CHAPTER 3

UNCERTAINTIES

Contents

Equations

Figures

Tables

Boxes

3 UNCERTAINTIES

UPDATE OF VOLUME 1, CHAPTER 3 OF THE 2006 IPCC GUIDELINES.

Sections and box to be developed and provided in the Second-order Draft (SOD) are highlighted in yellow.

3.1 INTRODUCTION

This section elaborates on the importance on uncertainty assessment as a means of improving emission inventories over time and attempt at a step-by-step guidance to its implementation.

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

Elaboration of Section 3.1.1 of the *2006 IPCC Guidelines*.

Uncertainty assessment is at the core of the effort of compiling an inventory of anthropogenic emissions and removals of GHGs (GHG inventory) and to assess its evolution over time. Since the GPG2000 report, the IPCC has adopted the concept of "*Good Practice*" in developing a GHG inventory, defined as an inventory that "contain neither over- nor under-estimates so far as can be judged, and in which uncertainties are reduced as far as practicable".

The first notion that emerges from this concept is that it is impossible to eliminate uncertainty completely, leading to the immediate conclusion that for every value reported in an inventory there will exist an associated uncertainty. Knowledge of this uncertainty is an integral part of the inventory compilation effort.

The second notion that follows is that, as a priority, effort should focus on accuracy, ensuring that emissions and removals are neither over- nor under-estimated. In short, bias should be eliminated as far as can be judged. Figure 3.2 of the *2006 IPCC Guidelines* gives a good illustration of the difference between accuracy and precision clearly showing that a precise estimate is of limited value if it is not accurate.

The key word is "knowledge". Knowing the processes involved and the information available is key to quantify and reduce uncertainty. While variability is a characteristic of the process and cannot be eliminated, broad uncertainty is associated with lack of knowledge. Causes of uncertainty are described in chapter 3.1.5 of the *2006 IPCC Guidelines* and further discussed in section 3.1.5 of this report.

Uncertainty calculation is strongly linked to the methods used to estimate emissions and removals. Simple methods are based on the multiplication of activity data (AD) by an emission factor (EF). More generally, both AD and EF can be result of several different parameters (see section 3.2.3 for a discussion). For some complex systems, models are developed for their description, evaluation of emissions and calculation of uncertainty.

Regardless of the complexity of the approaches, uncertainty of the results is a function of the uncertainty of data (activity or emission factors) used to compile the inventory. Hence, data collection and uncertainty evaluation are strongly linked. In short, every collected data value need to have an associated uncertainty assessment (further discussed in section 3.2).

Finally, it is important to point that producing an uncertainty analysis result (level or trend) is not a goal per se. The uncertainty values are not an absolute measure of the overall quality of the inventory. Even if they depend on the level of the complexity of the estimation methods and uncertainty calculation approaches, they are also a function of the share of sectors and categories in each country. Moreover, the uncertainty analysis as a whole is an important tool in the process of improvement of the inventory. Together with the *key category* analysis, it helps the inventory compilers in prioritizing the improvements in methodology development and data collection for the different source and sink categories (see section 3.1.2).

3.1.2 Uncertainty assessment as part of inventory management

New guidance in Section 3.1.2 of the *2019 Refinement*.

The uncertainty assessment is one of the instruments that will be used by the inventory compiler in the effort of improving the inventory over time. Regardless of the framework, under which national GHG inventories are developed and reported this will not be a one-time task. Inventories will be reported annually, biannually or over longer periods but will be updated and extended periodically.

