2 can take explicit account of this when setting up the component PDFs. In Approach 1 the current percentage uncertainties should be entered in the table, and in cases where changes throughout the time series mean that the assumption of good correlation of uncertainty in emission factors between years is no longer valid, Type A sensitivity should be used in place of Type B. If annual data are autocorrelated, then there will typically be less difference when comparing two years than if they are not autocorrelated, assuming that the autocorrelation is positive.

The issue of 'time series' can refer to inter-annual comparison of emissions in a Year *t* versus a base year, as has been set forth in Table 3.2 and the general reporting table given in Table 3.3, or it can refer to a broader set of

statistical methodologies for taking into account temporal autocorrelation. With regard to the latter interpretation, statistical time series techniques can be used to more accurately take into account temporal autocorrelations in order to reduce the estimate of uncertainty. For example, if emissions vary on a short term basis, such as for power plant emissions, the emissions at a given time period often depend on what the emissions were in the immediately preceding time periods as well as on the emissions at previous points in a cycle. For example, a power plant may require some amount of time in order to achieve a significant change in load. Thus, the emissions in a current hour are constrained somewhat depending on what emissions were in the previous hour. Furthermore, a power plant may respond to daily fluctuations in load, which are similar from one day to the next. Thus, the emissions at a given hour of the day may be correlated to those at a given hour of a preceding day. Likewise, there can be longer-term seasonal cycles, such as from one year to the next, that might induce a temporal correlation. Statistical time series methods can be fit to an adequate sample of empirical data in order to explain these temporal correlations. The unexplained portion of the model response is referred to as a random or white noise term. The white noise term is an indication of uncertainty in the ability to predict the emissions output. A detailed example of the application of time series models to emissions estimation is given by Abdel-Aziz and Frey (2003).

3.4 USE OF OTHER APPROPRIATE TECHNIQUES

The guidance offered here is not meant to preclude the use of other improved methods. For example, in applying Approach 1, an inventory compiler may wish to derive a similar approach from the generalised error propagation equations in order to account for more complex correlations or for differences in ranges of uncertainty in year *t* versus the base year. Such improvements are consistent with *good practice* as long as they are appropriately documented and justified. Furthermore, this document does not cover all situations that may be faced by an analyst. Therefore, the inventory compiler is encouraged to refer to the references cited in the end of the chapter for additional suggestions on how to perform uncertainty analyses.

3.5 REPORTING AND DOCUMENTATION

A great deal of effort can be expended in collecting information and data for a quantified uncertainty assessment and implementing a model to combine uncertainties across parameters, categories, and the entire inventory. However, all that effort can result in little benefit to a country's inventory if steps are not also taken to report and document the findings of an uncertainty assessment so that they can lead to real improvements in the quality of data collected and the inventory as a whole. The integration of a country's uncertainty assessment efforts with the implementation of data quality investigations within its QA/QC system can help solve this problem.

Given the large number of inputs and assumptions needed to document an uncertainty analysis, it is not feasible to report all information. The information reported should be sufficient to provide the key assumptions, choice of methods and detailed results. Overall, documentation should be sufficient to support the estimates and enable duplication of the uncertainty estimates. In particular, the documentation should touch upon the following issues (as they pertain to a particular variable):

- Which causes of uncertainty are addressed (see Table 3.1).
- Which methods for addressing uncertainty were used (see Table 3.1).
- What is the source of any data or models that were used as the basis for estimating uncertainty.
- For an estimate of bias, explain what is the magnitude of the error expressed on a relative or absolute basis as appropriate (specify which and give proper units).
- If uncertainty was estimated based upon data, explain how uncertainty was distinguished from variability and how the appropriate geographic extent, averaging time (e.g., annual), and other representativeness considerations were addressed in selecting and analyzing data. Provide a brief summary of the data itself, including the mean, sample standard deviation, and sample size. Provide additional detail as appropriate if data were stratified or contain other components of uncertainty (e.g., precision and accuracy of measurement methods used to obtain the data).
- For an estimate of random error in the form of a range or a distribution, provide sufficient information to uniquely specify the range (e.g., plus or minus percentage variation relative to the mean, or parameters of a PDF).

- For estimates of uncertainty based upon expert judgement, the following information should be documented and archived:
  - (i) reference number for judgement;
  - (ii) date;
  - (iii) name of expert(s) involved;
  - (iv) experts' background (references, roles, etc.);
  - (v) the variable being judged;
  - (vi) the logical basis for judgement, including any data taken into consideration. This should include the rationale for the high end, low end, and central tendency of the distribution;
  - (vii) the resultant PDF, or the range and most likely value and the PDF subsequently inferred;
  - (viii) identification of any external reviewers;
  - (ix) results of any external review;
  - (x) approval by inventory compiler, specifying date and person.
- Explanation of any correlation or dependencies that were accounted for between two or more inputs or with respect to autocorrelation.
- Explanation of any special considerations that might be unique to a particular country or situation, such as the use of various statistical techniques for dealing with non-detects, mixture distributions, extrapolation, and so on.
- Explanation of differences in results between Approach 1 and 2.

In addition to documentation of the uncertainty estimates for inputs to an inventory, documentation should be provided regarding the general approach used and whether it is based primarily on Approach 1 or Approach 2. Any modifications to these Approaches should be explained and appropriately justified.

The reporting of uncertainties also requires a discussion of limitations and caveats for any quantitative uncertainty estimates produced that are suspected to present an incomplete representation of all causes of uncertainty. During the process of collecting information on inputs for an uncertainty assessment (e.g., empirical or expert judgement as a basis for PDFs, characterisations of conceptualisation and model uncertainty), the likely causes of the various uncertainties identified—including potential biases—should be documented. These likely causes should be documented whether or not they were quantified and include any specific recommendations available regarding how they can be reduced.

Similarly, when reporting and interpreting the results from a quantitative uncertainty assessment, it is important to keep in mind the limitations of the approach used to combine uncertainties. For example, although Approach 1 can address some causes of correlation, any possible biases associated with other causes of correlations that might exist (e.g., between categories) that are identified in the course of an uncertainty assessment should be documented.

Table 3.3 is a generalised table for reporting uncertainty of an inventory, regardless of the Approach followed. If the point estimate and the mean estimate of emissions/removals are not the same value, it is *good practice* for the uncertainty ranges shown in Columns E, F, G, and J to be estimated relative to point estimates used when reporting the national inventory. If the point estimate and the mean estimates differ, then it is advisable to consider why they differ and possibly revisit the point estimate in order to identify and account for bias.

TABLE 3.3
GENERAL REPORTING TABLE FOR UNCERTAINTY

A	B	C	D	E		F		G		H	I	J		K
IPCC category	Gas	Base year emissions /removals	Year *t* emissions /removals	Activity data uncertainty		Emission factor/estimation parameter uncertainty (combined if more than one estimation parameter is used)		Combined uncertainty		Contribution to variance in Year *t*	Inventory trend in national emissions for year *t* increase with respect to base year	Uncertainty introduced into the trend in total national emissions with respect to base year		Approach and Comments
		Gg CO_2 equivalent	Gg CO_2 equivalent	(-) %	(+) %	(-) %	(+) %	(-) %	(+) %	(fraction)	(% of base year)	(-) %	(+) %	
E.g., 1.A.1. Energy Industries Fuel 1	CO_2													
E.g., 1.A.1. Energy Industries Fuel 2	CO_2													
Etc...	…													
Total										1.000				

Notes:

Column C: Base year emissions in Gg CO_2 equivalent by category and gas.

Column D: Year *t* emissions in Gg CO_2 equivalent by category and gas. Year *t* is the year of interest or the current year.

Columns E and F: Uncertainties in activity and emission factor estimates (Columns E and F) should be reported where possible, but it is understood that some calculation methods for some categories may not be amenable to this type of reporting. Thus, where this information is not available, the table entry can be left blank.

Column G: An uncertainty estimate for each category should be reported, relative to the mean estimate, even if uncertainties cannot be further disaggregated by activity and emission factor in a particular case. For the foot of the table, report the uncertainty in the total inventory. This must be obtained through the Approach 1 or 2 calculations, and cannot be determined simply by summing quantities in the columns.

Column H: Report the 'contribution to uncertainty'. It is estimated dividing the variance of each category by the total variance of the inventory ($\sigma_x^2 / \sum \sigma_x^2$). If Approach 1 has been used, it is calculated dividing each entry in Column H of Table 3.2 by the value, in the same column, in the line 'Total' of Table 3.2. General methodology to be applied when Approach 2 is used and when uncertainty is asymmetric is provided in Section 3.2.3.

Column I: Report the inventory trend, estimated as:

$$Mean\,Trend\,(\%) = \left(\frac{Year\,t\,emissions - Base\,year\,emissions}{Base\,year\,emissions} \right) \bullet 100 \ .$$

Report separately for each category by row and report for the total inventory at the foot of the column.

Column J: This is the uncertainty in the trend by category. For the 'total' at the foot of the table, the overall uncertainty in the trend for the entire inventory should be given. The uncertainty in the trend is based on *percentage points* with respect to the inventory trend. For example, if the inventory trend is -5%, and if the 95% probability range of the trend is -8% to -3%, then the uncertainty in trend is reported as -3% to +2%.

Column K: Indicate whether Approach 1 or Approach 2 was used, and include any other comments that might help clarify the methodology or information sources.

General Comments on Columns E, F, G, and J: For each of these columns, two subcolumns are provided to facilitate reporting of uncertainty ranges that are asymmetric. For example, if the uncertainty range is -50% to +100%, then '50' should be reported the subcolumn headed as '(-)%' and '100' should be reported in the column headed as '(+)%'.

3.6 EXAMPLES

This section presents two examples of uncertainty estimates for inventories, both based upon the Finnish 2003 greenhouse gas emission inventory. These examples are country-specific and are shown here only for the purpose of illustrating procedures and general insights. The specific uncertainty estimates and results will differ among countries.

The example of Table 3.4 is based upon Approach 1, and is shown in the general format of the Approach 1 worksheet (Table 3.2). The results indicate that the net emissions in year *t*, which is 2003 in this example, is 67,730 Gg CO_2 equivalent with an uncertainty of ±15.9%, which corresponds to a 95 percent probability range of 56,970 to 78,490 Gg CO_2 equivalent. Based upon the total base year and year *t* inventories reported in the table, the average trend is a 42 percent increase in emissions from 1990 to 2003. The uncertainty in the trend is ±19% (percentage points), which corresponds to a 95 percent probability range for the trend of 24% to 61% with respect to the base year emissions.

