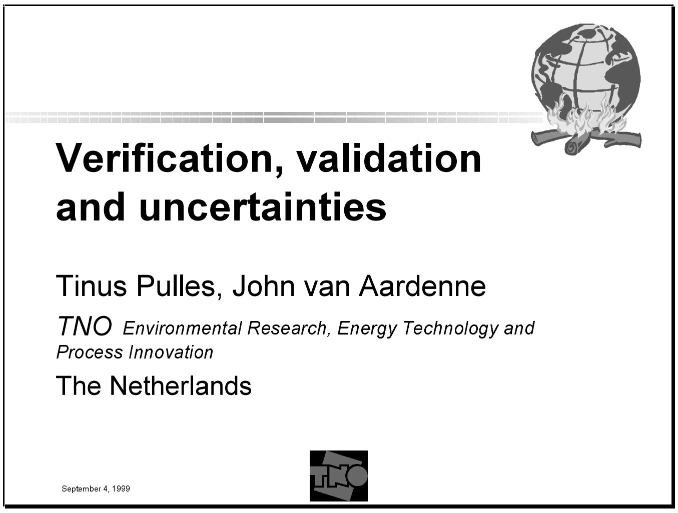
2. VALIDATION AND VERIFICATION

2.1. VERIFICATION, VALIDATION AND UNCERTAINTIES

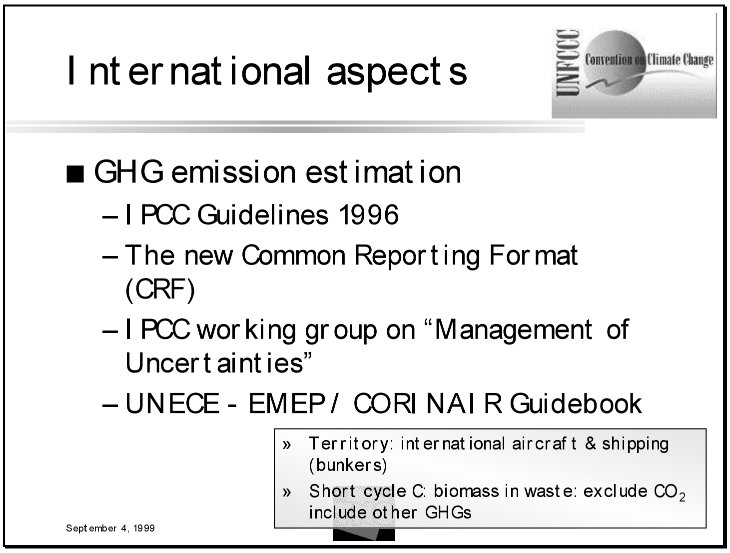
Tinus Pulles and John van Aardenne

TNO Environmental Research, Energy Technology and Process Innovation
P.O. Box 342, 7300 AH Apeldoorn, The Netherlands
E-mail: m.p.j.pulles@mep.tno.nl

Slide 1



Slide 2



Greenhouse gas (GHG) inventories are produced in response to obligations derived from the Framework Convention on Climate Change (UNFCCC).
Parties to UNFCCC have accepted the IPCC 1996 Revised Guidelines to guide estimating GHG emissions.
Guidelines:

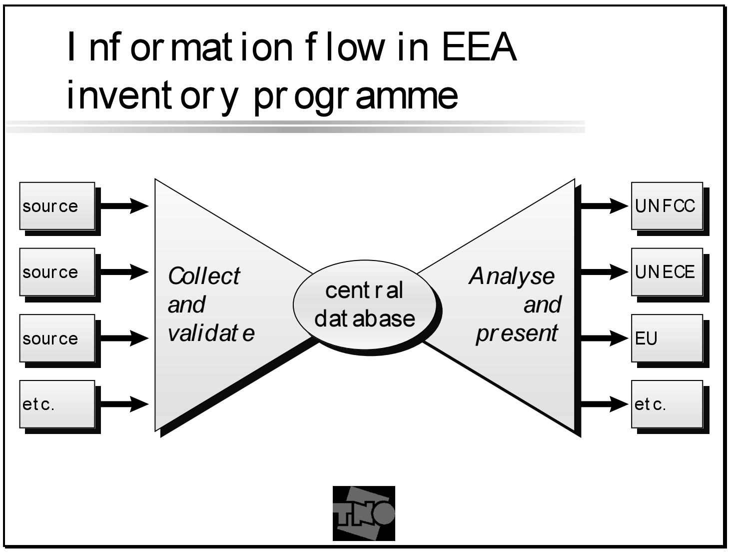
- **list sources and sources sectors**
- **propose default emission factors**
- **propose Summary Tables and Working Tables**

