Use of Facility–Specific Data in National Greenhouse Gas Inventories

Introduction

The 2006 IPCC Guidelines for National Greenhouse Gas Inventories give users methods for estimating emissions from most sources of emissions and removals by sinks. While they allow the use of detailed plant-specific data where this is available, they provide few details on the practical considerations to guide inventory compilers. The Guidelines provide general principles on including plant-specific data but not detailed guidance. With the development of plant and facility reporting systems in a number of different countries and regions, interest is growing in the use of this data. This Technical Bulletin is intended to explain the application of the IPCC Guidelines for National Greenhouse Gas Inventories in the light of users' experiences, and to provide inventory compilers practical approaches they may wish to adopt. This Technical Bulletin does not alter or add to the Guidelines themselves.

This Technical Bulletin is an output from an expert meeting convened by the IPCC's Task Force on National Greenhouse Gas Inventories (TFI) on this issue in Wellington, New Zealand, 18th-20th July 2011. This document has been reviewed by the participants and the Bureau of the TFI; however it has not been subject to the formal IPCC review and approval process and is therefore IPCC *Supporting Material* prepared for consideration by the Intergovernmental Panel on Climate Change.

This document should help to inventory compilers using either the Revised 1996 Guidelines for National Greenhouse Gas Inventories (IPCC 1997) or the 2006 IPCC Guidelines for National Greenhouse Gas Inventories (IPCC 2006). The Revised 1996 Guidelines were supplemented by two sets of Good Practice Guidance (IPCC 2000; IPCC 2003), and the Good Practice Guidance was incorporated into the main body of the 2006 Guidelines. In either case, Good Practice Guidance (GPG) is intended to provide a way to ensure that inventories are as unbiased, and as precise, as possible, within available resources. This document follows GPG.

Users of plant or facility level data will need to consider this document in light of their specific situation. Such data will come in many different forms and so the discussion here must remain fairly general.

There are a number of key issues that have been identified and are discussed in the following chapters:

- Ensuring data is of sufficient quality

- Consistency with definitions

- Estimating uncertainties

- Consistent use of data

- Dealing with incomplete coverage and time series consistency

Facility level data is information related to emissions from an individual plant, installation or factory. It can come in a number of forms:

Measurement data. Emissions can be measured directly (e.g. stack measurements) or indirectly (e.g. fuel carbon contents leading to CO_2 estimates). Measurements may be continuous (e.g. CEMS) or periodic, leading to plant specific emission factors.

Plant-Specific activity data may also be available. This is particularly useful where plant can be stratified into different groups with known emission rates (e.g. presence of different mitigations options). Where the plant cannot be grouped in this way the data may still be useful to compare with national sectoral estimates.

Emission factors for specific plant require plant-specific activity data to give plant-specific emissions. If plant specific emission factors are not accompanied by plant-specific activity data, it may be difficult to use such data to improve the inventory.

Using facility-level data would mean replacing an emission estimate for a category (the "sectoral emission estimate") with the sum of all the facility level data from all the facilities in that category, or by a combination of the sum of facility–level data and an estimate for the plants that have not reported their emissions.

Use of facility level data usually results in Tier 3 (or mixed Tier2/3) estimates. Inventory compilers should avoid using facility level data if more accurate estimates can be made without using it – but in these cases facility level data could still be useful for QC purposes.
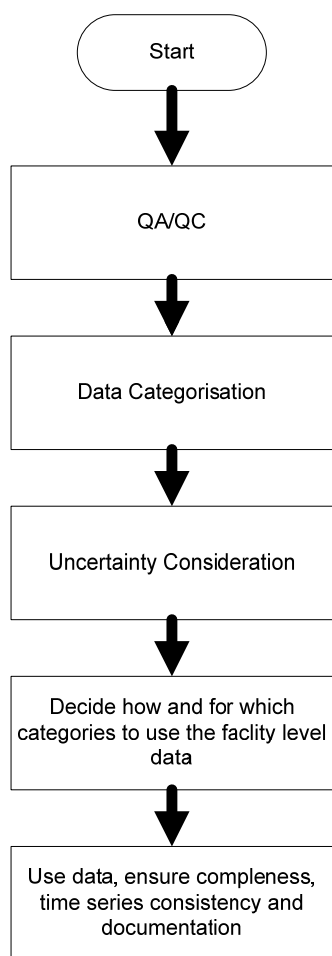
Start