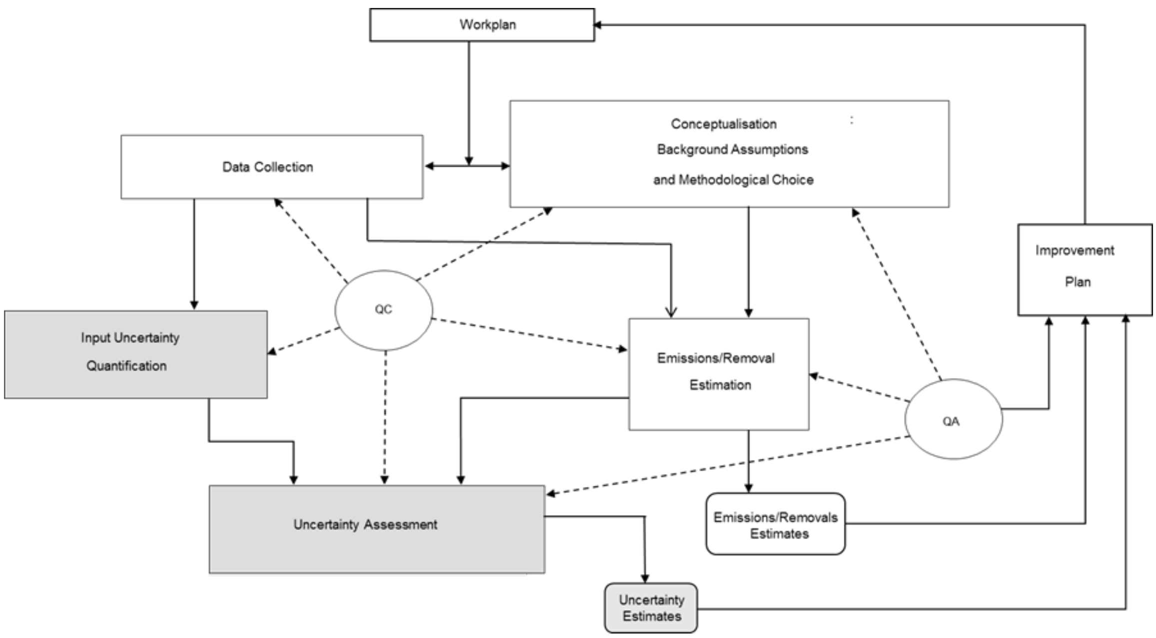
Between two reporting occasions, it is *good practice* to evaluate the institutional arrangements and their functions. All decisions that have been taken and options that have been adopted should be revisited. Ideally the inventory would have been verified by a third party and recommendations produced (e.g. reviews under the UNFCCC). Figure 1.1 in Chapter 1.1 of Volume 1 of the *2006 IPCC Guidelines* illustrates the steps of a typical inventory cycle and Chapter 1 of this report covers the steps to put in place the institutional arrangements necessary to manage the process, providing the organization and resources for planning and preparation of the inventory. Figure 3.1 below, builds from Figure 3.1 in Chapter 3 of Volume 1 of the *2006 IPCC Guidelines* to show how the uncertainty assessment fits in this improvement cycle.

The process of producing an uncertainty analysis can pragmatically be divided into four parts: (1) the rigorous investigation of the likely causes of data uncertainty and quality; (2) the creation of quantitative uncertainty estimates and parameter correlations; (3) the mathematical combination of those estimates when used as inputs to a statistical model (e.g., first-order error propagation or Monte Carlo method); and (4) the selection of inventory improvement actions (improvement plan) to take in response to the results of the previous three parts.

The improvement plan will assess the opportunities and prioritize the ways to improve the inventory based on the *key category* analysis, the uncertainty assessment, the recommendations from quality assurance and verification processes (including review process) and available resources.

Particularly in relation to the uncertainty analysis, the improvement plan will investigate ways to improve accuracy that would have been identified and ways to enhance precision for categories with high contribution to the overall uncertainty of the inventory. The approach 2 for *key category* analysis is a useful tool for this prioritization.

Figure 3.1 Overall structure of a generic uncertainty analysis



3.1.3 Overall structure of uncertainty analysis

Elaboration of Section 3.1.2 of the *2006 IPCC Guidelines*.

As part of the planning process, an improvement plan will be developed selecting the categories for which changes would be implemented in the new inventory. The changes would cover both methodological choice and data definition and collection. Most frequently, the improvement will focus of getting better data for the same methodology (e.g. collecting country specific data). The goal will include increasing the accuracy of the inventory with a better representation of the emissions/removals processes.