The Guidelines will not be changed soon. The reporting format will. Parties did so far not deliver the inventories in full detail according to the reporting format of the IPCC guidelines, Most parties reported using (derivatives of) the 8 Summary Tables only.
To increase transparency and comparability, SBSTA 10 has decided to adopt a new Common Reporting Format (CRF). Parties are encouraged to use this new CRF for the third national communication. Parties are requested to use the tables unchanged (do not delete or add rows or columns; do not change labels)

SBSTA asked IPCC Working Group on Management of Uncertainties to propose "Good Practice" and to produce guidance for estimating and reporting uncertainties.
Other reporting requirements (UNECE CRLTAP) use different definitions and a different sector split.
The major differences being:

- Territorial: international air traffic & shipping (bunkers); cruising aircraft
- Short cycle C in IPCC: biomass in waste: exclude CO_2 include other greenhouse gases

Slide 3

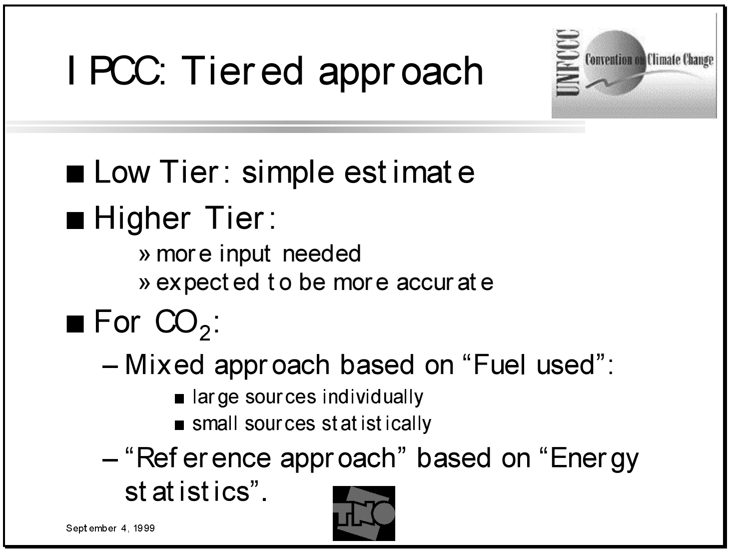


The inventory collection and reporting process is given in the scheme A clear distinction is made between collecting the data and reporting enabling:

- **consistency between different reports**
- **UNFCCC,**
- **UNECE/CLRTAP,**
- **European Union:**
  - ✓ (Amended) CO_2 Monitoring Mechanism,
  - ✓ Large Combustion Plant (LCP) Directive,
  - ✓ Integrated Pollution Prevention and Control (IPPC) Directive,
  - ✓ Emission Ceilings Directive etc.
- **OSPARCOM/HELCOM**
- **relatively simple adaptation to changing reporting formats of individual protocols and**

This approach is fully compatible with the concept of a Common Reporting Format (CRF), now being proposed by the UNFCCC Secretariat and accepted by SBSTA for a trial phase.
This means that where the CRF contains sectoral back ground data, this should not be interpreted as the activity data and emission factors as used by estimating and collecting emissions, but as information that increases transparency in the sense that it allows cross country comparisons.