↓

QA/QC

↓

Data Categorisation

↓

Uncertainty Consideration

↓

Decide how and for which categories to use the faclity level data

↓

Use data, ensure compleness, time series consistency and documentation

Figure 1 Overview of process of using facility level data in national greenhouse gas inventories.

Ensuring Data Quality - QA/QC

Clearly the data should pass quality checks to ensure it contributes to the overall inventory quality. However, there is no one-size-fits-all definition for data quality, and expert judgement from experienced inventory compilers will be needed, taking into account national circumstances, uncertainty of existing sectoral estimates, and anticipated improvement of the overall inventory. Inventory compilers need to lay down criteria for quality covering: methodology for deriving data; QA/QC procedures in line with IPCC Guidelines; sample size and uncertainty; transparency and representativeness.

In the 2006 IPCC Guidelines in Volume 1, Chapter 6 there is a considerable body of information on QA/QC that lays out what QA/QC is expected of data used in national GHG inventories. In many cases the facility level data will have to comply with QA/QC

procedures of whichever national statistical system initially collected the data. Where the inventory compilers believe that the national system QA/QC meets the requirements of the GPG then they do not need to duplicate all these QA/QC checks again. However, they should be able to demonstrate and document that these checks have taken place. In any case, examination of the data by inventory compilers will increase their understanding and ability to best use the data.
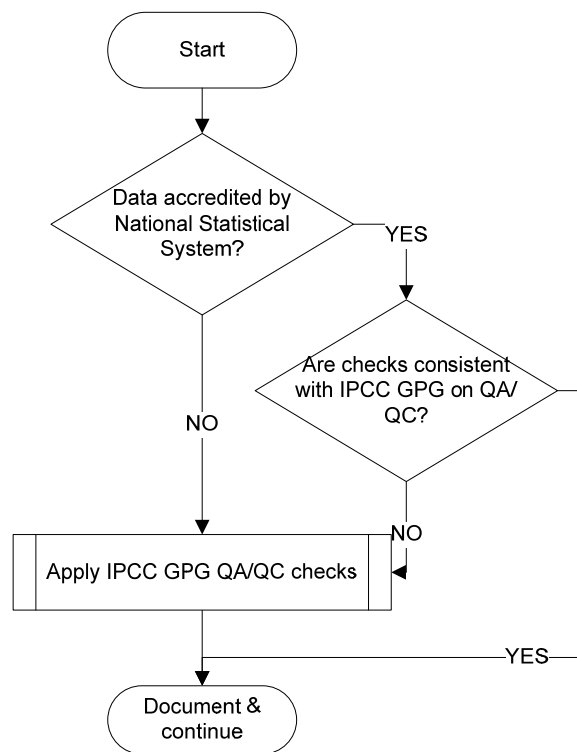
Start

↓

Data accredited by National Statistical System? — YES →

NO ↓

Are checks consistent with IPCC GPG on QA/QC?

NO ↓

Apply IPCC GPG QA/QC checks — NO

YES →

↓

Document & continue

Figure 2 QA/QC Checks

Specific QA/QC activities for facility level data include:

• Checks of the internal consistency between the detailed activity data, emissions factors and reported emissions. If the inventory compiler cannot get access to the detailed source data behind the estimates of facility level emissions a way of checking the data needs to be established. For example, the inventory agency can try to calculate implied emission factors[1] using an alternative source of activity data to see if the emission estimates are reasonable. Note that it may not be good practice to accept facility level data if the quality of the data cannot be understood. Therefore access to estimated emissions alone, while potentially useful, is of often of limited value.