The example of Table 3.5 is based upon Approach 2, and is shown in the format of the General Reporting Table for Uncertainty shown in Table 3.3. The results indicate that the net emissions in year *t* are 67,730 Gg CO_2 equivalent with an uncertainty range of -14 to +15 percent, which corresponds to a 95 percent probability range of 58,490 to 78,130 Gg CO_2 equivalent. Based upon the total base year and year *t* inventories reported in the table, the average trend is a 42 % increase in emissions from 1990 to 2003. The uncertainty in the trend is -18 to +23% (percentage points) which corresponds to a 95 percent probability range for the trend of 25% to 65% with respect to the base year emissions.

These examples illustrate that the results from Approaches 1 and 2 can be very similar when the overall uncertainty is relatively small. However, Approach 2 is a more flexible approach that enables quantification of asymmetry in probability ranges, such as for the year *t* inventory.

TABLE 3.4
EXAMPLE OF AN APPROACH 1 UNCERTAINTY ANALYSIS FOR FINLAND (BASED ON STATISTICS FINLAND, 2005)

The aggregation level and uncertainty estimates are country-specific for Finland, and do not represent recommended uncertainties or level of aggregation for other countries.

A	B	C	D	E	F	G	H	I	J	K	L	M
IPCC category	**Gas**	**Base year emissions or removals** Gg CO_2 equivalent	**Year *t* emissions or removals** Gg CO_2 equivalent	**Activity data uncertainty** %	**Emission factor /estimation parameter uncertainty** %	**Combined uncertainty** %	**Contribution to variance by source/sink category in year *t***	**Type A sensitivity** %	**Type B sensitivity** %	**Uncertainty in trend in national emissions introduced by emission factor / estimation parameter uncertainty** %	**Uncertainty in trend in national emissions introduced by activity data uncertainty** %	**Uncertainty introduced into the trend in total national emissions** %
1.A Fuel Combustion Activities												
Liquid	CO_2	27 232	27 640	2%	2%	3%	0.0001	0.2320	0.5806	0.46%	1.64%	0.03%
Solid	CO_2	15 722	22 753	2%	3%	3%	0.0001	0.0080	0.4780	0.02%	1.08%	0.01%
Gas	CO_2	5 073	9 350	1%	1%	1%	0.0000	0.0447	0.1964	0.04%	0.28%	0.00%
Peat	CO_2	5 656	10 676	4%	5%	7%	0.0001	0.0552	0.2243	0.28%	1.36%	0.02%
1.A.1 Energy Industries												
Liquid	CH_4	6	7	2%	75%	75%	0.0000	0.0000	0.0001	0.00%	0.00%	0.00%
	N_2O	26	30	2%	75%	75%	0.0000	0.0001	0.0006	0.01%	0.00%	0.00%
Solid	CH_4	9	16	2%	75%	75%	0.0000	0.0001	0.0003	0.01%	0.00%	0.00%
	N_2O	85	162	2%	50%	50%	0.0000	0.0009	0.0034	0.04%	0.01%	0.00%
Gas	CH_4	4	9	1%	75%	75%	0.0000	0.0001	0.0002	0.01%	0.00%	0.00%
	N_2O	18	51	1%	50%	50%	0.0000	0.0005	0.0011	0.03%	0.00%	0.00%
Biomass	CH_4	2	31	20%	50%	54%	0.0000	0.0006	0.0006	0.03%	0.02%	0.00%
	N_2O	10	80	20%	150%	151%	0.0000	0.0014	0.0017	0.21%	0.05%	0.00%
Peat	CH_4	5	7	5%	50%	50%	0.0000	0.0000	0.0002	0.00%	0.00%	0.00%
	N_2O	141	226	5%	150%	150%	0.0000	0.0005	0.0047	0.08%	0.03%	0.00%
1.A.2 Manufacturing Industries and Construction												
Liquid	CH_4	9	7	2%	75%	75%	0.0000	0.0001	0.0001	0.01%	0.00%	0.00%
	N_2O	39	41	2%	75%	75%	0.0000	0.0003	0.0009	0.02%	0.00%	0.00%
Solid	CH_4	4	2	2%	75%	75%	0.0000	0.0001	0.0001	0.01%	0.00%	0.00%
	N_2O	108	90	2%	50%	50%	0.0000	0.0013	0.0019	0.07%	0.01%	0.00%

TABLE 3.4 (CONTINUED)
EXAMPLE OF AN APPROACH 1 UNCERTAINTY ANALYSIS FOR FINLAND (BASED ON STATISTICS FINLAND, 2005)

The aggregation level and uncertainty estimates are country-specific for Finland, and do not represent recommended uncertainties or level of aggregation for other countries.

A	B	C	D	E	F	G	H	I	J	K	L	M
IPCC category	**Gas**	**Base year emissions or removals** Gg CO_2 equivalent	**Year *t* emissions or removals** Gg CO_2 equivalent	**Activity data uncertainty** %	**Emission factor /estimation parameter uncertainty** %	**Combined uncertainty** %	**Contribution to variance by source/sink category in year *t***	**Type A sensitivity** %	**Type B sensitivity** %	**Uncertainty in trend in national emissions introduced by emission factor / estimation parameter uncertainty** %	**Uncertainty in trend in national emissions introduced by activity data uncertainty** %	**Uncertainty introduced into the trend in total national emissions** %
Gas	CH_4	5	6	1%	75%	75%	0.0000	0.0000	0.0001	0.00%	0.00%	0.00%
	N_2O	17	19	1%	50%	50%	0.0000	0.0001	0.0004	0.01%	0.00%	0.00%
Biomass	CH_4	20	19	15%	50%	52%	0.0000	0.0002	0.0004	0.01%	0.01%	0.00%
	N_2O	111	81	15%	150%	151%	0.0000	0.0016	0.0017	0.24%	0.04%	0.00%
Peat	CH_4	4	3	5%	50%	50%	0.0000	0.0001	0.0001	0.00%	0.00%	0.00%
	N_2O	56	29	5%	150%	150%	0.0000	0.0011	0.0006	0.16%	0.00%	0.00%
1.A.3 Transport												
a. Civil Aviation	CH_4	0.4	0.3	5%	100%	100%	0.0000	0.0000	0.0000	0.00%	0.00%	0.00%
	N_2O	4	4	5%	150%	150%	0.0000	0.0000	0.0001	0.01%	0.00%	0.00%
b. Road Transportation												
Gasoline	CH_4	78	40	1%	50%	50%	0.0000	0.0015	0.0008	0.07%	0.00%	0.00%
Cars with Catalytic Converters	N_2O	32	410	1%	378%	378%	0.0005	0.0076	0.0086	2.89%	0.01%	0.08%
Cars without Catalytic Converters	N_2O	59	22	1%	259%	259%	0.0000	0.0013	0.0005	0.34%	0.00%	0.00%
Diesel	CH_4	12	6	1%	50%	50%	0.0000	0.0002	0.0001	0.01%	0.00%	0.00%
	N_2O	68	84	1%	158%	158%	0.0000	0.0003	0.0018	0.04%	0.00%	0.00%
Natural gas	CH_4	0.0	2	1%	50%	50%	0.0000	0.0000	0.0000	0.00%	0.00%	0.00%
	N_2O	0.0	0.0	1%	150%	150%	0.0000	0.0000	0.0000	0.00%	0.00%	0.00%
c. Railways	CH_4	0.2	0.2	5%	110%	110%	0.0000	0.0000	0.0000	0.00%	0.00%	0.00%
	N_2O	2	1	5%	150%	150%	0.0000	0.0000	0.0000	0.00%	0.00%	0.00%

TABLE 3.4 (CONTINUED)
EXAMPLE OF AN APPROACH 1 UNCERTAINTY ANALYSIS FOR FINLAND (BASED ON STATISTICS FINLAND, 2005)

The aggregation level and uncertainty estimates are country-specific for Finland, and do not represent recommended uncertainties or level of aggregation for other countries.

A	B	C	D	E	F	G	H	I	J	K	L	M
IPCC category	**Gas**	**Base year emissions or removals** Gg CO_2 equivalent	**Year *t* emissions or removals** Gg CO_2 equivalent	**Activity data uncertainty** %	**Emission factor /estimation parameter uncertainty** %	**Combined uncertainty** %	**Contribution to variance by source/sink category in year *t***	**Type A sensitivity** %	**Type B sensitivity** %	**Uncertainty in trend in national emissions introduced by emission factor / estimation parameter uncertainty** %	**Uncertainty in trend in national emissions introduced by activity data uncertainty** %	**Uncertainty introduced into the trend in total national emissions** %
d. Water-borne Navigation												
Residual Oil & Gas/Diesel Oil	CH_4	0.5	1	10%	100%	100%	0.0000	0.0000	0.0000	0.00%	0.00%	0.00%
	N_2O	2	3	10%	150%	150%	0.0000	0.0000	0.0001	0.00%	0.00%	0.00%
Gasoline	CH_4	7	4	20%	100%	102%	0.0000	0.0001	0.0001	0.01%	0.00%	0.00%
	N_2O	0.4	0.6	20%	150%	151%	0.0000	0.0000	0.0000	0.00%	0.00%	0.00%
e. Other Transportation												
Gasoline&Diesel	CH_4	5	6	30%	50%	58%	0.0000	0.0000	0.0001	0.00%	0.01%	0.00%
Gasoline	N_2O	1	1	30%	150%	153%	0.0000	0.0000	0.0000	0.00%	0.00%	0.00%
Diesel	N_2O	4	4	30%	150%	153%	0.0000	0.0000	0.0001	0.01%	0.00%	0.00%
1.A.4 Other Sectors												
Liquid	CH_4	19	15	3%	75%	75%	0.0000	0.0002	0.0003	0.02%	0.00%	0.00%
	N_2O	56	47	3%	75%	75%	0.0000	0.0007	0.0010	0.05%	0.00%	0.00%
Solid	CH_4	2	0.6	10%	75%	76%	0.0000	0.0001	0.0000	0.00%	0.00%	0.00%
	N_2O	0.5	0.3	10%	50%	51%	0.0000	0.0000	0.0000	0.00%	0.00%	0.00%
Gas	CH_4	0.1	0.3	5%	75%	75%	0.0000	0.0000	0.0000	0.00%	0.00%	0.00%
	N_2O	1	1	5%	50%	50%	0.0000	0.0000	0.0000	0.00%	0.00%	0.00%
Biomass	CH_4	282	307	15%	150%	151%	0.0000	0.0020	0.0064	0.30%	0.14%	0.00%
	N_2O	56	61	15%	150%	151%	0.0000	0.0004	0.0013	0.06%	0.03%	0.00%
Peat	CH_4	1	1	25%	50%	56%	0.0000	0.0000	0.0000	0.00%	0.00%	0.00%
	N_2O	1	2	25%	150%	152%	0.0000	0.0000	0.0000	0.00%	0.00%	0.00%

TABLE 3.4 (CONTINUED)
EXAMPLE OF AN APPROACH 1 UNCERTAINTY ANALYSIS FOR FINLAND (BASED ON STATISTICS FINLAND, 2005)

The aggregation level and uncertainty estimates are country-specific for Finland, and do not represent recommended uncertainties or level of aggregation for other countries.