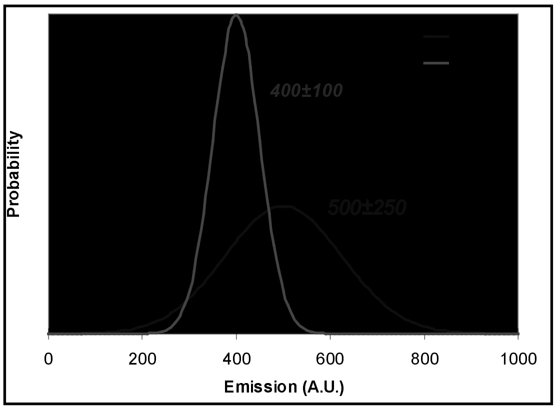
Slide 4



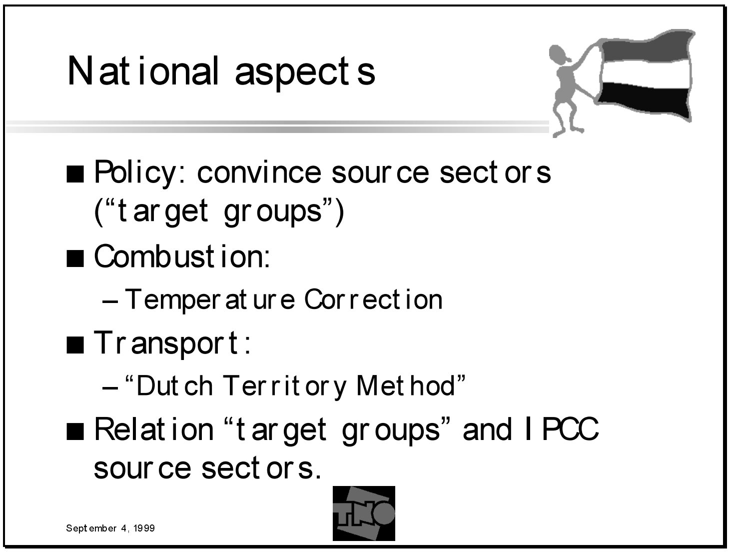
IPCC Guidelines provide different “Tier”s:

- **Lower Tier: “quick and dirty”**
- **Higher Tier: more detail, more input data required,**
  - ✓ expected to have higher scientific quality: smaller 95 % confidence interval

So different tiers produce different estimates with (expectedly) different uncertainties. So by applying higher tier methods, the number might (but need not) change, but the uncertainties should decrease.
For CO_2 most parties (including the Netherlands) apply a mixed approach, using detailed estimates for large fuel using activities and a statistical approach for the smaller sources.
The IPCC Summary tables request in addition reporting the “reference approach”, based on energy statistics (Energy Balance).



Slide 5



At the national level some additional aspects are relevant:

- **in preparing, performing and monitoring national policy, emission data are needed that**
  - ✓ enables prioritisation of abatement measures
  - ✓ will convince all actors involved of the effectiveness, efficiency and necessity of measures proposed
  - ✓ show continuity in reporting at the national level:
    - ∗ temperature correction
    - ∗ road transport emissions

These aspects are of low importance for the international community.

A second item, recognisable from this workshop's agenda is the relation between the "target groups" defined in national environmental policy and the "IPCC source sectors".
These issues can be more easily solved when indeed a clear distinction is made between data collection and reporting. While preparing a report the database should be left unchanged, but necessary extra analyses, like a temperature correction, should be applied while interpreting and aggregating the data in the inventory towards the different reports.

Slide 6

Perspectives on data quality and intended use of the data

	Perspective	*High quality if ...*
"Scientist"	Scientific debate: search for weaknesses and errors; falsification.	...it produces predictions that are confirmed
"Policy maker"	Political debate: search for consensus and agreement; compromise	...everybody involved agrees
"Lawyer"	Judicial debate: search for proof or doubt; persuasion	...it convinces a judge or jury

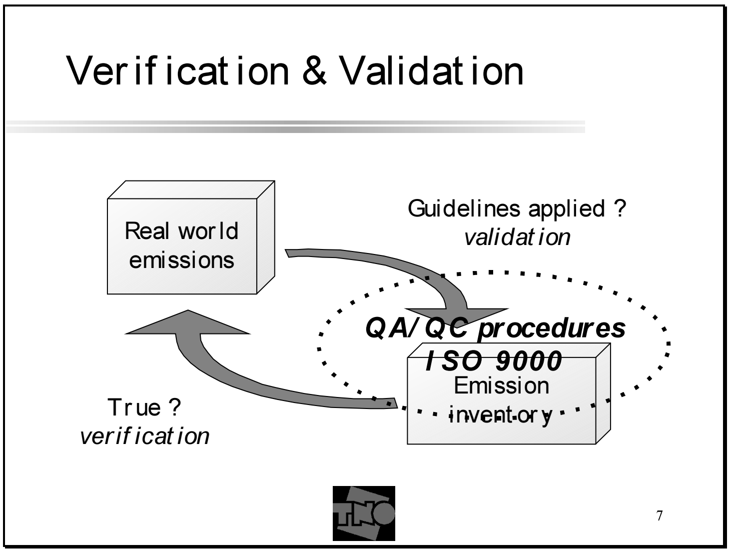
6

Three perspectives on data quality:

- The scientist is looking for the truth by trying to find weak spots in theory and data and by falsification. The data quality will be high if the data or predictions based on them are confirmed by independent estimates. If falsification occurs, the scientist will work on it until he or she understands the reasons and has derived better data or a better theory.
- A policy maker is looking for agreement and will therefore be more inclined towards reaching consensus and compromise. In many cases a policy maker does not have enough time to wait until all scientific problems are solved: a company might have asked for a permit for a new activity and regulations prescribe a decision to be made within a given period of time; or a country has to report its emissions according to a protocol before a certain fixed date. The policy maker will have to decide although a number of uncertainties are still present and a number of phenomena might not be fully understood.
- The lawyer has again a different perspective. He or she might be involved in compliance checking and will regard data to have high quality if the data are convincing.

These perspectives on data quality will also influence the perspective on "truth" and "quality" and hence on verification and validation of emission inventory data and of models.

Slide 7

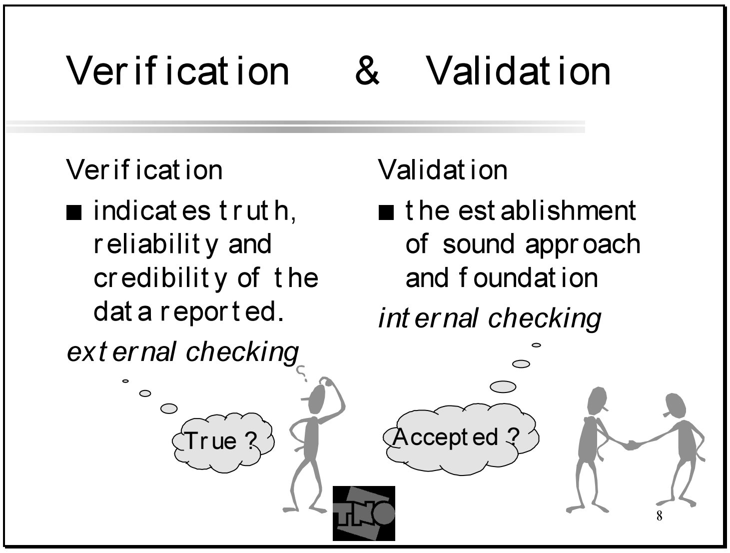


We recognise three levels of inventory quality:

- **procedural quality, to be established by applying "good practice" and "quality assurance / quality control" procedures, assuring adequate documentation and error free arithmetic;**
- **transparency and comparability, to be established in some kind of interaction between the responsible inventory compiler and the receiving body assuring that the inventory is compiled according to the guidelines and that it is comparable to other countries' inventories**
- **accuracy, completeness etc., to be established in primarily scientific assessments of the inventory involving independent estimates or measurements assuring the "truth" in the values stored in the inventory.**
- These three levels of inventory quality will be relevant with respect to national inventories in a cyclic way. Scientific knowledge is brought in while defining the guidelines, including the default emission factors. Once guidelines are accepted, good practice and QA/QC procedures can help in compiling the inventories in such a way that application in national and international environmental policy is possible. The international bodies, receiving national inventories, must validate them with respect to transparency and comparability. Once emission inventories are available, verification studies might be used to check the methods as prescribed in the guidelines and "science" can suggest amendments to the guidelines and the default emission factors therein.

In such a cyclic process, inventory quality can be improved. The development of the tools within the CORINAIR program of the European Environment Agency might very well contribute to this.

Slide 8



Definitions below are derived from the Glossary developed within the IPCC Working Group on Managing Uncertainties:

Verification

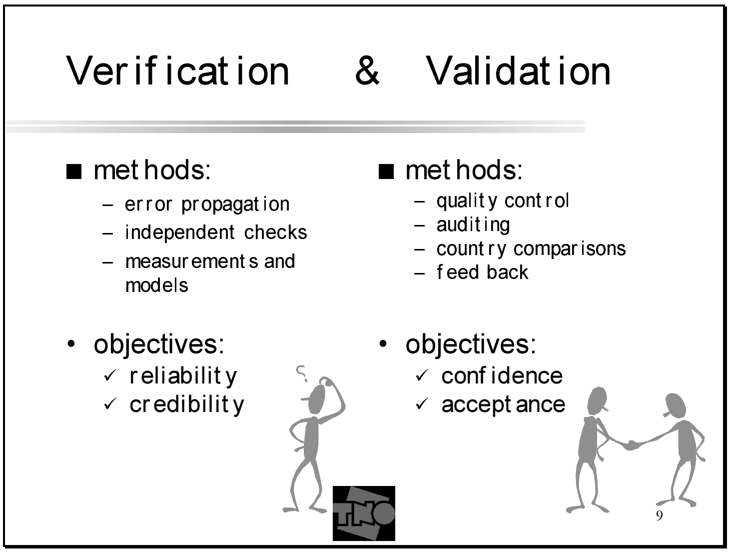
The term verification is used to indicate truth or to confirm accuracy and is used to represent the ultimate reliability and credibility of the data reported.
This will call for external checking, using independent estimates by other organisations, models, measurements, etc.