• If the concept of materiality[2] and/or a risk-based approach has been used in the verification process for facility level

---

[1] An *implied emission factor* is the emission divided by some activity data (which may not be the same as data used to estimate the emission). For example emission measured emissions from coal combustion can be divided by the amount burnt and the implied emission factor compared with published default emission factors.

[2] Materiality is a concept (derived from accounting) relating to the importance or significance of a source or sink. This can be expressed in different ways. For example, in some reporting requirements, a source may

data, users should document this and explain what materiality level has been applied. Note that application of a materiality level would mean there still could be unidentified errors in the data. Users should also note that where plant less than a specific threshold have not reported facility-level data, these smaller plant may be systematically different to the larger plant and have different emission rates. Thus it may be difficult to apply data derived from the larger plant to small plant. There is no materiality criterion that exempts emissions from small plant from national greenhouse gas inventory reporting consistent with good practice.

- Understand and document the quality checks that have been applied to the facility level data.

Categorisation

Facility level data needs to be allocated to individual IPCC sub-categories. If the facility level data are not clearly in one IPCC category then inventory compilers should consider:

1) Do facility emissions cover more than one IPCC category (e.g. do plant emissions combine combustion and process emissions, or do the refinery emission estimates cover both fuel conversion and chemical production), if so additional information is needed to correctly allocate emissions.

2) Do estimates include off-site emissions? These are commonly called SCOPE 2[3] or SCOPE 3 (World Resources Institute and World Business Council for Sustainable Development 2004). National Greenhouse Gas inventories include emission from electricity generation, waste disposal, travel etc., under other inventory categories so there is a potential for double counting.

3) Is sufficient information available to categorise individual plant? For example, size of boiler may be known but not its use, making it difficult to allocate these emissions to inventory categories without double counting or omissions. The ISIC code (or similar) for the business may help in allocate the plant to an IPCC category, if it is available.

4) How do data relate to national statistics (energy balance, production)?

---

only be reported if it is larger than a materiality threshold which could be either an absolute size or a fraction of the total emission.

[3] Produced by World Resources Institute and the World Business Council for Sustainable Development, the GHG protocol is designed for corporate emissions and recognises that a company's activities result in emissions that occur offsite as well as those within a plant boundary (e.g. emissions form the generation of electricity). The GHG Protocol defines three types of emission estimates: Scope 1 (direct GHG emissions) – emissions from sources that the company owns or controls; Scope 2 (indirect GHG emission from purchased electricity, steam, or heat) – emissions associated with the generation of electricity, steam, or heat purchased and consumed by facilities or equipment that the company owns or controls; and Scope 3 (other indirect GHG emissions) – emissions from other sources the company does not own or control. This may include waste disposal, leased/outsourced activities, or emissions such as those related to business travel and employee commuting.

It may be possible to suggest changes to reporting to make matching to IPCC categories easier. Where facility level reporting is being considered, input from inventory compilers can make the resultant data more useful without significant additional costs.

If the inventory compiler considers the facility level data to be high quality then the most important issues are ensuring that there is no double counting, that all emissions are reported, and that where the emissions are reported is transparently documented. Small errors in allocation across source categories are less important because they will not affect the national total.

Uncertainties

The impact of the use of facility level data on the uncertainty of an emission estimate for an inventory category should be estimated.

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. Uncertainty comprises:

- random variation around a mean value (precision); and
- bias (accuracy).

The 2006 IPCC Guidelines provide general guidance on how to estimate uncertainty for sources estimated with facility level data. A reduction in uncertainty of the source's emission estimate is a desirable outcome of any inventory improvement.

The collection of facility data based on measurements increases the knowledge and information about the uncertainty associated with an emission source. Data on facility measurements provide more information in relation to the:

- precision and accuracy of emission estimates for individual facilities;
- distribution of emissions across facilities for a source; and
- variability of emission estimates over time.