A	B	C	D	E	F	G	H	I	J	K	L	M
IPCC category	**Gas**	**Base year emissions or removals** Gg CO_2 equivalent	**Year *t* emissions or removals** Gg CO_2 equivalent	**Activity data uncertainty** %	**Emission factor /estimation parameter uncertainty** %	**Combined uncertainty** %	**Contribution to variance by source/sink category in year *t***	**Type A sensitivity** %	**Type B sensitivity** %	**Uncertainty in trend in national emissions introduced by emission factor / estimation parameter uncertainty** %	**Uncertainty in trend in national emissions introduced by activity data uncertainty** %	**Uncertainty introduced into the trend in total national emissions** %
1.A.5 Non-Specified												
Liquid	CH_4	2	2	7%	75%	75%	0.0000	0.0000	0.0000	0.00%	0.00%	0.00%
	N_2O	6	9	7%	75%	75%	0.0000	0.0000	0.0002	0.00%	0.00%	0.00%
Gas	CH_4	0.3	0.4	13%	75%	76%	0.0000	0.0000	0.0000	0.00%	0.00%	0.00%
	N_2O	1	2	13%	50%	52%	0.0000	0.0000	0.0000	0.00%	0.00%	0.00%
1.B Fugitive Emissions from Fuels												
1.B.2 Oil and Natural Gas												
a.ii. Oil – Flaring	CO_2	123	63	50%	0%	50%	0.0000	0.0024	0.0013	0.00%	0.09%	0.00%
a.iii.4 Oil – Refining	CH_4	8	10	2%	90%	90%	0.0000	0.0000	0.0002	0.00%	0.00%	0.00%
b.iii.4 Natural Gas - Transmission and Storage	CH_4	4	12	3%	0%	3%	0.0000	0.0001	0.0003	0.00%	0.00%	0.00%
b. iii.5 Natural gas - Distribution	CH_4	0	40	5%	0%	5%	0.0000	0.0008	0.0008	0.00%	0.01%	0.00%
2 Industrial Processes and Product Use												
2.A.1 Cement Production	CO_2	786	500	2%	5%	5%	0.0000	0.0130	0.0105	0.06%	0.03%	0.00%
2.A.2 Lime Production	CO_2	383	513	2%	3%	4%	0.0000	0.0007	0.0108	0.00%	0.03%	0.00%
2.A.3 and 2.A.4 Limestone and Dolomite Use[1]	CO_2	99	148	7%	9%	11%	0.0000	0.0002	0.0031	0.00%	0.03%	0.00%
2.A.3 and 2.A.4 Soda Ash Use[1]	CO_2	18	20	7%	2%	7%	0.0000	0.0001	0.0004	0.00%	0.00%	0.00%
2.B.2 Nitric Acid Production	N_2O	1 595	1 396	5%	100%	100%	0.0004	0.0184	0.0293	1.84%	0.21%	0.03%
2.B.8.b Ethylene	CH_4	4	5	5%	20%	21%	0.0000	0.0000	0.0001	0.00%	0.00%	0.00%
2.B.10 Other	CO_2	60	147	12%	5%	13%	0.0000	0.0013	0.0031	0.01%	0.05%	0.00%
2.C.1 Iron and Steel Production	CH_4	5	9	3%	20%	20%	0.0000	0.0000	0.0002	0.00%	0.00%	0.00%

TABLE 3.4 (CONTINUED)
EXAMPLE OF AN APPROACH 1 UNCERTAINTY ANALYSIS FOR FINLAND (BASED ON STATISTICS FINLAND, 2005)

The aggregation level and uncertainty estimates are country-specific for Finland, and do not represent recommended uncertainties or level of aggregation for other countries.

A	B	C	D	E	F	G	H	I	J	K	L	M
IPCC category	**Gas**	**Base year emissions or removals** Gg CO_2 equivalent	**Year *t* emissions or removals** Gg CO_2 equivalent	**Activity data uncertainty** %	**Emission factor /estimation parameter uncertainty** %	**Combined uncertainty** %	**Contribution to variance by source/sink category in year *t***	**Type A sensitivity** %	**Type B sensitivity** %	**Uncertainty in trend in national emissions introduced by emission factor / estimation parameter uncertainty** %	**Uncertainty in trend in national emissions introduced by activity data uncertainty** %	**Uncertainty introduced into the trend in total national emissions** %
2.D Non-Energy Products from Fuels and Solvent Use	CO_2	640	830	50%	5%	50%	0.0000	0.0017	0.0174	0.01%	1.23%	0.02%
2.F.1 Refrigeration and Air Conditioning	HFCs	0	578	26%	0%	26%	0.0000	0.0121	0.0121	0.00%	0.45%	0.00%
2.F.2 Foam Blowing Agents	HFCs	0	25	24%	0%	24%	0.0000	0.0005	0.0005	0.00%	0.02%	0.00%
2.F.4 Aerosols	HFCs	0	63	10%	0%	10%	0.0000	0.0013	0.0013	0.00%	0.02%	0.00%
2.G.1 Electrical Equipment	SF_6	87	22	88%	0%	88%	0.0000	0.0021	0.0005	0.00%	0.06%	0.00%
2.G.3.a Medical Applications	N_2O	62	40	30%	20%	36%	0.0000	0.0010	0.0008	0.02%	0.04%	0.00%
2.H.3 Other (grouped data of f-gases)	HFCs, PFCs, SF_6	8	21	38%	0%	38%	0.0000	0.0002	0.0004	0.00%	0.02%	0.00%
3 AFOLU												
3.A.1 Enteric Fermentation	CH_4	1 868	1 537	0%	31%	31%	0.0000	0.0235	0.0323	0.72%	0.00%	0.01%
3.A.2 Manure Management	CH_4	215	222	0%	16%	16%	0.0000	0.0018	0.0047	0.03%	0.00%	0.00%
3.A.2 Manure Management	N_2O	623	461	0%	83%	83%	0.0000	0.0089	0.0097	0.74%	0.00%	0.01%
3.B.1.a Forest Land Remaining Forest Land												
carbon stock change in biomass	CO_2	-23 798	-21 354	0%	35%	35%	0.0122	0.2640	0.4486	9.24%	0.00%	0.85%
3.B.2.a Cropland Remaining Cropland												
net carbon stock change in mineral soils	CO_2	-535	-1 113	0%	100%	100%	0.0003	0.0074	0.0234	0.74%	0.00%	0.01%
net carbon stock change in organic soils	CO_2	1 813	1 324	20%	90%	92%	0.0003	0.0264	0.0278	2.37%	0.79%	0.06%

TABLE 3.4 (CONTINUED)
EXAMPLE OF AN APPROACH 1 UNCERTAINTY ANALYSIS FOR FINLAND (BASED ON STATISTICS FINLAND, 2005)

The aggregation level and uncertainty estimates are country-specific for Finland, and do not represent recommended uncertainties or level of aggregation for other countries.

A	B	C	D	E	F	G	H	I	J	K	L	M
IPCC category	**Gas**	**Base year emissions or removals** Gg CO_2 equivalent	**Year *t* emissions or removals** Gg CO_2 equivalent	**Activity data uncertainty** %	**Emission factor /estimation parameter uncertainty** %	**Combined uncertainty** %	**Contribution to variance by source/sink category in year *t***	**Type A sensitivity** %	**Type B sensitivity** %	**Uncertainty in trend in national emissions introduced by emission factor / estimation parameter uncertainty** %	**Uncertainty in trend in national emissions introduced by activity data uncertainty** %	**Uncertainty introduced into the trend in total national emissions** %
3.B.3.a Grassland Remaining Grassland												
net carbon stock change in mineral soils	CO_2	-1 181	2 907	0%	100%	100%	0.0018	0.0964	0.0611	9.64%	0.00%	0.93%
net carbon stock change in organic soils	CO_2	109	67	30%	90%	95%	0.0000	0.0019	0.0014	0.17%	0.06%	0.00%
3.B.4.ai Peatlands Remaining Peatlands	CO_2	503	547	15%	208%	208%	0.0003	0.0036	0.0115	0.74%	0.08%	0.01%
3.B.4.ai Peatlands Remaining Peatlands	CH_4	5	6	15%	208%	208%	0.0000	0.0000	0.0001	0.01%	0.00%	0.00%
3.C.1.a Biomass Burning in Forest Lands	CO_2	180	91	10%	70%	71%	0.0000	0.0035	0.0019	0.24%	0.03%	0.00%
3.C.1.a Biomass Burning in Forest Lands	CH_4	16	8	10%	70%	71%	0.0000	0.0003	0.0002	0.02%	0.00%	0.00%
3.C.1.a Biomass Burning in Forest Lands	N_2O	2	1	10%	70%	71%	0.0000	0.0000	0.0000	0.00%	0.00%	0.00%
3.C.2 Liming	CO_2	618	277	20%	20%	28%	0.0000	0.0127	0.0058	0.25%	0.16%	0.00%
3.C.4 Direct N_2O Emissions from Managed Soils: Agricultural Soils	N_2O	3 486	2 608	0%	227%	227%	0.0077	0.0494	0.0548	11.23%	0.00%	1.26%
3.C.4 Direct N_2O Emissions from Managed Soils: N Fertilizers Application, Forest Land	N_2O	27.0	11.3	10%	380%	380%	0.0000	0.0006	0.0002	0.22%	0.00%	0.00%
3.C.5 Indirect N_2O Emissions from Managed Soils	N_2O	735	592	0%	334%	334%	0.0009	0.0095	0.0124	3.18%	0.00%	0.10%
4 Waste												
4.A Solid Waste Disposal	CH_4	3 678	2 497	0%	43%	43%	0.0003	0.0574	0.0525	2.47%	0.00%	0.06%

TABLE 3.4 (CONTINUED)
EXAMPLE OF AN APPROACH 1 UNCERTAINTY ANALYSIS FOR FINLAND (BASED ON STATISTICS FINLAND, 2005)

The aggregation level and uncertainty estimates are country-specific for Finland, and do not represent recommended uncertainties or level of aggregation for other countries.