Validation

Validation is the establishment of sound approach and foundation. The legal use of validation is to give an official confirmation or approval of an act or product.
This internal checking of the inventory is meant to ensure that it has been compiled correctly in line with reporting instructions and guidelines and that the calculations are arithmetically correct.

In some earlier reports and papers, both words have been used differently, but at present the definitions as reproduced here are being accepted widely. We propose to also accept in the Netherlands the definitions as above.

Slide 9



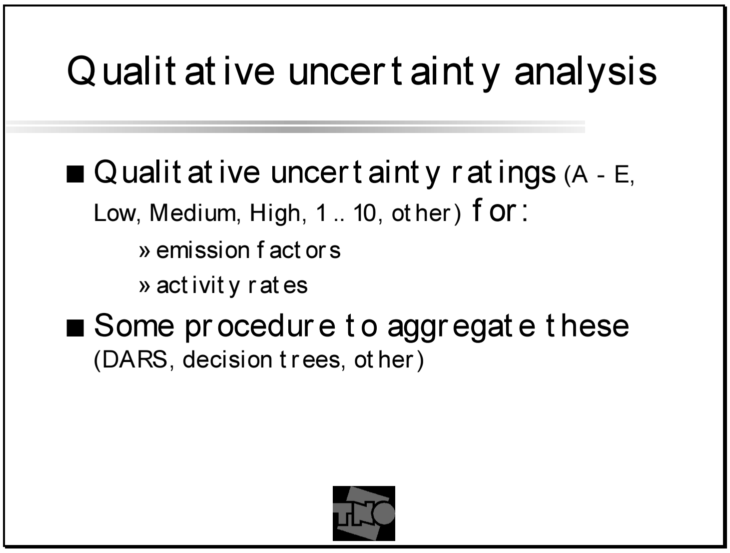
Verification can be performed by applying:

- **estimating errors in variables and parameters and establishing the resulting error distribution in the total number (error propagation)**
- **compare with other independent estimates and calculations**
- **compare with measured values, using models**

Validation uses methods like:

- **QA/QC systems (ISO 9000 certification or equivalent, assuring full documentation and allowing to redo the inventory and to reproduce all numbers**
- **auditing (part of ISO 9000 certification): allowing independent auditors to check the application of the correct procedures and the origin of input data**
- **between country comparisons ("All countries make the same errors").**
- **feed back from the Convention's secretariat**

Slide 10



Uncertainty analysis

- **Qualitative**
- **Quantitative**

Qualitative:

- **needs quantitative uncertainty ratings for**
  - ✓ activity data
  - ✓ emission factors
- **needs a procedure to aggregate the individual ratings of emission factors and activity rates towards an overall qualitative rating.**

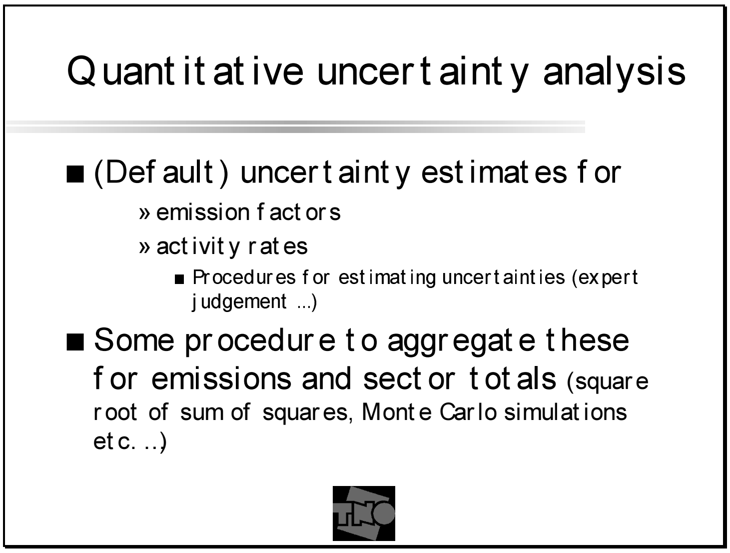
IPCC Guidelines now request a rating “High”, “Medium”, “Low” to be given to reported emissions at the level of the summary tables. No detailed guidance is provided on how to do that. The guidelines however also indicate that uncertainties should be expressed as 95 % confidence intervals.