It is likely that the use of data based on facility measurements will decrease the uncertainty of an emission estimate compared with the uncertainty associated with simpler Tier 1 or Tier 2 approaches.

In some cases, the improvement in the knowledge and information from the use of facility data may lead to an apparent increase in the estimate of uncertainty. This is possible if previously the uncertainty associated with a lower Tier emission estimate was underestimated.

The IPCC Guidelines explain how to estimate uncertainties. The first step is determining the uncertainties in all the input data. For measured data, this should be clear from the measurement method. The IPCC Guidelines provide guidance on how to do this[4]. For the overall sectoral emission estimate the process is fairly simple, the uncertainties of the input data (e.g. emission

---

[4] See Chapter 6 of the IPCC GPG 2000 or Chapter Volume 1 Chapter 3 of the 2006 Guidelines

factor and activity data) are combined according to the guidance provided. The same approach can be used for category estimates derived from facility-level data.

However, using facility level data does add to the complexity of the calculations and may depend on how the data is to be used. Where all the facilities in a category report their emissions, the overall uncertainty can be readily derived from the individual facility data. Where there is a residual of emission sources not covered by the reporting facilities then the uncertainty of this residual must be estimated.

If the collected facility-level data are considered representative of the whole source, the uncertainty value for the collected data should be applied to the residual. However, if the collected facility data are not representative of the residual or if the characteristics of the residual are unknown, the situation is more complex.

One approach is to consider stratification of the collected data into homogenous sub-groups, if appropriate. This can help understand the representativeness of that data in relation to the unknown residual. Homogenous sub-groups of a source may reflect regional influences such as climate or geological factors or technological drivers such as a common mitigation technique.

Another approach is to try to identify why the collected facility data may not be representative of the residual. This may help understanding of the uncertainty of the overall emission source. The representativeness of the collected facility data in relation to the unknown residual may be tested through the collection of more information. A sampling procedure can be designed to collect relevant information from the unknown residual of the source while balancing reporting burdens on facilities and administrative cost considerations. For example, where the facility data collection system is designed for facilities that exceed certain emission thresholds, or where reporting is undertaken on a voluntary basis, a targeted sample may be conducted to collect information from the missing facilities.

Deciding how to use Facility Level Data

Useful Facility Level Data

The most useful facility level data is likely to be where the emissions data has been derived from plant-specific information or measurements. For example, measured emissions or measured plant specific fuel quality (e.g. carbon content of fuel) should give improved emission estimates compared to sectoral estimates based on default or country specific emission factors for the reporting facilities.

Where the facility level data are produced using sector-wide emission factors and plant level activity data (e.g. fuel consumption or output) the data will probably not lead to more accurate estimates of emissions than sectoral approach (and may, in fact be equivalent). This does not mean that such data cannot be used:

- It may be the case that the facility-level data allows inventory compilers to stratify the category (e.g. into different levels of

abatement) and to apply appropriate emission factors to each strata.

- Where the emission factor used differs from the default emission factor used by the inventory agency, or from the IPCC default, then, in accordance with GPG, this difference should be investigated and the reasons documented. This may lead to improved emission factors.

- Such facility level activity data may be helpful in supplementing other data e.g. national energy balance data (in the case of fuel), and other national statistics. In the case of facility level fuel consumption, these data may help more accurately estimate fuel use in individual categories (such as 1A2f), and may be useful in helping understand the range and characteristics of fuels used (e.g. in refineries).

- There may be examples of facilities reporting only limited sets of data – e.g. emissions mitigated. With care, such data may be used to improve estimates made using other methods.