A	B	C	D	E	F	G	H	I	J	K	L	M
IPCC category	**Gas**	**Base year emissions or removals** Gg CO_2 equivalent	**Year *t* emissions or removals** Gg CO_2 equivalent	**Activity data uncertainty** %	**Emission factor /estimation parameter uncertainty** %	**Combined uncertainty** %	**Contribution to variance by source/sink category in year *t***	**Type A sensitivity** %	**Type B sensitivity** %	**Uncertainty in trend in national emissions introduced by emission factor / estimation parameter uncertainty** %	**Uncertainty in trend in national emissions introduced by activity data uncertainty** %	**Uncertainty introduced into the trend in total national emissions** %
4.D.1 Domestic Wastewater Treatment and Discharge												
sparsely populated areas	CH_4	118	95	15%	32%	35%	0.0000	0.0015	0.0020	0.05%	0.04%	0.00%
densely populated areas	CH_4	12	13	5%	104%	105%	0.0000	0.0001	0.0003	0.01%	0.00%	0.00%
sparsely populated areas	N_2O	21	18	10%	380%	380%	0.0000	0.0002	0.0004	0.09%	0.01%	0.00%
densely populated areas	N_2O	84	66	5%	380%	380%	0.0000	0.0011	0.0014	0.43%	0.01%	0.00%
4.D.2 Industrial Wastewater Treatment and Discharge	CH_4	22	19	10%	104%	105%	0.0000	0.0003	0.0004	0.03%	0.01%	0.00%
4.D.2 Industrial Wastewater Treatment and Discharge	N_2O	28	17	5%	380%	380%	0.0000	0.0005	0.0004	0.17%	0.00%	0.00%
4.E Other: N input from Fish Farming	N_2O	8	3	10%	380%	380%	0.0000	0.0002	0.0001	0.07%	0.00%	0.00%
Total		**47 604**	**67 730**				**0.0252**					**0.0349**
					Percentage uncertainty in total inventory		**15.9%**				**Trend uncertainty**	**18.7%**

[1] Uncertainty assessment was made at the aggregation level used by Finland in 2003 inventory, and therefore glass production could not be separated.

TABLE 3.5
EXAMPLE OF REPORTING OF APPROACH 2 UNCERTAINTY ANALYSIS USING GENERAL REPORTING TABLE FOR UNCERTAINTY
Emissions, removals and uncertainties are from National Inventory of Finland for year 2003 (Statistics Finland, 2005). The aggregation level and uncertainty estimates are country-specific for Finland, and do not represent recommended uncertainties or level of aggregation for other countries.

A	B	C	D	E		F		G		H	I	J		K
IPCC category	**Gas**	**Base year emissions or removals**	**Year *t* emissions or removals**	**Activity data uncertainty**		**Emission factor uncertainty**		**Combined uncertainty**		**Contribution to variance in year *t*[a]**	**Inventory trend in national emissions for year *t* increase with respect to base year**	**Uncertainty introduced into the trend in total national emissions with respect to base year**		**Approach and Comments**
		Gg CO_2 equivalent	Gg CO_2 equivalent	(-) %	(+) %	(-) %	(+) %	(-) %	(+) %	(fraction)	(% of base year)	(-) %	(+) %	Approach 2
1.A Fuel Combustion Activities														
Liquid	CO_2	27 232	27 640	2	2	2	2	3	3	0.0061	1	-3	3	
Solid	CO_2	15 722	22 753	2	2	3	3	3	3	0.0061	45	-3	3	
Gas	CO_2	5 073	9 350	1	1	1	1	1	1	0.0002	84	-3	3	
Peat	CO_2	5 656	10 676	4	4	5	5	6	7	0.0050	89	-11	11	
1.A.1 Energy Industries														
Liquid	CH_4	6	7	2	2	75	10	75	12	0.0000	18	-32	39	
	N_2O	26	30	2	2	75	10	75	12	0.0000	15	-30	39	
Solid	CH_4	9	16	2	2	75	10	75	12	0.0000	91	-43	59	
	N_2O	85	162	2	2	50	50	50	50	0.0001	91	-23	25	
Gas	CH_4	4	9	1	1	75	10	76	11	0.0000	140	-57	87	
	N_2O	18	51	1	1	50	50	51	50	0.0000	188	-37	39	
Biomass	CH_4	2	31	20	20	50	50	52	57	0.0000	1 370	-398	544	
	N_2O	10	80	20	20	70	150	71	154	0.0001	729	-260	374	
Peat	CH_4	5	7	5	5	50	50	50	50	0.0000	37	-18	21	
	N_2O	141	226	5	5	70	150	70	148	0.0007	60	-33	41	
1.A.2 Manufacturing Industries and Construction														
Liquid	CH_4	9	7	2	2	75	10	75	12	0.0000	-19	-21	27	
	N_2O	39	41	2	2	75	10	75	12	0.0000	4	-25	30	

TABLE 3.5 (CONTINUED)
EXAMPLE OF REPORTING OF APPROACH 2 UNCERTAINTY ANALYSIS USING GENERAL REPORTING TABLE FOR UNCERTAINTY
Emissions, removals and uncertainties are from National Inventory of Finland for year 2003 (Statistics Finland, 2005). The aggregation level and uncertainty estimates are country-specific for Finland, and do not represent recommended uncertainties or level of aggregation for other countries.

A	B	C	D	E		F		G		H	I	J		K
IPCC category	**Gas**	**Base year emissions or removals**	**Year *t* emissions or removals**	**Activity data uncertainty**		**Emission factor uncertainty**		**Combined uncertainty**		**Contribution to variance in year *t*[a]**	**Inventory trend in national emissions for year *t* increase with respect to base year**	**Uncertainty introduced into the trend in total national emissions with respect to base year**		**Approach and Comments**
		$Gg\ CO_2$ equivalent	$Gg\ CO_2$ equivalent	(-) %	(+) %	(-) %	(+) %	(-) %	(+) %	(fraction)	(% of base year)	(-) %	(+) %	Approach 2
Solid	CH_4	4	2	2	2	75	10	74	12	0.0000	-44	-13	20	
	N_2O	108	90	2	2	50	50	50	50	0.0000	-17	-11	12	
Gas	CH_4	5	6	1	1	75	10	75	11	0.0000	35	-35	45	
	N_2O	17	19	1	1	50	50	50	50	0.0000	13	-14	16	
Biomass	CH_4	20	19	15	15	50	50	51	53	0.0000	-7	-20	26	
	N_2O	111	81	15	15	70	150	70	151	0.0001	-28	-20	27	
Peat	CH_4	4	3	5	5	50	50	50	50	0.0000	-29	-9	11	
	N_2O	56	29	5	5	70	150	70	150	0.0000	-49	-11	14	
1.A.3 Transport														
a. Civil Aviation	CH_4	0.4	0.3	5	5	57	100	57	100	0.0000	-12	-12	15	
	N_2O	4	4	5	5	70	150	70	148	0.0000	-1	-17	21	
b. Road Transportation														
Gasoline	CH_4	78	40	1	1	50	50	50	50	0.0000	-49	-6	6	
Cars with Catalytic Converters	N_2O	32	410	1	1	94	378	94	392	0.0174	1 176	-446	643	
Cars without Catalytic Converters	N_2O	59	22	1	1	86	259	86	259	0.0000	-63	-11	16	
Diesel	CH_4	12	6	1	1	50	50	50	50	0.0000	-51	-5	5	
	N_2O	68	84	1	1	99	158	99	157	0.0001	23	-59	94	
Natural gas	CH_4		2	1	1	50	50	49	50					
	N_2O		0.0	1	1	70	150	70	149					
c. Railways	CH_4	0.2	0.2	5	5	60	110	60	110	0.0000	-30	-11	13	
	N_2O	2	1	5	5	70	150	70	149	0.0000	-30	-13	17	

TABLE 3.5 (CONTINUED)
EXAMPLE OF REPORTING OF APPROACH 2 UNCERTAINTY ANALYSIS USING GENERAL REPORTING TABLE FOR UNCERTAINTY
Emissions, removals and uncertainties are from National Inventory of Finland for year 2003 (Statistics Finland, 2005). The aggregation level and uncertainty estimates are country-specific for Finland, and do not represent recommended uncertainties or level of aggregation for other countries.

A	B	C	D	E		F		G		H	I	J		K
IPCC category	**Gas**	**Base year emissions or removals**	**Year *t* emissions or removals**	**Activity data uncertainty**		**Emission factor uncertainty**		**Combined uncertainty**		**Contribution to variance in year *t*[a]**	**Inventory trend in national emissions for year *t* increase with respect to base year**	**Uncertainty introduced into the trend in total national emissions with respect to base year**		**Approach and Comments**
		Gg CO_2 equivalent	Gg CO_2 equivalent	(-) %	(+) %	(-) %	(+) %	(-) %	(+) %	(fraction)	(% of base year)	(-) %	(+) %	Approach 2
d. Water-borne Navigation														
Residual Oil & Gas/Diesel Oil	CH_4	1	1	10	10	57	100	57	99	0.0000	2	-19	22	
	N_2O	2	3	10	10	70	150	70	149	0.0000	36	-30	39	
Gasoline	CH_4	7	4	20	20	57	100	59	104	0.0000	-42	-16	22	
	N_2O	0.4	1	20	20	70	150	71	154	0.0000	56	-49	71	
e. Other Transportation														
Gasoline & Diesel	CH_4	5	6	30	30	50	50	54	63	0.0000	15	-43	67	
Gasoline	N_2O	1	1	30	30	70	150	72	156	0.0000	9	-41	67	
Diesel	N_2O	4	4	30	30	70	150	72	158	0.0000	-5	-37	60	
1.A.4 Other Sectors														
Liquid	CH_4	19	15	3	3	75	10	74	13	0.0000	-19	-18	20	
	N_2O	56	47	3	3	75	10	76	13	0.0000	-15	-21	25	
Solid	CH_4	2	1	10	10	75	10	76	20	0.0000	-72	-6	8	
	N_2O	0.5	0.3	10	10	50	50	51	52	0.0000	-27	-12	14	
Gas	CH_4	0.1	0.3	5	5	75	10	75	15	0.0000	132	-49	62	
	N_2O	1	1	5	5	50	50	50	50	0.0000	124	-27	32	
Biomass	CH_4	282	307	15	15	70	150	71	151	0.0013	9	-28	38	
	N_2O	56	61	15	15	70	150	71	150	0.0000	9	-28	38	
Peat	CH_4	1	1	25	25	50	50	53	60	0.0000	1	-32	46	
	N_2O	1	2	25	25	70	150	71	155	0.0000	13	-38	57	
1.A.5 Non-Specified														
Liquid	CH_4	2	2	7	7	75	10	75	17	0.0000	43	-31	46	
	N_2O	6	9	7	7	75	10	75	17	0.0000	45	-33	43	

TABLE 3.5 (CONTINUED)
EXAMPLE OF REPORTING OF APPROACH 2 UNCERTAINTY ANALYSIS USING GENERAL REPORTING TABLE FOR UNCERTAINTY
Emissions, removals and uncertainties are from National Inventory of Finland for year 2003 (Statistics Finland, 2005). The aggregation level and uncertainty estimates are country-specific for Finland, and do not represent recommended uncertainties or level of aggregation for other countries.