US EPA developed the DARS system:

- **using a decision tree to attribute a rating (1..10) to each emission factor and activity rate**
- **defining a detailed algorithm to combine these towards an overall rating.**

The Dutch Emission Inventory (in Dutch: Emissieregistratie) is experimenting with a 5 point scale (A, B, C, D, E) originally proposed in US EPA’s AP42

Slide 11



Quantitative

For a quantitative uncertainty analysis quantitative error ranges for activity rates and emission factors are needed. This means that basically also the IPCC Guidelines should contain:

- **guidance to estimate uncertainties quantitatively, i.e.**
- **procedures to estimate uncertainties e.g. expert judgement?**
- **and / or**
- **default values for these error ranges.**

Part of the activities of the IPCC Working Group on Management of Uncertainties is directed towards providing these:

- **guidance on how to combine quantitative uncertainties in variables and parameters towards a quantitative uncertainty estimate for the full inventory:**
- **some algorithm ("square root of sum of squares")**
  - ✓ Monte Carlo simulations
  - ✓ other?
- **guidance on how to report uncertainties.**

The present guidelines mention reporting in 95% confidence intervals. It is expected that this will not be changed.

Slide 12

Default quality ratings

Rating	Definition	typical error range
A	an estimate based on a large number of measurements made at a large number of facilities that fully represent the sector	10 to 30 %
B	an estimate based on a large number of measurements made at a large number of facilities that represent a large part of the sector	20 to 60 %
C	an estimate based on a number of measurements made at a small number of representative facilities, or an engineering judgement based on a number of relevant facts	50 to 150 %
D	an estimate based on single measurements, or an engineering calculation derived from a number of relevant facts	100 to 300 %
E	an estimate based on an engineering calculation derived from assumptions only	order of magnitude

12

Little is known on uncertainties in emission estimates. An attempt has been made to produce default uncertainty ratings at the main sector level in the UNECE EMEP/CORINAIR Guidebook, using the qualitative A, B, C, D, E ratings of US EPA's AP42 manual. The table reproduces the definitions of these ratings.

To help modellers assess the uncertainties in calculated air pollution concentrations, typical error ranges (standard deviations) have been proposed (Guidance Report on Preliminary Assessment under EC Air Quality Directives, Roel van Aalst, Lynne Edwards, Tinus Pulles, Emile De Saeger, Maria Tombrou5, Dag Tønnesen, January 1998, ETC/AQ).

These numbers

- **might provide a feeling of what these qualitative error ranges mean.**
- **pose a challenge to the experts to come up with better default uncertainty ranges**
- **could also provide a link from a qualitative assessment to a quantitative one!**

Slide 13

Default quality ratings

Source sector	CO_2	CH_4	N_2O
1. public power, cogeneration and district heating	A	C	E
2. commercial, institutional & residential combustion	B	C	E
3. industrial combustion	A	C	E
4. industrial processes	B	D	D
5. extraction & distribution of fossil fuels	D	D	
6. solvent use			
7. road transport	B	C	E
8. other mobile sources and machinery	C	D	D
9. waste treatment	B	C	E
disposal activities	C	D	E
10. agriculture activities	C	D	E
11. nature	D	E	E