Choosing between sectoral and facility level data

Where inventory compilers have choices between use of facility-level data (e.g. ETS data) and other data (e.g. national energy statistics), they need to develop guidance on making consistent decisions. A key consideration is the quality of the activity data and how to judge this quality. Steps to this can include:

1. Assess the difference between two different data sets and document the assessment and its results.

2. Consider whether the uncertainty of the facility level data is lower or higher than the sectoral emission factor-based estimates made by the inventory agency. Try and determine the uncertainties associated with emissions, and if possible emission factors and activity data. Where a Tier-based approach to facility level estimates has been used, determine the uncertainty level associated with each Tier. This will necessitate access to the detailed data behind the calculations. Given a choice between facility-level data based heavily on default emission factors and any associated assumptions compared with sectoral estimates that made with country specific emission factors then the country specific data sets should normally be used.

3. Consider any QA/QC verification that has been applied and how this relates to the requirements of good practice. The

---

Deciding How to Use Facility-Level Data - Key Questions

- Are the data based on measurements?
- Is additional information available to allow stratification?
- Can the data be used to derive EFs or be used for QA/QC?
- Has clear guidance been developed on deciding how to use facility-level data?
- For combustion emissions: how does this data compare with national energy statistics?
- Is the sample size adequate compared with the number of facilities in the IPCC category?

fact that QA/QC and verification has been applied should not be always taken to indicate low uncertainty.

4. Make a comparison between the activity data of the facility-level data and comparable national and/or international statistics, if available.

Inventory compilers may wish to develop decision trees to guide their decision to use a specific body of data. This would be particularly useful where there are facility-level data covering a number of inventory categories. The use of a decision tree would enable the compiler to demonstrate that consistent criteria had been used to decide on how to use the data. Compilers can devise separate decision trees for emissions factor data and activity data but this may not be needed.

Specific Fuel Combustion Considerations

For combustion emissions, compilers should consider how these facility-level data improve emission estimates. In particular:

1) How do these data sets compare with national energy balance?

2) Will using these data improve the national total or reallocate emissions between sub-categories? (e.g. where you have only part of a combustion category covered by facility-level data)

If the energy balance and fuel characteristics are well known then it is unlikely that facility-level data will significantly improve the national total fuel combustion CO_2 emission estimate, but will improve the split between sub-categories. However emission estimates of non-CO_2 gases from combustion, fugitive emissions and other processes can be improved.

Sample Size

Often the facility-level data do not cover all the facilities in an IPCC category (e.g. see discussion of materiality above). Perhaps only facility greater that a size threshold have to report, or maybe only larger facility have to produce measured emissions while smaller facility can report estimates based on emission factors. In such cases the usefulness of the data may be improved by dividing available data into homogeneous sub-categories. Guidance on stratification is given in Volume 1 Chapter 3 of the 2006 Guidelines.

However, caution is required when applying EFs from a limited sample of facilities to a much larger group: compilers need to understand the actual processes and methods used in the facilities to be able to demonstrate they are similar in their emission characteristics. Users should:

- determine the sample size (data available for each category) compared to number of facilities in the category;
- consider the probability distribution for each category – a narrow range of measured emission rates implies that the emission rate can be extrapolated to the unmeasured facility with more accuracy;
- assess uncertainty and compare with the uncertainty of sectoral emission factor approaches

In addition, as a QA/QC measure, facility-level data should be compared with a proxy measure such as fuel use, production, vehicle kilometres or passenger numbers. Any differences need to be explained and documented.

Consistent use of data

In deciding if and how to use facility-level data a formal, documented, decision process leads to clear, well justified decisions and a consistent treatment of facility-level data (Figure 3). This is especially the case where facility-level data cover several inventory categories and choices must be made, category-by-category, on whether or not to use the data, and if used, how to use the data.

Time Series Issues

Usually facility-level data do not provide a complete time series. The 2006 Guidelines, in Volume 1, Chapter 5 "Time Series Consistency" provide guidance on how to splice data sets and to fill any gaps. It is possible to splice together facility-level data with data from other sources.