A	B	C	D	E		F		G		H	I	J		K
IPCC category	**Gas**	**Base year emissions or removals**	**Year *t* emissions or removals**	**Activity data uncertainty**		**Emission factor uncertainty**		**Combined uncertainty**		**Contribution to variance in year *t*[a]**	**Inventory trend in national emissions for year *t* increase with respect to base year**	**Uncertainty introduced into the trend in total national emissions with respect to base year**		**Approach and Comments**
		Gg CO_2 equivalent	Gg CO_2 equivalent	(-) %	(+) %	(-) %	(+) %	(-) %	(+) %	(fraction)	(% of base year)	(-) %	(+) %	Approach 2
Gas	CH_4	0.3	0.4	13	13	75	10	75	23	0.0000	64	-41	55	
	N_2O	1	2	13	13	50	50	51	52	0.0000	64	-31	37	
1.B Fugitive Emissions from Fuels														
1.B.2 Oil and Natural Gas														
a.ii Oil – Flaring	CO_2	123	63					50	50	0.0000	-49	-29	85	b
a.iii.4 Oil – Refining	CH_4	8	10	2	2	90	90	90	90	0.0000	27	-41	53	
b.iii.4 Natural Gas - Transmission and storage	CH_4	4	12					3	3	0.0000	236	-113	334	b
b.iii.5 Natural Gas - Distribution	CH_4	0	40					5	5	0.0000				b,c
2 Industrial Processes														
2.A.1 Cement Production	CO_2	786	500	2	2	5	5	5	5	0.0000	-36	-2	2	
2.A.2 Lime Production	CO_2	383	513	2	2	3	3	4	4	0.0000	34	-4	4	
2.A.3 and 2.A.4 Limestone and Dolomite Use	CO_2	99	148	4	7	9	5	10	10	0.0000	50	-13	14	d
2.A.3 and 2.A.4 Soda Ash Use	CO_2	18	20	4	7	2	1	5	7	0.0000	10	-9	10	d
2.B.2 Nitric Acid Production	N_2O	1 595	1 396	5	5	57	100	57	100	0.0126	-13	-7	8	
2.B.8.b Ethylene	CH_4	4	5	5	5	20	20	20	21	0.0000	32	-9	10	
2.B.10 Other	CO_2	60	147	8	12	5	5	10	13	0.0000	145	-35	40	
2.C.1 Iron and Steel Production	CH_4	5	9	3	3	20	20	20	20	0.0000	85	-8	8	
2.D Non-Energy Products from Fuels and Solvent Use	CO_2	640	830	50	50	5	5	50	50	0.002	30	-71	156	
2.F.1 Refrigeration and Air Conditioning	HFCs, PFCs	0	578					11	26	0.0001	4 584 122	-519 745	1 206 234	b
2.F.2 Foam Blowing Agents	HFCs		25					24	24	0.0000				b,c
2.F.4 Aerosols	HFCs		63					10	10	0.0000				b,c

TABLE 3.5 (CONTINUED)
EXAMPLE OF REPORTING OF APPROACH 2 UNCERTAINTY ANALYSIS USING GENERAL REPORTING TABLE FOR UNCERTAINTY
Emissions, removals and uncertainties are from National Inventory of Finland for year 2003 (Statistics Finland, 2005). The aggregation level and uncertainty estimates are country-specific for Finland, and do not represent recommended uncertainties or level of aggregation for other countries.

A	B	C	D	E		F		G		H	I	J		K
IPCC category	**Gas**	**Base year emissions or removals**	**Year *t* emissions or removals**	**Activity data uncertainty**		**Emission factor uncertainty**		**Combined uncertainty**		**Contribution to variance in year *t*[a]**	**Inventory trend in national emissions for year *t* increase with respect to base year**	**Uncertainty introduced into the trend in total national emissions with respect to base year**		**Approach and Comments**
		Gg CO_2 equivalent	Gg CO_2 equivalent	(-) %	(+) %	(-) %	(+) %	(-) %	(+) %	(fraction)	(% of base year)	(-) %	(+) %	Approach 2
2.G.1 Electrical Equipment	SF_6	87	22					88	88	0.0000	-75	-22	41	b
2.G.3.a Medical Applications	N_2O	62	40	30	30	20	20	34	38	0.0000	-36	-23	35	
2.H.3 Other (grouped data of f-gases)	HFCs, PFCs, SF_6	8	21					38	38	0.0000	164	-123	292	b
3 AFOLU														
3.A.1 Enteric Fermentation	CH_4	1 868	1 537					20	31	0.0015	-18	-3	3	b
3.A.2 Manure Management	CH_4	215	222					16	16	0.0000	3	-4	5	b
3.A.2 Manure Management	N_2O	623	461					83	27	0.0006	-26	-15	17	b
3.B.1.a Forest Land remaining Forest Land														
carbon stock change in biomass	CO_2	-2 3798	-2 1354					35	35	0.5662	-10	-19	25	b
3.B.2.a Cropland Remaining Cropland														
net carbon stock change in mineral soils	CO_2	-535	-1 113					99	101	0.0125	108	-242	393	b
net carbon stock change in organic soils	CO_2	1 813	1 324	20	20	90	90	89	95	0.0152	-27	-32	54	
3.B.3.a Grassland Remaining Grassland														
net carbon stock change in mineral soils	CO_2	-1 181	2 907					99	100	0.0852	-346	-2223	1067	b
net carbon stock change in organic soils	CO_2	109	67	30	30	90	90	90	103	0.0000	-39	-29	50	
3.B.4.ai Peatlands Remaining Peatlands	CO_2	503	547	15	15	80	208	80	212	0.0074	9	-32	48	
3.A.4.ai Peatlands Remaining Peatlands	CH_4	5	6	15	15	80	208	80	208	0.0000	6	-32	46	
3.C.1.a Biomass Burning in Forest Lands	CO_2	180	91	10	10	70	70	71	71	0.0000	-50	-12	15	
3.C.1.a Biomass Burning in Forest Lands	CH_4	16	8	10	10	70	70	70	71	0.0000	-49	-12	15	
3.C.1.a Biomass Burning in Forest Lands	N_2O	2	1	10	10	70	70	70	72	0.0000	-50	-11	15	
3.C.2 Liming	CO_2	618	277	20	20	20	3	25	22	0.0000	-55	-11	15	

TABLE 3.5 (CONTINUED)
EXAMPLE OF REPORTING OF APPROACH 2 UNCERTAINTY ANALYSIS USING GENERAL REPORTING TABLE FOR UNCERTAINTY
Emissions, removals and uncertainties are from National Inventory of Finland for year 2003 (Statistics Finland, 2005). The aggregation level and uncertainty estimates are country-specific for Finland, and do not represent recommended uncertainties or level of aggregation for other countries.

A	B	C	D	E		F		G		H	I	J		K
IPCC category	**Gas**	**Base year emissions or removals**	**Year *t* emissions or removals**	**Activity data uncertainty**		**Emission factor uncertainty**		**Combined uncertainty**		**Contribution to variance in year *t*[a]**	**Inventory trend in national emissions for year *t* increase with respect to base year**	**Uncertainty introduced into the trend in total national emissions with respect to base year**		**Approach and Comments**
		Gg CO_2 equivalent	Gg CO_2 equivalent	(-) %	(+) %	(-) %	(+) %	(-) %	(+) %	(fraction)	(% of base year)	(-) %	(+) %	Approach 2
3.C.4 Direct N_2O Emissions from Managed Soils: Agricultural Soils	N_2O	3 486	2 608					76	227	0.2170	-25	-19	29	b
3.C.4 Direct N_2O Emissions from Managed Soils: N Fertilizers Application, forest land	N_2O	27	11	10	10	94	380	94	386	0.0000	-58	-17	32	
3.C.5 Indirect N_2O Emissions from Managed Soils	N_2O	735	592					81	334	0.0303	-19	-19	25	b
4 Waste														
4.A Solid Waste Disposal	CH_4	3 678	2 497					43	43	0.012	-32	-14	16	b
4.D.1 Domestic Wastewater Treatment and Discharge														
sparcely populated areas	CH_4	118	95	15	15	32	20	34	27	0.000	-20	-16	20	
densely populated areas	CH_4	12	13					60	109	0.000	9	-16	20	b
sparcely populated areas	N_2O	21	18	10	10	94	380	94	378	0.000	-13	-29	40	
densely populated areas	N_2O	84	66	5	5	94	380	94	378	0.000	-21	-25	34	
4.D.2 Industrial Wastewater Treatment and Discharge	CH_4	22	19					61	109	0.000	-15	-17	22	b
4.D.2 Industrial Wastewater Treatment and Discharge	N_2O	28	17	5	5	94	380	94	388	0.000	-37	-19	27	
4.E Other: N input from Fish Farming	N_2O	8	3	10	10	94	380	94	391	0.000	-62	-12	18	
Total		47 604	67 730					14	15		42	-18	23	

[a] Entries to Column H are obtained by dividing the variance of each category (obtained from the Monte Carlo simulation tool) by the total variance of the inventory.