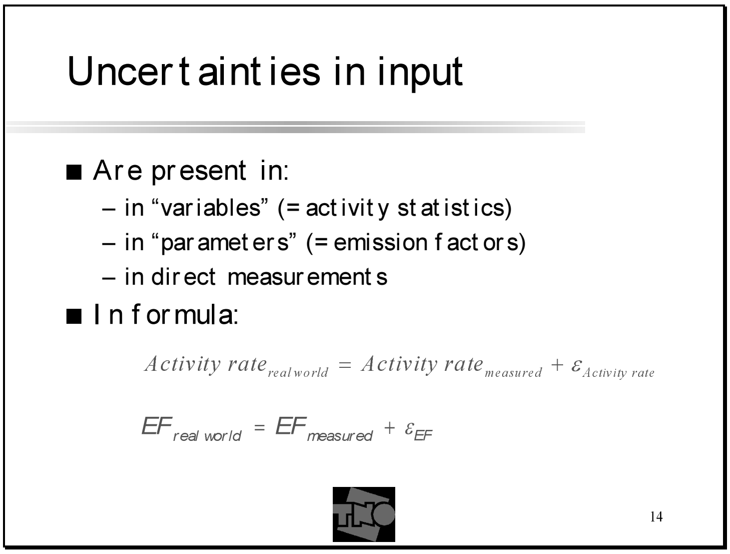
13

The table, derived from the UNECE-EMEP / CORINAIR Guidebook, shows the default data quality ratings by main source sector.
In greenhouse gas emission inventories:

- **uncertainties are the smallest for the most important gas CO_2**
- **uncertainties are smallest in the most important sectors (combustion related: 1, 2, 3, 7 and 8)**

Nevertheless, the reduction targets for 2008 - 2012 (Kyoto) might be well in the range of or even below the over all uncertainty!
These default uncertainty ranges, combined with the typical error ranges, suggested in the previous table, have been used tentatively in a study made for the IPCC Working Group on Management of Uncertainties, using the simple mathematical treatment that follows.

Slide 14



In the following a simple mathematical treatment of error propagation in an emission inventory is presented.

Any emission inventory will be inaccurate by its very nature. The data collected are mostly based upon extrapolation of sample measurements or upon the use of emission factors and activity data. The accuracy of the data will be determined by uncertainties occurring in all stages of the inventorying process. Four sources of uncertainties are discerned:

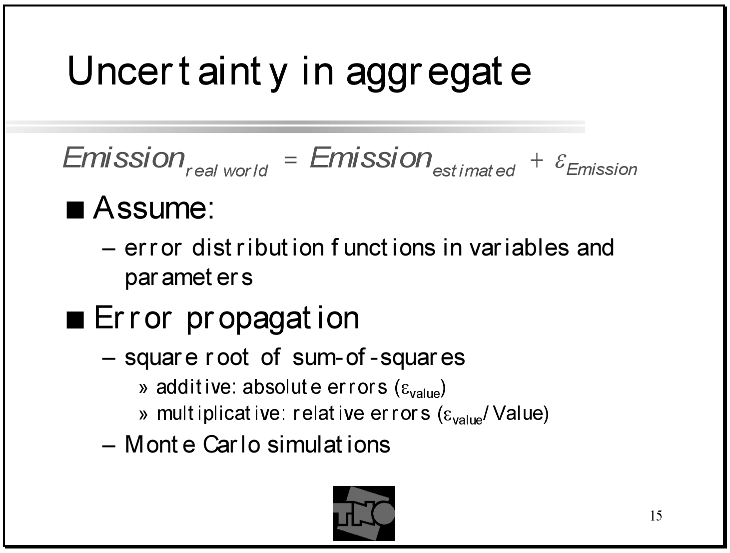
- **Uncertainties originating from the *real variance* of the emissions in time and between different comparable units: some cars have higher emissions than others and emissions of individual cars will vary depending on the state of maintenance.**
- **Uncertainties originating from *variability in the external conditions* in which the units are working: heating emissions will be higher in a cold winter as compared to a warm winter.**

  These uncertainties are due to "naturally" occurring variances in the emissions. As emissions are mostly being estimated by means of sampling, extrapolation of the sample to the total emission might induce errors. Two other sources of uncertainties stem from the fact that no measurement and no inventory can be perfect:

- **Uncertainties *in the measurements* of emissions, emission factors and activity data, including "unknown sources".**
- **Possible *errors in the databases themselves*, including "forgetting" sources.**

The uncertainties will cause a certain level of inaccuracy of the data collected: any value in the inventory may contain an error. All uncertainties in both energy data and emission factors applied to calculate a national emission inventory will be reflected in uncertainties in the final result.