One issue that may arise is that facility-level data is normally not recalculated after the emission estimates have been accepted – this may be a problem if errors are detected by the inventory agency or other organisations. This is a particular problem where emission factors are used. Good practice requires that emission estimates are unbiased and so, if the inventory agency detects a bias in future years, it should correct the error, recalculate, and report the deviation between the new estimate of emission and the original estimate. Recalculations will also be needed if other errors are detected in previous years' estimates.

Where the facility-level data is only periodically available (e.g. measurements made every other year) then the inventory compiler could integrate the available data into the time series using the procedures set out in the 2006 Guidelines, or alternatively the facility-level data could be used as QA/QC check against the compiler's own estimates.

Documentation of the approach used to ensure time series consistency and of the assumptions and data used is needed to support transparent and verifiable estimates.

Problems with time series consistency are likely. For years prior to facility-level data becoming available, inventory compilers will need to calculate using IPCC category method and splicing the two time series using the methods outline in the good practice guidance on time series consistency..

1) Usually compilers should use the overlap method to recalculate earlier years. Alternatively, they may consider using new emission factors for earlier years based on facility-level data or from elsewhere.

2) Compilers may want to wait until there are several years' data before using facility-level data so that the two data sets can be compared.

3) Compilers need to consider how far back in time to recalculate. Current facility may not exist in base year, abatement, management and fuel quality may have changed and so earlier, sectorial data may be more reliable.

4) Trading scheme data may not be recalculated as new information and corrections of errors become available (e.g. EUETS). Compilers should use their judgement on how to deal with these depending on the individual circumstances.

5) Compilers need to consider uniformity of methods used to generate facility-level data over time and ensure that changes do not introduce time series inconsistency

.Dealing with incomplete coverage

One key consideration to apply when selecting facility-level data is completeness and consistency of category coverage. The following causes for incomplete or inconsistent data should be considered:

- *Reporting thresholds* – many data collection systems contain corporate, facility or source thresholds. Data below these thresholds are not reported by facilities.

- *Unknown facilities* – a complete list of facilities within a category may not be known.

- *Non-compliance* – facilities may choose not to report their data. Data collection systems with compliance provisions can help mitigate non-compliance.

- *Incorrect reporting* – facilities may make errors in their facility-level reporting. Data collection systems with clear guidance, data evaluation systems and compliance provisions can help mitigate incorrect reporting.

- *Geographic incompleteness* – regional factors may lead to incomplete coverage of facilities.

- *Selection of methods* – Some emissions reporting systems provide facilities with the option of estimating facility specific emissions data or using default factors. This self-selection could lead to a bias in the reported data.

These factors may lead to a bias in the reported data if it is used to represent the whole category. However, it may still provide representative data for one or more sub-categories.

Characterising the size of the overall sector is an iterative and continuing process where facility-level data can play an important role in identifying otherwise unknown emission sources. Other sources of data to help assess the overall size of the category could be national statistics, trade association data and national inventory stakeholder committees.

Consistent Use of Data - Key Questions

- Has clear guidance been developed on deciding how to use facility-level data?
- Does the data provide a complete time series? If not, how will time series consistency be ensured?
- Is the data consistent?
- Does the data cover all facilities in an IPCC Category?
- How are recalculations dealt with?

Where facility data is well characterised but represents a subset of a category a decision tree may be used to assess how to use this data.