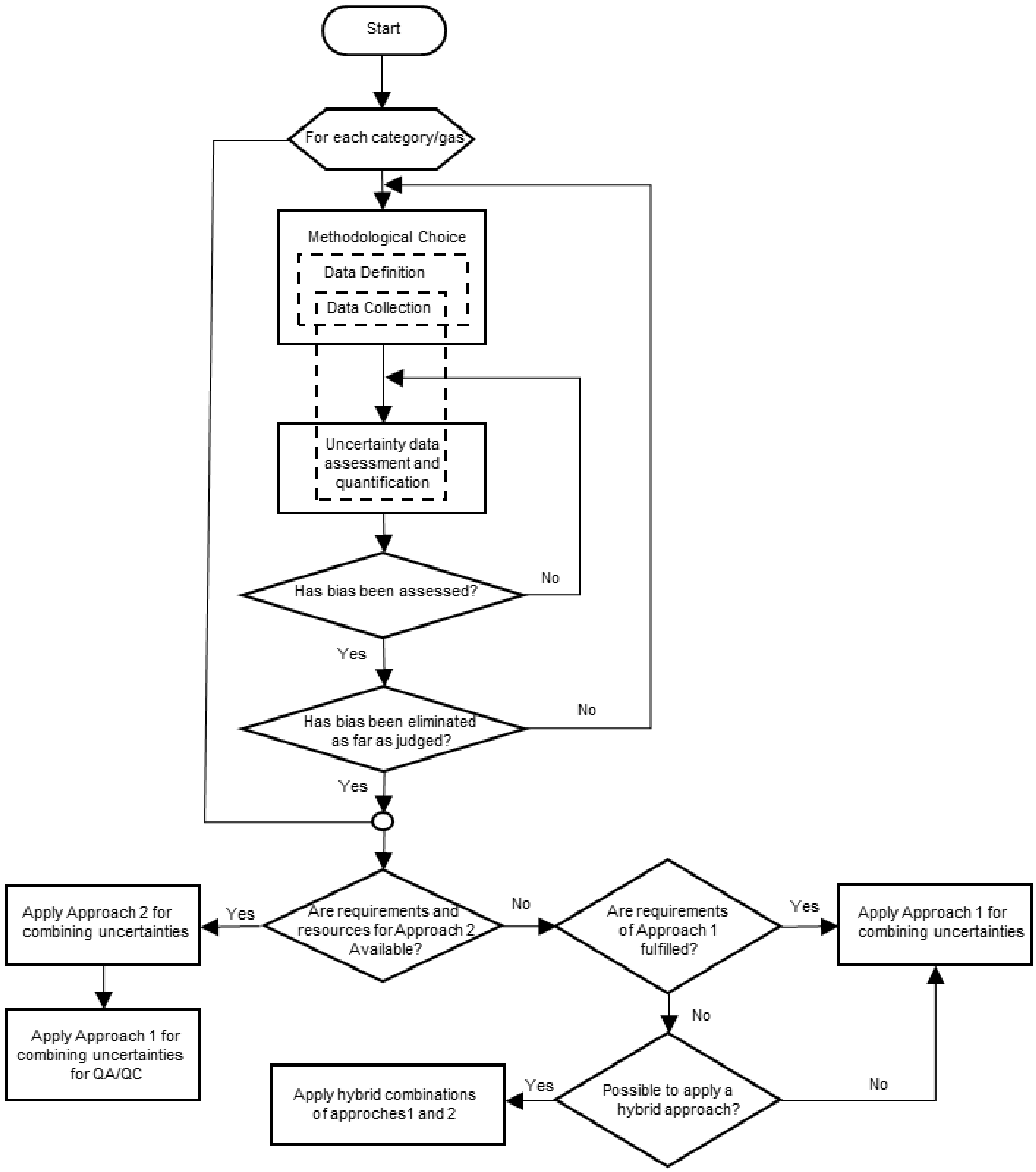
Figure 4.1 of Volume 1 of the *2006 IPCC Guidelines*, shows the steps of methodology choice that will depend on the selection of the category for improvement, the data availability, the possibility of data obtaining and the resources involved.

Figure 3.2 below show the general steps of an uncertainty assessment. It is important to note the strong link among these steps that usually need to be taken in conjunction and frequently reevaluated. This is true between the data definition and collection and between the data collection and the associated data uncertainty.

When assessing data uncertainty, it is essential to identify the causes of uncertainty involving the data estimation. In particular, priority should be given to identifying and correcting causes of bias.

Following the assessment of the uncertainty of the pieces of data used in emissions/removals estimation, the next step is to combine these findings, producing uncertainty assessment for a source of sink category that is subsequently propagated with uncertainties in all categories to determine the uncertainty in the whole inventory. Figure 3.2 show a simple scheme for the choice of approach but it is important to note that choices may vary among categories and usually a hybrid approach would be recommended. It is also important to note that even when requirements for application of approach 1 are not fully present it still can provide useful information about the uncertainty of the inventory. Because of its simplicity when compared with approach 2, it is also recommended to apply approach 1 as a QC/QA tool even when it is possible to apply approach 2.

Figure 3.2 Uncertainty analysis decision tree



3.1.4 Key concepts and terminology

No refinement.

3.1.5 Basis for uncertainty analysis

No refinement.

3.1.6 Causes of uncertainty

Elaboration of Section 3.1.5 of the *2006 IPCC Guidelines*.

Section 3.1.5 of the *2006 IPCC Guidelines* provide a description of the causes of uncertainty. It covers eight possible causes: lack of completeness, model, lack of data, lack of representativeness of data, statistical random sample error, measurement error, misreporting or misclassification and missing data. Depending on the cause, the result can be bias, random errors or both. Lack of completeness, lack of representativeness of data, misreporting or misclassification will typically lead to bias while model uncertainty and lack of data can lead to both.