[b] A more complex method for estimation of uncertainties is used, and therefore activity data and emission factor uncertainties are left blank. The resulting uncertainty is shown in Column G

[c] Trend not calculated, when base year emissions are zero

[d] Uncertainty assessment was made in the aggregation level used by Finland in 2003 inventory, and therefore glass production could not be separated.

3.7 TECHNICAL BACKGROUND INFORMATION

3.7.1 Approach 1 variables and equations

This section covers the basis for the statistical calculation methods used in Approach 1 as complementary information to Section 3.2.3.1, Approach 1: Propagation of Error, and Table 3.2, Approach 1 Uncertainty Calculation. Key variables and equations used for the calculation are defined in this section.

Explanation of the variables

$C_{x,}$ = Value of an entry in Column C and row *x*, emissions or removals of each category of base year inventory

$\sum C_i$ = Sum of emissions and removals over all categories (rows) of the base year inventory

$D_{x,}$ = Value of an entry in Column D and row *x*, emissions or removals of each category of the year of inventory *t*

$\sum D_i$ = Sum of emissions and removals over all categories (rows) of the year of inventory *t*

Column A-F

Input data of emissions and removals, activity data and emission factor uncertainties of each category

Column G

Combined uncertainty using error propagation equation. See Equation 3.1 in Section 3.2.3.1.

$$G_x = \sqrt{E_x^{\,2} + F_x^{\,2}}$$

Column H

Contribution to uncertainty. See also Equation 3.2 in Section 3.2.3.1.

$$H_x = \frac{(G_x \bullet D_x)^2}{(\sum D_i)^2}$$

The total emission uncertainty is obtained using the error propagation equation:

$$\frac{\sqrt{\sum (G_i \bullet D_i)^2}}{\sum D_i} = \sqrt{\sum H_i}$$

Column I

Entries in Column I show how the difference in emissions between the base year and the year *t* changes in response to a 1 percent increase in emissions of category *x* emissions in the base year and year *t*. This shows the sensitivity of the trend in emissions to a systematic uncertainty in the emission estimate – i.e., one that is correlated between the base year and year *t*. This sensitivity is described as Type A sensitivity.

I_x = percentage trend if category *x* is increased by 1 percent in both years – percentage trend without increase

$$= \frac{0.01 \bullet D_x + \sum D_i - (0.01 \bullet C_x + \sum C_i)}{(0.01 \bullet C_x + \sum C_i)} \bullet 100 - \frac{\sum D_i - \sum C_i}{\sum C_i} \bullet 100$$

Column J

Entries in Column J show how the difference in emissions between the base year and year *t* changes in response to a 1 percent increase in the emissions of category *x* in year *t* only. This shows the sensitivity of the trend in emissions to random uncertainty error in the emissions estimate – i.e., one that is not correlated between the base year and year Y. This sensitivity is described as Type B sensitivity.

J_x = percentage trend if category *x* is increased by 1 percent in year *t* – percentage trend without increase

$$= \frac{0.01 \bullet D_x + \sum D_i - \sum C_i}{\sum C_i} \bullet 100 - \frac{\sum D_i - \sum C_i}{\sum C_i} \bullet 100$$

$$= \frac{D_x}{\sum C_i}$$

Column K
Under the assumption that the same emission factor is used in both years and the actual emission factors are fully correlated, the percent error introduced by it is equal in both years. Therefore the formula for the uncertainty introduced on the trend by the emission factor is:

$$K_x = \text{sensitivity A} \bullet \text{uncertainty of emission factor}$$
$$= I_x \bullet F_x$$

In case no correlation between emission factors is assumed, sensitivity B should be used and the result needs to be increased by √2 for the reason given below in the main derivation for Column L:

$$K_x = \text{sensitivity B} \bullet \text{uncertainty of emission factor} \bullet \sqrt{2}$$
$$= J_x \bullet F_x \bullet \sqrt{2}$$

Column L
The trend is the difference between the emissions in the base year and in the year *t*. Therefore the uncertainty of the activity data of the base year and year *t* has to be taken into account. The two uncertainties combined using the error propagation equation and the assumption that the uncertainty is the same in the base year and year *t* is:

$$= \sqrt{(\text{uncertainty (activity data, base year)})^2 + (\text{uncertainty (activity data, year } t))^2}$$
$$\approx \sqrt{(\text{uncertainty (activity data, year } t))^2 \bullet 2}$$
$$= E_x \bullet \sqrt{2}$$

Since activity data in both years are assumed to be independent, Column L equals:

L_x = sensitivity B • combined uncertainty of activity data of both years

$$= J_x \bullet E_x \bullet \sqrt{2}$$

In case correlation between activity data is assumed, sensitivity A should be used and the √2 factor does not apply.

$$L_x = I_x \bullet E_x$$

Column M
Uncertainty introduced on the trend by the uncertainty in the activity data and the emissions factor.

$$M_x = K_x^2 + L_x^2$$

The entries M_i in Column M are combined to obtain the total uncertainty of the trend using the error propagation equation as follows:

$$\text{Total uncertainty of the trend} = \sqrt{\sum M_i}$$

3.7.2 Approach 1 – details of the equations for trend uncertainty

The following steps show how to calculate trend uncertainty using Types A and B sensitivities (see also Section 3.2.3.1).

1) The method for assessing level uncertainty in year Y assumes that categories and gases are uncorrelated, or are aggregated until the aggregated categories can be treated as uncorrelated.

2) The uncertainty in the trend in total emissions from the country (the quantity at the foot of Column M) is estimated as:

$$U_T = \sqrt{\sum_{i=1}^{N} U_i^2}$$

where U_T is the uncertainty in the trend in total emissions from the country and U_i is the uncertainty introduced into U_T by the category *i* and gas.

3) We take

$$U_i = \sqrt{(U_{E,i}^2 + U_{A,i}^2)}$$

where $U_{E,i}$ is the uncertainty introduced into U_i by the uncertainty associated with the emission factor of the category i and gas, and $U_{A,i}$ is the uncertainty introduced into U_i by the uncertainty associated with the activity data of the category i and gas.

4) We know from Columns E and F what the uncertainties related to activity data and emission factors for the category i and gas are in percentage terms, but we do not yet know how these uncertainties affect the trend in the total emissions, which is what we need for $U_{E,i}$ and $U_{A,i}$. For this we write

$$U_{E,i} = A_i u_{e,i} \qquad \text{and} \qquad U_{A,i} = B_i u_{a,i}$$

Where A_i is the Type A sensitivity associated with the category i and gas, and $u_{e,i}$ the percentage uncertainty associated with the emission factor in Column F, and B_i is the Type B sensitivity associated with the category i and gas, and $u_{a,i}$ the percentage uncertainty associated with the activity data in Column E. Essentially Type A and Type B sensitivities are elasticities relating respectively a percentage difference that is self-correlated between the base year and year Y, and one which is uncorrelated, to the percentage change in total emissions. The method allows for this assumption to be inverted, or for both emission factor and activity data to be self-correlated between years, or for neither to be self-correlated.

5) The Type A and Type B sensitivities are calculable from formulae for the trend in terms of sums over categories and gases in the base year and in year Y. The additional factor of √2 is introduced because an uncorrelated uncertainly might affect either the base year or the year Y. The current formulation assumes for Type B sensitivity that the emissions in year Y are not too different from those in the base year; if this were not the case we would have to introduce separate consideration of the base year and year Y for the uncorrelated uncertainties, rather than using the √2 factor.

DERIVATION OF TYPE A SENSITIVITY

The trend can be written as (assuming that 1990 is a base year):

$$100 \bullet \left(\frac{\sum_{i=1}^{N} e_{i,y} - \sum_{i=1}^{N} e_{i,1990}}{\sum_{i=1}^{N} e_{i,1990}} \right)$$

If the category i and gas is increased by 1 percent throughout (consistent with the assumption that Type A sensitivity captures the effect of uncertainties which are correlated between years) the trend becomes:

$$100 \bullet \left(\frac{\sum_{i=1}^{N} e_{i,y} + 0.01\, e_{i,y} - \left(\sum_{i=1}^{N} e_{i,1990} + 0.01\, e_{i,1990} \right)}{\sum_{i=1}^{N} e_{i,1990} + 0.01\, e_{i,1990}} \right)$$

and the sensitivity A_i becomes:

$$100 \bullet \left(\frac{\sum_{i=1}^{N} e_{i,y} + 0.01\, e_{i,y} - \left(\sum_{i=1}^{N} e_{i,1990} + 0.01\, e_{i,1990} \right)}{\sum_{i=1}^{N} e_{i,1990} + 0.01\, e_{i,1990}} \right) - 100 \bullet \left(\frac{\sum_{i=1}^{N} e_{i,y} - \sum_{i=1}^{N} e_{i,1990}}{\sum_{i=1}^{N} e_{i,1990}} \right)$$

This is the same as the expression given for the Type A sensitivity in Note B on page 6.18 of the *GPG2000*.

TYPE B SENSITIVITY

The Type B sensitivity we assume that the category i and gas is increased by 1 percent in year y only. In this case the trend becomes:

$$100 \bullet \left(\frac{\sum_{i=1}^{N} e_{i,y} + 0.01\, e_{i,y} - \sum_{i=1}^{N} e_{i,1990}}{\sum_{i=1}^{N} e_{i,1990}} \right)$$

So the sensitivity B_i becomes:

$$100 \bullet \left(\frac{\sum_{i=1}^{N} e_{i,y} + 0.01\, e_{i,y} - \sum_{i=1}^{N} e_{i,1990}}{\sum_{i=1}^{N} e_{i,1990}} \right) - 100 \bullet \left(\frac{\sum_{i=1}^{N} e_{i,y} - \sum_{i=1}^{N} e_{i,1990}}{\sum_{i=1}^{N} e_{i,1990}} \right)$$

All the terms on the numerator cancel out between the brackets except for $0.01\, e_{i,y}$ which becomes $e_{i,y}$ when multiplied by 100. So the expression for B_i simplifies to $e_{i,y} \Big/ \sum_{i=1}^{N} e_{i,1990}$ which is the expression at the top of Column J on page 6.16 of the *GPG2000*.

3.7.3 Dealing with large and asymmetric uncertainties in the results of Approach 1

This section provides guidance on how to correct for biases in large estimates of uncertainty from Approach 1 and how to convert the uncertainty ranges into asymmetric 95 percent probability ranges based upon a lognormal distribution.