Because representativeness will vary according to a Party's national circumstances (e.g. the size of the category, whether mitigation and technology changes are included in the facility-level data), ultimately expert judgement will need to be used in assessing whether and how facility-level data should be incorporated into the national inventory.
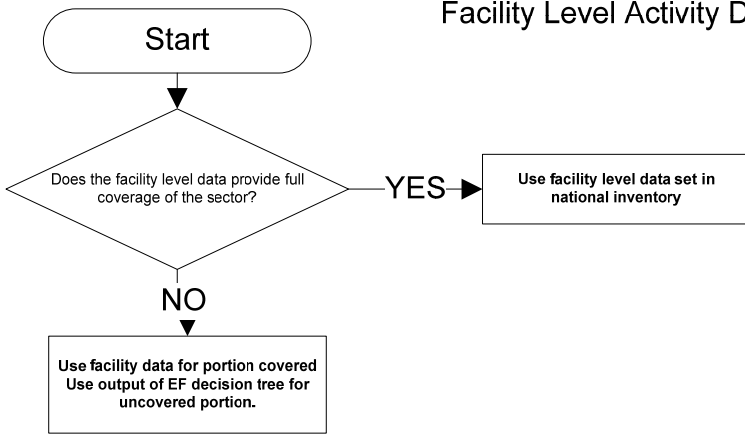
Use of data in EFDB

The IPCC's Emission Factor database (EFDB) is a library of data to be used in conjunction with IPCC Guidelines by inventory compilers to improve their estimates by using factors that better represent their national circumstances than the IPCC default data. While countries are encouraged to derive their own country-specific data, the ability to do this is limited by available resources and time. Compilers can look at the EFDB data to find material that is appropriate to their national circumstances.

Scientific verification is not done by EFDB Editorial board but they ensure that that the data is robust and that adequate documentation is available so that the users can judge whether EFs can be used for their inventory. It is the responsibility of the users themselves to ensure the data they use is appropriate to their specific circumstances.

Facility-level EFs may provide a valuable source of such data and should be submitted to the EFDB.

Facility Level Activity Data

```
         ┌──────────────┐
         │    Start     │
         └──────┬───────┘
                │
                ▼
         ╱────────────╲
        ╱  Does the    ╲         ┌─────────────────────┐
       ╱ facility level ╲  YES   │ Use facility level  │
       ╲ data provide   ╱───────▶│  data set in        │
        ╲ full coverage╱         │  national inventory │
         ╲of the sector╱         └─────────────────────┘
           ╲────────╱
                │
               NO
                │
                ▼
     ┌──────────────────────┐
     │ Use facility data for│
     │ portion covered      │
     │ Use output of EF     │
     │ decision tree for    │
     │ uncovered portion.   │
     └──────────────────────┘
```