For each category, the identification of causes of uncertainty is fundamental for elimination of bias and quantification of random errors. A poor identification step will entirely compromise an uncertainty reducing effort.

An investigation-focused approach to uncertainty is also one that is tightly integrated with QA/QC processes. In many ways, an investigation-focused approach to uncertainty is an in-depth approach to quality management. It rigorously identifies the causes of data quality problems, especially ones that general QC processes are unlikely to identify. These problems will often involve issues of incomplete data or other systematic biases in the data, which also happen to be key issues for developing a quantitative uncertainty analysis. (Gillenwater *et al.*, 2007)[1]

Both QA/QC and uncertainty analysis are part of a learning process. While the uncertainty analysis provides a standalone quantitative assessment of the inventory, its primary function is to understand what produces uncertainty and how to improve inventory quality. Conversely, the outcome of QA/QC procedures may result in a reassessment of individual category or parameter uncertainty estimates. Procedures to check quality and analyse uncertainties overlap and should work together because both processes are intended to understand the causes of uncertainty and identify potential areas of improvement (US-EPA, 2002)[2].

3.1.7 Reducing uncertainty

Elaboration of Section 3.1.6 of the *2006 IPCC Guidelines*.

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,

[1] http://www.springerlink.com/content/w238417357k02887/.

[2] http://nepis.epa.gov/Adobe/PDF/P1005GXH.PDF.

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 data that are measured*: 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).

For example, Tier 1 emission factors that are considered global defaults may be biased when they are applied in a specific country where emission rates deviate by a specific amount from the global defaults. Moving to a higher tier method in this case, will remove the bias associated with the default emission factor. Applying a higher tier method may also improve the precision of estimates as shown in Box 3.1.

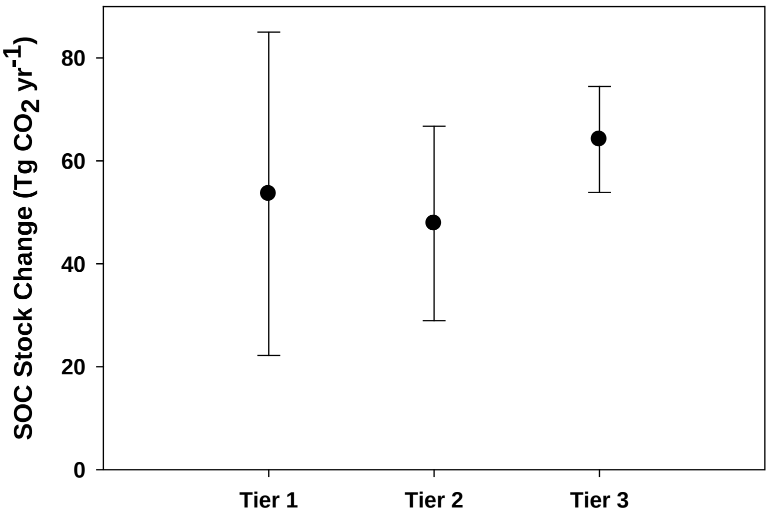
BOX 3.1

EXAMPLE OF REDUCING UNCERTAINTY IN AS SOURCE CATEGORY BY ADOPTING HIGHER TIER METHODS

Mineral soil C stock changes for *Cropland Remaining Cropland* have been estimated with all three methodological tiers for the United States, and this box provides information about how uncertainty has been reduced by moving to higher tiers. As with other source categories, the Tier 1 method is relatively simple with default emission factors provided in the IPCC guidance, but does require compilation of activity data for a simple classification of the lands, climate and soils. By moving to Tier 2, the compilers derived country-specific emission factors (i.e., stock change factors) that were based on experimental data from the region (Ogle *et al.*, 2003). Specifically, the new factors were derived using a linear mixed-effect modelling approach from 46 experiments evaluating the effect of tillage management on soil C, 19 experiments evaluating the impact of variation in carbon input to soils, and 35 experiments evaluating the impact of land use change between native conditions and long-term cultivation. Compilers also had the option of refining the classification of the activity data into a country-specific set of climate and soils types, in addition to management classes, but chose not to in this inventory. Regardless, flexibility in deriving new emission factors improved the precision of the estimates, reducing the confidence interval for the estimated soil C stock changes from ±59% using the Tier 1 method to a ±40% for the Tier 2 method (Figure, US-EPA 2017).