Correction of uncertainty estimate for large uncertainties: The approximate error propagation method of Approach 1 produces an estimate of the uncertainty half range (U), expressed as a percentage relative to the mean, of the inventory results. As the uncertainty in the total inventory uncertainty becomes larger, the error propagation approach systematically underestimates the uncertainty unless the model is purely additive. However, most inventories are estimated based upon the sum of terms, each of which is a product (e.g., of emission factors and activity data). The error propagation approach is not exact for such multiplicative terms. Results from empirical studies show that in some cases uncertainty estimated using Approach 1 could be underestimated; the analyst could use a correction factor, for example that proposed in Frey (2003). Frey (2003) evaluated the performance of an analytical approach for combining uncertainty in comparison to a Monte Carlo simulation with large sample sizes for many cases involving different ranges of uncertainty for additive, multiplicative, and quotient models. Error propagation and Monte Carlo simulated estimates of the uncertainty half-range of the model output agreed well for values of less than 100 percent. As the uncertainty in the total inventory increased to higher levels, there was a systematic under-estimation of uncertainty in the total inventory by the error propagation approach. The relationship between the simulated and error propagation estimates was found to well-behaved. Thus, a correction factor was developed from the comparison that is applicable if U for the total inventory uncertainty is large (e.g., greater than 100 percent) and is given by:

EQUATION 3.3
CORRECTION FACTOR FOR UNCERTAINTY HALF-RANGE

$$F_C = \left[\frac{\left(-0.720 + 1.0921U - 1.63 \bullet 10^{-3} U^2 + 1.11 \bullet 10^{-5} U^3 \right)}{U} \right]^2$$

Note: Use if $U > 100\%$ and if the model contains multiplicative or quotient terms

Not necessarily reliable for $U > 230\%$

Not necessary for models that are purely additive.

Where:

U = ½-range for uncertainty estimated from error propagation, in units of percent

F_c = Correction factor for analytical estimate of the variance, dimensionless ratio of corrected to uncorrected uncertainty

The empirically-based correction factor produces values from 1.06 to 1.69 as *U* varies from 100% to 230%. The correction factor is used to develop a new, corrected, estimate of the total inventory uncertainty half-range, $U_{corrected}$, which in turn is used to develop confidence intervals.

EQUATION 3.4
CORRECTED UNCERTAINTY HALF-RANGE

$$U_{corrected} = U \bullet F_C$$

Where:

$U_{corrected}$ = Corrected ½-range for uncertainty estimated from error propagation, in units of %

The errors in the analytical estimate of the variance are generally small for uncertainty half-ranges (*U*) of less than approximately 100 percent. If the correction factor is applied for *U* > 100% for values of *U* up to 230%, the typical error in the estimate of *U* is expected to be within plus or minus 10 percent in most cases. The correction factor will not necessarily be reliable for larger uncertainties because it was calibrated over the range of 10% to 230%.

Calculation of asymmetric confidence intervals for large uncertainties: In order to calculate confidence intervals for the model output based upon only the mean and half-range for uncertainty, a distribution must be assumed. For models that are purely additive, and for which the half range of uncertainty is less than approximately 50 percent, a normal distribution is often an accurate assumption for the form of the model output. In this case, a symmetric uncertainty range with respect to the mean can be assumed. For multiplicative models, or when the uncertainty is large for a variable that must be non-negative, a lognormal distribution is typically an accurate assumption for the form of the model output. In such cases, the uncertainty range is not symmetric with respect to the mean, even though the variance for the total inventory may be correctly estimated from Approach 1. Here, we provide a practical methodology for calculating approximate asymmetric uncertainty ranges based upon the results of error propagation, based upon a methodology developed by Frey (2003). A key characteristic of the 95 percent confidence intervals is that they are approximately symmetric for small ranges of uncertainty and they are positively skewed for large ranges of uncertainty. The latter result is necessary for a non-negative variable.

The parameters of the lognormal distribution can be defined in several ways, such as in terms of the geometric mean and geometric standard deviation. The geometric mean can be estimated based upon the arithmetic mean and the arithmetic standard deviation:

EQUATION 3.5
ASYMMETRIC CONFIDENCE INTERVALS – GEOMETRIC MEAN

$$\mu_g = \exp\left\{\ln(\mu) - \frac{1}{2}\ln\left(1 + \left[\frac{U}{200}\right]^2\right)\right\}$$

Where:

μ_g = geometric mean

μ = arithmetic mean

The geometric standard deviation is given by:

EQUATION 3.6
ASYMMETRIC CONFIDENCE INTERVALS – GEOMETRIC STANDARD DEVIATION

$$\sigma_g = \exp\left\{\sqrt{\ln\left(1 + \left[\frac{U}{200}\right]^2\right)}\right\}$$

Where:

σ_g = geometric standard deviation

A confidence interval can be estimated based upon the geometric mean, geometric standard deviation, and the inverse cumulative probability distribution of a standard normal distribution (with a logarithmic transformation):

EQUATION 3.7
LOWER/UPPER UNCERTAINTY HALF-RANGE FROM ERROR PROPAGATION

$$U_{low} = \left(\frac{\exp\left\{\ln\left(\mu_g\right) - 1.96\ln\left(\sigma_g\right)\right\} - \mu}{\mu} \right) \times 100$$

$$U_{high} = \left(\frac{\exp\left\{\ln\left(\mu_g\right) + 1.96\ln\left(\sigma_g\right)\right\} - \mu}{\mu} \right) \times 100$$

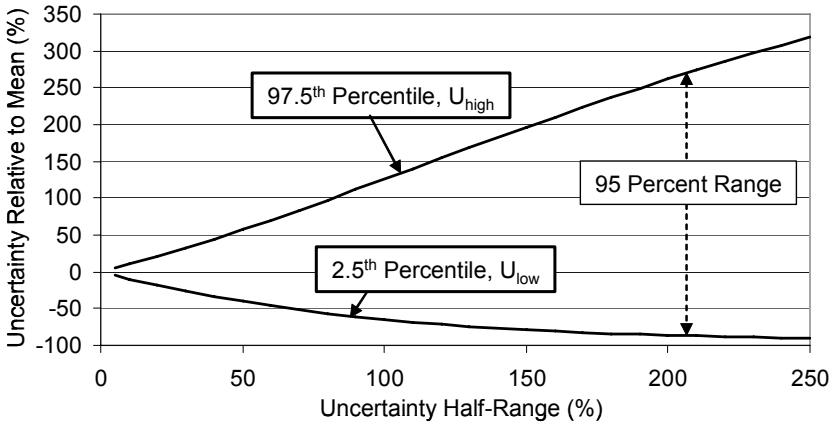
Where:

U_{low} = Lower ½-range for uncertainty estimated from error propagation, in units of %.

U_{high} = Upper ½-range for uncertainty estimated from error propagation, in units of %.

To illustrate the use of these equations, consider an example. Suppose the mean is 1.0 and the ½-range of uncertainty estimated from error propagation is 100 percent. In this case, the geometric mean is 0.89 and the geometric standard deviation is 1.60. The 95 percent probability range as a percentage relative to the mean is given by the interval from U_{low} to U_{high} of Equations 3.7. In the example, the result is -65% to +126%. In contrast, if a normal distribution had been used as the basis for uncertainty estimation, the range would have been estimated as approximately ±100% and there would be a probability of approximately two percent of obtaining negative values. Figure 3.9 illustrates the sensitivity of the lower and upper bounds of the 95 percent probability range, which are the 2.5th and 97.5th percentiles, respectively, calculated assuming a lognormal distribution based upon an estimated uncertainty half-range from an error propagation approach. The uncertainty range is approximately symmetric relative to the mean up to an uncertainty half-range of approximately 10 to 20 percent. As the uncertainty half-range, *U*, becomes large, the 95 percent uncertainty range shown in Figure 3.9 becomes large and asymmetric. For example, if *U* is 73 percent, then the estimated probability range is approximately -50% to +100%, or a factor of two.

Figure 3.9 Estimates of asymmetric ranges of uncertainty with respect to the arithmetic mean assuming a lognormal distribution based upon uncertainty half-range calculated from a propagation of error approach



3.7.4 Methodology for calculation of the contribution to uncertainty

The methodology for calculation of contribution to uncertainty is based upon apportioning the variance of the inventory to the variance of each category.

If the uncertainty is symmetric, then the variance is estimated, on a category basis, as:

EQUATION 3.8
CONTRIBUTION OF CATEGORY X – VARIANCE FOR SYMMETRIC UNCERTAINTY

$$\sigma_x^2 = \left(D_x \frac{U_x}{200} \right)^2$$

Where:

U_x = uncertainty half-range for category *x*, in units of percent;

D_x = the total emissions or removals for category *x*, corresponding to the entries in Column D of Table 3.5.

σ_x^2 = the variance of emissions or removals for category *x*.

Even if the uncertainty is asymmetric, the variance can be estimated based on the arithmetic standard deviation or the coefficient of variation. The variance is simply the square of the arithmetic deviation. The variance for the category can be estimated from the coefficient of variation, v_x, as:

EQUATION 3.9
CONTRIBUTION OF CATEGORY X – VARIANCE FOR ASYMMETRIC UNCERTAINTY

$$\sigma_x^2 = (D_x \upsilon_x)^2$$

Once the variance is known for a category, the variances should be summed over all categories. The result is the approximate total variance in the inventory. However, this result is not likely to agree exactly with a Monte Carlo simulation result for the inventory for at least one and possibly more reasons: (1) because of sample fluctuations in the Monte Carlo simulation, the Monte Carlo estimate of the variance may differ somewhat from the true value; (2) the analytical calculation is based upon assumptions of normality or lognormality of the distributions for combined uncertainty for individual categories, whereas Monte Carlo simulation can accommodate a wide variety of distribution assumptions; and (3) the Monte Carlo simulation may account for nonlinearities and dependencies that are not accounted for in the analytical calculation for contribution to variance. If the emission inventory calculations are linear or approximately linear, without any substantial correlations, then the results should agree fairly well. Furthermore, methods for estimating 'contribution to variance' for Monte Carlo methods are approximate. For those methods that potentially can account for all contributions to variance (e.g., Sobol's method, Fourier Amplitude Sensitivity Test), the measures of sensitivity are more complex (e.g., Mokhtari *et al*., 2006). Thus, the methodology described here is a practical compromise.