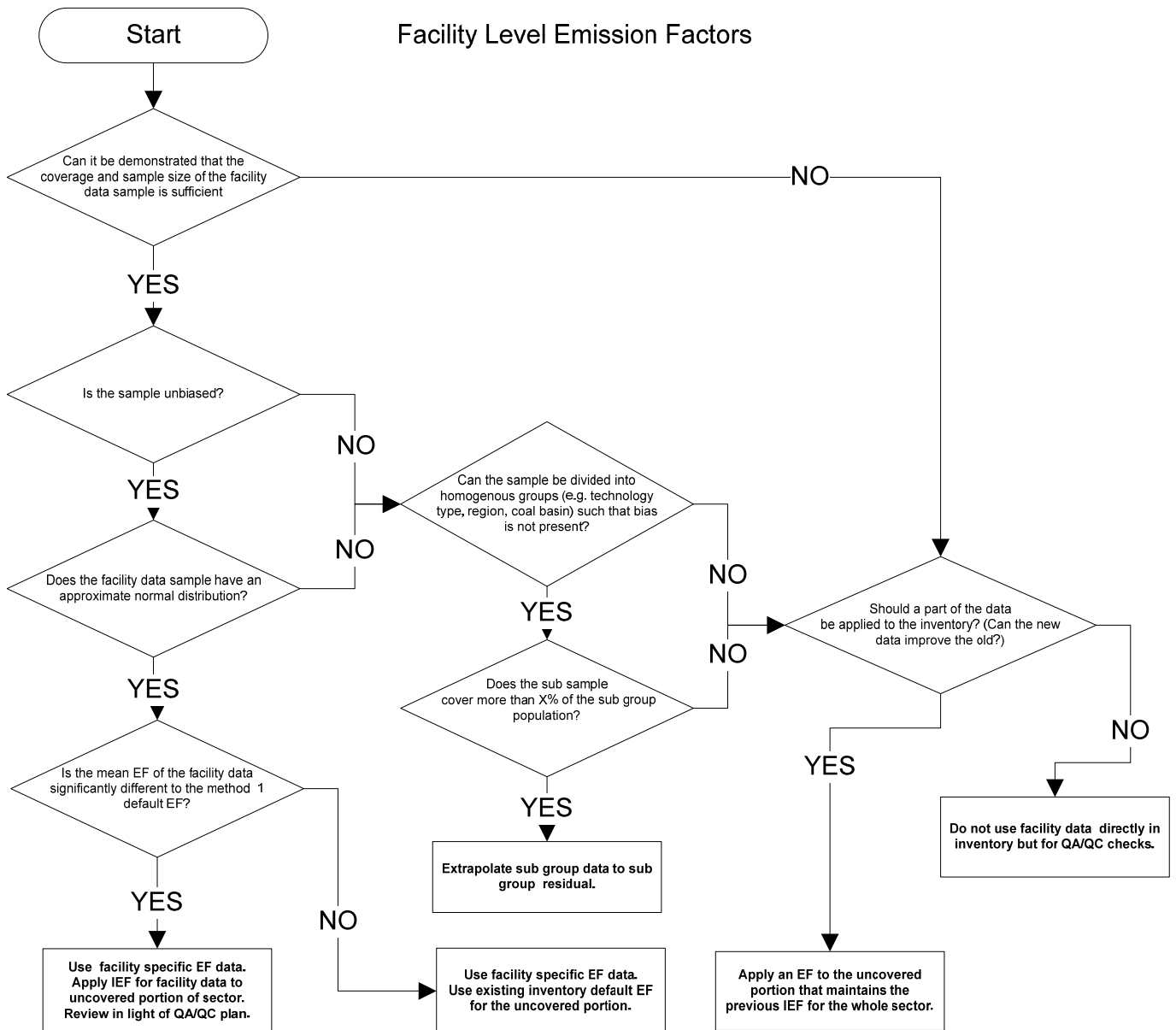## Facility Level Emission Factors

Figure 3 Examples of Decision Trees for Treatment of Facility-Level Data

7

Bibliography

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

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

IPCC (2003). Good Practice Guidance for Land Use, land-Use Change and Forestry. Penman J., Gytarsky M., Hiraishi T., Krug, T., Kruger D., Pipatti R., Buendia L., Miwa K., Ngara T., Tanabe K., Wagner F.,. Hayama Japan, IGES.

IPCC (2006). 2006 IPCC Guidelines for National Greenhouse Gas Inventories. Eggleston H.S., Buendia L., Miwa K., Ngara T., Tanabe K Hayama Japan, IGES.

World Resources Institute and World Business Council for Sustainable Development (2004). The Greenhouse Gas Protocol. A corporate Accounting and Reporting Standard. Revised Edition. USA, WRI & WBCSD.

Further Information

Further information, copies of the IPCC Guidelines and additional material are available for the Technical Support Unit of the TFI.

www.ipcc-nggip.iges.or.jp

or e-mail

nggip-tsu@iges.or.jp

Technical Bulletin, are intended to explain the application of the IPCC Guidelines for National Greenhouse Gas Inventories in the light of users' experience, and to provide inventory compilers practical approaches they may adopt if they so wish. This Technical Bulletin does not alter or add to the Guidelines themselves.

This Technical Bulletin is an output from an expert meeting convened by the IPCC's Task Force on National Greenhouse Gas Inventories (TFI) on this issue in Wellington, New Zealand, 18th-20th July 2011. This document has been reviewed by the participants and the Bureau of the TFI; however it has not been subject to the formal IPCC review and approval process and is therefore IPCC Supporting Material prepared for consideration by the Intergovernmental Panel on Climate Change.