The compilers further improved the inventory for *Cropland Remaining Cropland* by developing a Tier 3 method. This method is based on applying the Century Ecosystem Model, and later the DayCent Ecosystem Model (Ogle *et al.* 2010, US-EPA 2017). These models incorporate a more mechanistic representation of the processes influencing soil organic matter dynamics, including water flows through the soil, crop production, organic matter decomposition, and nutrient cycling (Parton *et al.* 1987). With a more advanced representation of processes, the inventory was able to capture a broader suite of drivers influencing the change in soil C stocks. In addition, the inventory incorporated more detailed information on activity data and environmental variables, such as weather, soils, and management practices. In general, Tier 3 methods allow compilers to develop a methodology that is more specific to national circumstances, and ultimately an approach meeting *good practice* that is working towards the goal of neither over nor under-estimating emissions (or removals) as far as can be judged. The compilers also evaluated biases in the Century/DayCent model predictions of soil C stock changes by comparing results to independent data, and developed an empirical model to adjust for biases in the inventory results (Ogle *et al.* 2007). The Tier 3 inventory reduced uncertainty in soil C stock change estimates over 5 years from a ±40% with the Tier 2 method to ±16% for the Tier 3 method (Figure).

Figure: Estimates and 95% confidence intervals for mineral soil C stock changes in *Cropland Remaining Cropland* in the United States using Tier 1, 2 and 3 methods.



3.1.8 Implications of methodological choice

No refinement.

3.2 QUANTIFYING UNCERTAINTIES

Elaboration of section 3.2 of the *2006 IPCC Guidelines*.

Regardless of the methodology used to estimate emissions/removals for a category, the evaluation will be based on the underlying data. The overall uncertainty of the emissions/removals will depend on the uncertainty associated with each and every piece of data that is used to inform the inventory. As such, *good practice* uncertainty assessment begins with *good practice* in data collection. Uncertainty consideration will need to be an integral part of the data collection effort, including selection of data sources and choice of methods following the guidance in Chapter 2 of Volume 1 of the *2006 IPCC Guidelines*.

Chapter 3.2 of Volume 1 of the *2006 IPCC Guidelines* covers the different techniques for quantifying uncertainties depending on the availability of information and ways of data collection. These include measured data, published information, model outputs and expert judgement. Usually, the pragmatic approach will be a combination of the techniques.

Again, regardless of the approach, it is *good practice* to follow strictly the procedures for QA/QC according to the guidance in Chapter 6 of Volume 1 of the *2006 IPCC Guidelines*. This will be fundamental in preventing mistakes and misreporting and misclassification errors and approach deviations.

Ultimately, the measure of uncertainty will be a 95 percent confidence interval around a point estimate for the value. In order to develop this information a probability density function (pdf) will be associated with each quantity. The development of that pdf is an essential part of the uncertainty assessment. Section 3.2.2.4 of Volume 1 of the *2006 IPCC Guidelines* provide guidance on how to select the pdf. The representativeness of the pdf will depend on the characteristics of the quantity, including domain (e.g., if it can have both positive or negative values, or only non-negative values), range (e.g., is the range narrow or does it cover orders-of-magnitude) and shape (e.g., symmetry). The same characteristics will be fundamental when the approaches for combining uncertainties are selected.

Where the probability distribution of the emission factor (pdf) is believed to be 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%). In both cases, the understanding is that the confidence interval has a 95 percent probability of enclosing the true but unknown value of the emission factor.

BOX 3.2

STANDARD DEVIATION X STANDARD ERROR

In case of data assumed normally distributed, the 95% confidence interval may be derived considering the standard deviation (σ) or the standard error (SE) around our central estimate. The uncertainty of our estimate (μ) may be expressed as:

Uncertainty = ± (2σ/μ) *100%, where:

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2}$$

or

Uncertainty = ± (2SE/μ) *100%, where:

$$SE = \frac{\sigma}{\sqrt{n}}$$

Some practical examples may help the inventory compiler choose between these two statistics.

The standard deviation is a measure of variability. When we calculate the standard deviation of a sample, we are using it as an estimate of the variability of the population/individuals from which the sample was drawn. For data with a normal distribution, about 95% of individuals will have values within 2 standard deviations of the mean, the other 5% being equally scattered above and below these limits.

When we calculate the sample mean we are usually interested not in the mean of this particular sample, but in the mean for individuals of the same type, of the population from which the sample comes. For a sectoral category in the inventory, if we are interested in a specific emission factor, we usually collect data in order to generalize from them and use the sample mean as an estimate of the average emission factor for the whole category. Now the sample mean will vary from sample to sample; the way this variation occurs is described by the "sampling distribution" of the mean. We can estimate how much sample means will vary from the standard deviation of this sampling distribution, telling that we are interested in the variability of the mean, which is actually defined by the standard error of the estimate of the mean.