References

Abdel-Aziz, A., and Frey, H.C. (2003). 'Development of Hourly Probabilistic Utility NO_x Emission Inventories Using Time Series Techniques: Part I-Univariate Approach', *Atmospheric Environment*, 37:5379-5389 (2003).

Ang, A. H-S., and Tang, W.H., (1984). *Probability Concepts in Engineering Planning and Design, Volume 2: Decision, Risk, and Reliability*. John Wiley and Sons, New York .

Ang, A. H-S., and Tang, W.H., (1975). *Probability Concepts in Engineering Planning and Design, Volume 1.* John Wiley and Sons, New York.

Baggott, S.L., Brown, L., Milne, R., Murrells, TP., Passant, N., Thistlethwaite, G., Watterson, J.D. (2005) *"UK Greenhouse Gas Inventory, 1990 to 2003: Annual Report for submission under the Framework Convention on Climate Change",* April 2005. pub AEA Technology, UK ref AEAT/ENV/R/1971, ISBN 0-9547136-5-6.

Barry, T.M. (1996), Recommendations on the testing and use of pseudo-random number generators used in Monte Carlo analysis for risk assessment*, Risk Assessment*, 16(1):93-105.

Bevington, P.R. and Robinson, D.K. (1992). *Data Reduction and Error Analysis for the Physical Sciences*. McGraw-Hill: New York.

Cohen A.C. and Whitten B. (1998). *Parameter Estimation in Reliability and Life Span Models*, M. Dekker: New York.

Cullen, A.C. and Frey, H.C. (1999), *Probabilistic Techniques in Exposure Assessment: A Handbook for Dealing with Variability and Uncertainty in Models and Inputs*, Plenum: New York.

D'Agostino, R.B. and Stephens, M.A. (eds.) (1986). *Goodness-of-Fit Techniques*, Marcel Dekker, New York.

Efron, B. and Tibshirani, R.J. (1993). *An Introduction to the Bootstrap*, Chapman and Hall, New York.

Eggleston, S., *et al*. (1998). Treatment of Uncertainties for National Greenhouse Gas Emissions, Report AEAT 2688-1 for DETR Global Atmosphere Division, AEA Technology, Culham, UK.

Evans, J.S., Graham J.D., Gray, G.M., and Sielken Jr, R.L. (1994). "A Distributional Approach to Characterizing Low-Dose Cancer Risk," *Risk Analysis*, 14(1):25-34 (February 1994).

Falloon, P. and Smith, P. (2003). Accounting for changes in soil carbon under the Kyoto Protocol: need for improved long-term data sets to reduce uncertainty in model projections. *Soil Use and Management*, 19, 265-269.

Frey, H.C. and Rubin, E.S. (1991). *Development and Application of a Probabilistic Evaluation Method for Advanced Process Technologies*, Final Report, DOE/MC/24248-3015, NTIS DE91002095, Prepared by Carnegie-Mellon University for the U.S. Department of Energy, Morgantown, West Virginia, April 1991, 364p.

Frey, H.C. and Rhodes, D.S. (1996). "Characterizing, Simulating, and Analyzing Variability and Uncertainty: An Illustration of Methods Using an Air Toxics Emissions Example," *Human and Ecological Risk Assessment: an International Journal*, 2(4):762-797 (December 1996).

Frey, H.C. and Bammi, S. (2002). Quantification of Variability and Uncertainty in Lawn and Garden Equipment NO_x and Total Hydrocarbon Emission Factors, *J. Air & Waste Manage. Assoc*., 52(4), 435-448.

Frey, H.C., Zheng, J., Zhao, Y., Li, S., and Zhu, Y. (2002). Technical Documentation of the AuvTool Software for Analysis of Variability and Uncertainty, Prepared by North Carolina State University for the Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, NC. February 2002.

Frey, H.C. and Zheng, J. (2002). "Probabilistic Analysis of Driving Cycle-Based Highway Vehicle Emission Factors," *Environmental Science and Technology,* 36(23):5184-5191 (December 2002).

Frey, H.C. (2003), "Evaluation of an Approximate Analytical Procedure for Calculating Uncertainty in the Greenhouse Gas Version of the Multi-Scale Motor Vehicle and Equipment Emissions System," Prepared for Office of Transportation and Air Quality, U.S. Environmental Protection Agency, Ann Arbor, MI, May 30, 2003.

Frey, H.C. (2005), "Comparison of Approach 1 and Approach 2," January 2005, unpublished analysis done for this Chapter.

Gelfand, A. E. (1996). *Gibbs Sampling, The Encyclopedia of Statistical Sciences* (editors: Kotz J., Reed C. and Banks D.), John Wiley and Sons, New York, 283-292.

Hahn, G.J., and Shapiro, S.S. (1967) *Statistical Models in Engineering*, Wiley Classics Library, John Wiley and Sons, New York.

Holland, D.M and Fitz-Simons, T. (1982) "Fitting statistical distributions to air quality data by the maximum likelihood method," *Atmospheric Environment*, 16(5):1071-1076.

Hora, S.C. and Iman, R.L. (1989). Expert opinion in risk analysis: The NUREG-1150 methodology, *Nuclear Science and Engineering*, 102:323-331.

IPCC (1997). Houghton, J.T., Meira Filho, L.G., Lim, B., Tréanton, K., Mamaty, I., Bonduki, Y., Griggs, D.J. and Callander, B.A. (Eds). *Revised 1996 IPCC Guidelines for National Greenhouse Inventories*. Intergovernmental Panel on Climate Change (IPCC), IPCC/OECD/IEA, Paris, France.

IPCC (2000). Penman, J., Kruger, D., Galbally, I., Hiraishi, T., Nyenzi, B., Emmanuel, S., Buendia, L., Hoppaus, R., Martinsen, T., Meijer, J., Miwa, K., and Tanabe, K. (Eds). *Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories*. Intergovernmental Panel on Climate Change (IPCC), IPCC/OECD/IEA/IGES, Hayama, Japan.

ISO (1993). "Guide to the Expression of Uncertainty in Measurement (GUM)" prepared by ISO, IEC, BIPM, IFCC, OIML, IUPAC, IUPAP and published by ISO, Switzerland in 1993.

Kirchner, T.B. (1990). Establishing modeling credibility involves more than validation, Proceedings, On the Validity of Environmental Transfer Models, Biospheric Model Validation Study, Stockholm, Sweden, October 8-10.

Manly, B.F.J. (1997). *Randomization, Bootstrap, and Monte Carlo Methods in Biology, Second Edition*, Chapman and Hall.

McCann, T.J. and Associates, and Nosal, M. (1994). Report to Environmental Canada Regarding Uncertainties in Greenhouse Gas Emission Estimates, Calgary, Canada.

Merkhofer, M.W. (1987). Quantifying judgmental uncertainty: Methodology, experiences, and insights, *IEEE Transactions on Systems, Man, and Cybernetics*. 17(5):741-752.

Mokhtari, A., Frey H.C. and Zheng J. (2006). "Evaluation and recommendation of sensitivity analysis methods for application to Stochastic Human Exposure and Dose Simulation (SHEDS) models," *Journal of Exposure Assessment and Environmental Epidemiology*, Accepted December 2, 2005, In press.

Monni, S., Syri, S. and Savolainen I. (2004). 'Uncertainties in the Finnish greenhouse gas emission inventory'.*Environmental Science and Policy* 7, pp.87-98.

Monte, L, Hakanson, L., Bergstrom, U., Brittain, J. and Heling, R. (1996). Uncertainty analysis and validation of environmental models: the empirically based uncertainty analysis. *Ecological Modelling*, 91, 139-152.

Morgan, M.G., and Henrion, M. (1990). *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*, Cambridge University Press, New York.

NARSTO (2005). Improving Emission Inventories for Effective Air Quality Management Across North America, NARSTO, June 2005.

NCRP (National Council on Radiation Protection and Measurements). (1996). A Guide for Uncertainty Analysis in Dose and Risk Assessments Related to Environmental Contamination, NCRP Commentary No. 14, Bethesda, MD.

Ogle, S.M., Breidt, F.J., Eve, M.D. and Paustian, K. (2003). Uncertainty in estimating land use and management impacts on soil organic carbon storage for U.S. agricultural lands between 1982 and 1997. *Global Change Biology* 9:1521-1542.

Smith, A.E, Ryan, P.B. and Evans J.S. (1992). The effect of neglecting correlations when propagating uncertainty and estimating the population distribution of risk, *Risk Analysis*, 12:467-474.

Spetzler, C.S., and von Holstein, S. (1975). Probability Encoding in Decision Analysis, *Management Science*, 22(3).

Statistics Finland. (2005). *Greenhouse gas emissions in Finland 1990-2003. National Inventory Report to the UNFCCC*, 27 May 2005.

USEPA (1996). Summary Report for the Workshop on Monte Carlo Analysis, EPA/630/R-96/010, Risk Assessment Forum, U.S. Environmental Protection Agency, Washington, DC.

USEPA (1997). Guiding Principles for Monte Carlo Analysis, EPA/630/R-97/001, Risk Assessment Forum. U.S. Environmental Protection Agency, Washington, DC.

USEPA (1999). Report of the Workshop on Selecting Input Distributions for Probabilistic Assessments, EPA/630/R-98/004, U.S. Environmental Protection Agency, Washington, DC, January 1999. http://www.epa.gov/ncea/input.htm

Wackerly, D.D., Mendenhall III, W. and Scheaffer, R.L. (1996). *Mathematical Statistics with Applications*, Duxbury Press: USA.

Winiwarter, W. and Rypdal K. (2001). "Assessing the uncertainty associated with national greenhouse gas emission inventories: a case study for Austria," *Atmospheric Environment*, 35(22):5425-5440.

Zhao, Y. and Frey, H.C. (2004a). "Development of Probabilistic Emission Inventory for Air Toxic Emissions for Jacksonville, Florida," *Journal of the Air & Waste Management Association*, 54(11):1405-1421.

Zhao, Y., and Frey, H.C. (2004b). "Quantification of Variability and Uncertainty for Censored Data Sets and Application to Air Toxic Emission Factors," *Risk Analysis*, 24(3):1019-1034 (2004).

Zheng, J. and Frey H.C. (2004). "Quantification of Variability and Uncertainty Using Mixture Distributions: Evaluation of Sample Size, Mixing Weights and Separation between Components," *Risk Analysis*, 24(3):553-571 (June 2004).