The standard error falls as the sample size increases, by contrast the standard deviation will not tend to change as we increase the size of our sample.

The following examples are provided for emission factors but can be extended.

Availability of annual information to derive country specific emission factors of a specific category/gas/fuel. Data are yearly collected from the whole population or a representative sample/samples of the category we are estimating.

This situation may occur in case data collected from facilities.

In this case, the emission factor derived from the data is calculated as the average emission factor from repeated measurements and changes year by year. We are therefore interested in the variability of this factor.

Assuming a normal distribution of the data collected, the 95% confidence interval may be expressed with the standard error and the uncertainty of the estimated emission factor as:

Uncertainty = ± (2SE/μ) *100%

Availability of irregular information to derive country specific emission factors of a specific category/gas/fuel. Data are not regularly collected and the result of data collected for one single year for a specific category is used for a longer period of the time series.
This situation may occur in case data are sporadically collected from facilities, e.g. methane emissions and relevant activity data and parameters from landfills

In this case, the 95% confidence intervals can be measured by the standard deviation of the central value because, assuming the value representative of other years, we have to consider the variability of the population/individuals. The uncertainty will be:

Uncertainty = $\pm (2\sigma/\mu) *100\%$

Availability of annual information to derive country specific emission factors at an upper level than actually used.

This situation may occur if we have for instance an average emission factor and we apply this country specific value to a specific portion of area, e.g. carbon stock per hectare of deforested area.

We are in a similar situation as case 2), the variability of the population/individuals should be considered to derive the 95% confidence intervals and the uncertainty is to be estimated as:

Uncertainty = $\pm (2\sigma/\mu) *100\%$

3.2.1 Sources of data and information

No refinement.

3.2.2 Techniques for quantifying uncertainties

No refinement.

3.2.3 Methods to combine uncertainties

Update of Section 3.2.3 of the *2006 IPCC Guidelines*. It further elaborates on the two approaches to combine uncertainties: Approach 1, propagation of error, and Approach 2, Monte Carlo simulation. A tool for the implementation of Approach 1 is also included as an addendum.

Once the uncertainties in activity data, emission factor or other parameters for a category have been determined, they may be combined to provide uncertainty estimates for the category emissions. Once the uncertainties for 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.

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.

Figure 3.2 flowchart shows a basic step-by-step suggestion on how the choice of approach could be made. In practice, however, the options are not always straightforward.

Approach 1 is simpler to apply but requires assumptions that frequently are not entirely met, such as lack of significant correlations among the quantities used in the inventory, or uncertainties that are less than ±30% of the quantity value. Approach 2 requires more information on the probability distributions of the data involved in the calculations. As such, it also involves assumptions and more information on the underlying processes and its application depends on the capacity to acquire this information. In turn, approach 2 may provide a more representative confidence interval for the uncertainty in the category.

Approach 2 will be particularly appropriate to use when uncertainties are large, their distribution are non-Gaussian and algorithms are complex functions.

Biases should be addressed prior to applying either Approach 1 or 2, as these approaches focus on quantifying the random component of the uncertainty of the inventory results where known sources of bias have been removed.

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 3.1, Approach 1 Uncertainty Calculation. A tool set up on a commercial spreadsheet software is provided, as an addendum to this methodological report, to facilitate the implementation of Table 3.1. The table is completed at the category level using uncertainty ranges for activity data and emission factors consistent with the sectoral *good practice guidance*[3]. 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 uncertainty 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. When inputs are known to be fully (or mostly) correlated, modified equations should be used or a preliminary step should be performed to combine these inputs before the application of the basic rules.

Uncertainty of the inputs will represent a 95 percent confidence interval expressed as a percentage of the central estimate of the input (e.g. ± 20%). When the probability distribution function is known to be asymmetrical, upper and lower limits of the confidence interval need to be specified separately (e.g., -10%, +20%). In this case, approach 1 will provide only a rough approximation and in order to be used the interval needs to be replaced by a symmetrical interval (e.g. ± 20%).

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[4] yields two convenient rules for combining uncorrelated uncertainties under addition and multiplication:

[3] 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.

[4] 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).

Where uncertain quantities are to be combined by multiplication a simple equation (Equation 3.1) can then be derived for the uncertainty of the product, expressed in percentage terms. This rule is approximate for all random variables. Under typical circumstances, this rule is reasonably accurate as long as the percentage uncertainty is less than approximately 30%. This rule is not applicable to division.

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, a simple equation (Equation 3.2) can be derived for the uncertainty of the sum, expressed in percentage terms. This rule is exact for uncorrelated variables.

EQUATION 3.2
COMBINING UNCERTAINTIES – APPROACH 1 – ADDITION AND SUBTRACTION

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

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).

x_i and U_i = the 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 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.

Applying approach 1 (level) in practice

Simple methods for estimation of the emissions of a category are usually based on the multiplication of activity data (AD) by an emission factor (EF). In many cases, it will be a reasonable assumption that these values are uncorrelated. The uncertainty associated with the emissions can then be calculated by equation 3.3:

EQUATION 3.3
COMBINING UNCERTAINTIES – APPROACH 1 – AD X EF

$$\boldsymbol{U_{emissions} = \sqrt{U_{AD}^2 + U_{EF}^2}}$$

More generally, both AD and EF can be result of several different parameters and this frequently occurs for the EF (e.g. EF = a x b x c). The uncertainty of the EF will be calculated as:

EQUATION 3.4
COMBINING UNCERTAINTIES – APPROACH 1 – EF = A X B X C

$$U_{EF} = \sqrt{U_a^2 + U_b^2 + U_c^2}$$

BOX 3.3
EXAMPLE OF UNCERTAINTY CALCULATION: CH_4 EMISSIONS FROM MANURE MANAGEMENT

TBD

The uncertainties associated with the emissions for each subcategory will be combined to obtain the uncertainty associated with a whole category and further combined to obtain the uncertainty of the whole inventory. In these steps the uncertainties as the quantities are combined through addition, equation 3.2 should be applied.

Particular attention should be given to the correlation in this step. The subcategories can be highly correlated, because either the ADs are derived from the same source or the EFs have parameters in common. A special situation occurs when an input is entirely dependent on a set of other inputs. As noted in the *2006 IPCC Guidelines* this could occur, for example, if residential fuel is estimated as the difference between total consumption and usage in the transportation, industrial, and commercial sectors. Similarly, in the AFOLU sector, when land transitions are assessed, total area transitions depend on the total area of the country, resulting in less degrees of freedom for the variables.

Approach 1 has limitations to the consideration of correlation as it only allows for full correlation or independency between the variables. Still broad sensibility can be implemented, either for correlation between variables in the same year or different years. This flexibility is included in the tool described in section 3.6.2. It is important to note that in the case of full correlation among categories, aggregation of these categories is the recommended procedure.

Uncertainty in the Trend[5]

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

[5] Note: More detailed information based on Section 3.7.2 of the *2006 IPCC Guidelines* will be included in the SOD.

The columns of Table 3.1, Approach 1 Uncertainty Calculation, are labelled A to Q 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 shows the sector of the IPCC category.
- B shows the code of the IPCC category.
- C shows the name of the IPCC category.
- D shows the greenhouse gas.
- E and F are the inventory estimates in the base year and the current year[6] respectively, for the category and gas specified in Columns C and D, expressed in CO_2 equivalents.
- G and I 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 G and I 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.
- H indicates if the uncertainty in activity data is correlated across years
- J indicates if the uncertainty in emission factor is correlated across years
- K is the combined uncertainty by category derived from the data in Columns G and I using the error propagation equation (Equation 3.2). The entry in Column K is therefore the square root of the sum of the squares of the entries in Columns G and I.
- L shows the uncertainty in Column K as a percentage of total national emissions in the current year. The entry in each row of Column L is the square of the entry in Column K multiplied by the square of the entry in Column F, divided by the square of total at the foot of Column F. The value at the foot of Column L 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 L and taking the square root.
- M 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.
- N 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.
- O shows the uncertainty introduced into the trend in emissions by emission factor uncertainty. If the uncertainty in emission factors is correlated between years (J = Y) the result is the product of the information in Columns M and I. If the emission factor uncertainties are not correlated between years (J = N) then the entry in Column N should be used in place of that in Column M and the result multiplied by $\sqrt{2}$.
- P shows the uncertainty introduced into the trend in emissions by activity data uncertainty. If the uncertainty in activity data is not correlated between years (H = N) the result is the product of the information in Columns N and G multiplied by $\sqrt{2}$. If the activity data uncertainties are correlated between years (H = Y) then the entry in Column M should be used in place of that in Column N and the $\sqrt{2}$ factor does not then apply.
- Q 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 O and P using Equation 3.2. The entry in Column Q is therefore the sum of the squares of the entries in Columns O and P. 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 Q 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

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

year emissions are 10 percent greater than the base year emissions, and if the trend uncertainty at the foot of Column Q 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 3.1
APPROACH 1 UNCERTAINTY CALCULATION

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
Inventory sector	IPCC category code	IPCC category name	Gas	Base year emissions or removals	Year *t* emissions or removals	Activity data uncertainty	AD uncertainty correlated across years?	Emission factor / estimation parameter uncertainty	EF uncertainty correlated across years?	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 Default: N	Input data Note A	Input data Default: Y	$\sqrt{G^2+I^2}$	$\frac{(K \bullet F)^2}{(\sum F)^2}$	Note B	$\left\lvert\frac{F}{\sum E}\right\rvert$	If J = Y $M \bullet I$ If J = N $N \bullet I \bullet \sqrt{2}$	If H = N $N \bullet G \bullet \sqrt{2}$ If H = Y $M \bullet G$	O^2+P^2
				Gg CO_2 equivalent	Gg CO_2 equivalent	%	Y/N	%	Y/N	%		%	%	%	%	%
e.g. Energy	e.g. 1.A.1	e.g. Energy Industries Fuel 1	CO_2													
e.g.	e.g. 1.A.1	e.g. Energy Industries Fuel 2	CO_2													
Etc...	Etc.	Etc...	…													
Total				$\sum E$	$\sum F$						$\sum L$					$\sum Q$
										Percentage uncertainty in total inventory:	$\sqrt{\sum L}$				Trend uncertainty:	$\sqrt{\sum Q}$

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 I, and enter 0 in Column G; it is suggested to assume correlation across years if some of the parameters used in the estimates are the same in both years or derived from the same source.
- If uncertainty is not correlated across years, enter the uncertainty into Column G, and enter 0 in Column I; it is suggested to assume no correlation between years if the estimates for the two years are independent from each other, for example based on independent measurements.

Note B: Absolute value of: $$\frac{0.01 \bullet F_x + \sum F_i - \left(0.01 \bullet E_x + \sum E_i\right)}{\left(0.01 \bullet E_x + \sum E_i\right)} \bullet 100 - \frac{\sum F_i - \sum E_i}{\sum E_i} \bullet 100$$

Where:

E_x, F_x = entry from row x of the table from the corresponding column, representing a specific category;

$\sum E_i$, $\sum F_i$= sum over all categories (rows) of the inventory of the corresponding column.

DEALING WITH LARGE AND ASSYMMETRIC UNCERTAINTIES

No refinement.

3.2.3.2 APPROACH 2: MONTE CARLO SIMULATION

TBD.

3.2.3.3 HYBRID COMBINATIONS OF APPROACHES 1 AND 2

No refinement.

3.2.3.4 COMPARISON BETWEEN APPROACHES

No refinement.

3.2.3.5 GUIDANCE ON CHOICE OF APPROACH

No refinement.

3.3 UNCERTAINTY AND TEMPORAL AUTOCORRELATION

No refinement.

3.4 USE OF OTHER APPROPRIATE TECHNIQUES

No refinement.

3.5 REPORTING AND DOCUMENTATION

No refinement.

3.6 EXAMPLES

TBD.

3.7 TECHNICAL BACKGROUND INFORMATION

No refinement.

3.7.1 Approach 1 variables and equations

No refinement.

3.7.2 Approach 1 – details of the equations for trend uncertainty

No refinement.

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

No refinement.

3.7.4 Methodology for calculation of the contribution to uncertainty

No refinement.

References

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

EPA, U. (2002) Procedures manual for Quality Assurance/Quality Control and Uncertainty Analysis.

EPA, U. (2017) Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2015.

Gillenwater, M., Sussman, F., Cohen, J (2007) Practical Policy Applications of Uncertainty Analysis for National Greenhouse Gas Inventories. *Water, Air &Soil Pollution*: *Focus* **7**: 451-474.

Mandel, J. (1984) The Statistical Analysis of Experimental Data, Dover Publications New York, USA, p.410.

Ogle, S. M., Breidt, F.J., Easter, M., Williams, S., and Paustian, K. (2007) An Empirically based approach for estimating uncertainty associated with modeling carbon sequestration in soils. *Ecological Modelling* **2005**: 453-463.

Ogle, S. M., Breidt, F.J., Easter, M., Williams, S., Killian, K., and Paustian, K. (2010) Scale and Uncertainty in modeled soil organic carbon stock changes for US croplands using a process-based model. *Global Change Biology* **16**: 810-820.

Ogle, S. M., Eve, M.D, Breidt, F.J., Paustian, K. (2003) Uncertainty in estimating land use and management impacts on soil organic carbon storage for US agroecosystems between 1982 and 1987. *Global Change Biology* **9**: 1521-1542.

Parton, W.J., D.S. Schimel, C.V. Cole and D.S. Ojima (1987). Analysis of factors controlling soil organic matter levels in Great Plains grasslands. *Soil Sci. Soc. Am. J.* **51**, 1173-1179.