Slide 15

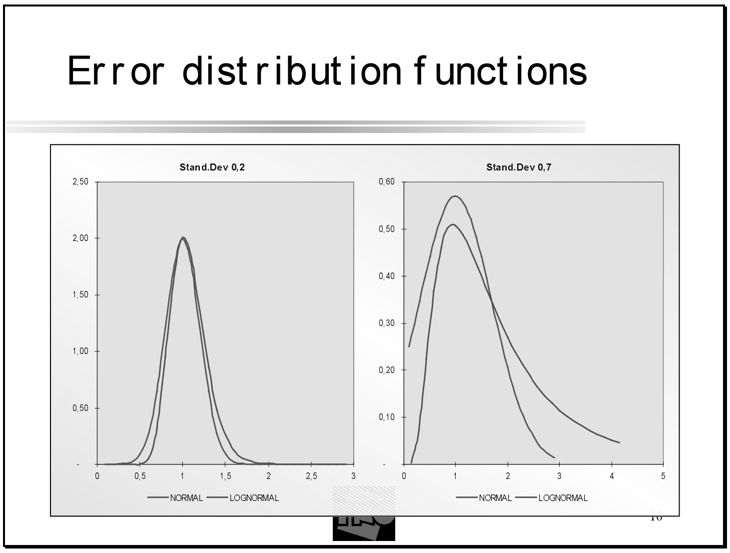


The algorithm as presented above can be applied whenever the standard deviations of the errors in all variables and parameters are available. To find the standard deviation, find an interval, which approximates a 95% confidence interval. In a normal distribution the 95% confidence interval is almost equal to two standard deviations to both sides of the mean value. If such data are not available, as a first estimate the values from the tables could be used. If standard deviations are available the following procedure could be applied:

- **While calculating the emission for each activity and fuel combination, the relative uncertainty in this emission should be calculated as the square root of the sum of squares of the relative uncertainties in both the fuel use and the emission factors (multiplicative operation: use relative errors!).**
- **The absolute uncertainty in the emission of each activity and fuel combination should be derived by multiplying the relative uncertainty with the emission value**
- **The absolute uncertainty in the inventory should be calculated as the square root of the sum of squares of the absolute uncertainties in each separate activity - fuel combination (additive operation: use absolute errors!).**
- **The relative uncertainty in the inventory should be calculated by dividing the absolute error by the total emission.**

This procedure can be applied at the complete inventory, but also by pollutant or for any sector separately. The procedure is not producing mathematically correct standard deviations, but produces a number that is higher when the uncertainty is higher and vice versa and therefore might be used as an indicator of uncertainty.

Slide 16



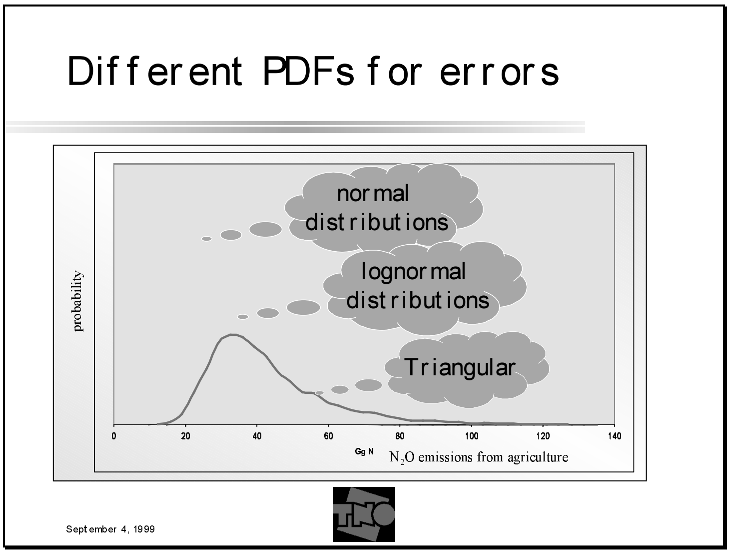
Using more complicated methods, like Monte Carlo simulations, a full description of probability distribution functions for all variables and parameters is needed. This asks for selecting such functions. Two possibilities are normal or lognormal. Both distributions only have two parameters.
Pros and cons between normal and lognormal:

- **With low uncertainty (left)**
  - ✓ not much difference between the two
  - ✓ normal distribution easier to use

- **With high uncertainty (right)**
  - ✓ error distributions look quite differently
  - ✓ normal: zero and negative values could occur; this is unrealistic
  - ✓ lognormal: higher tail: higher probability of high extremes.

Please note that in a normal distribution the mean value and the most probable value (“mode”) are equal. In a lognormal distribution they are not!

Slide 17

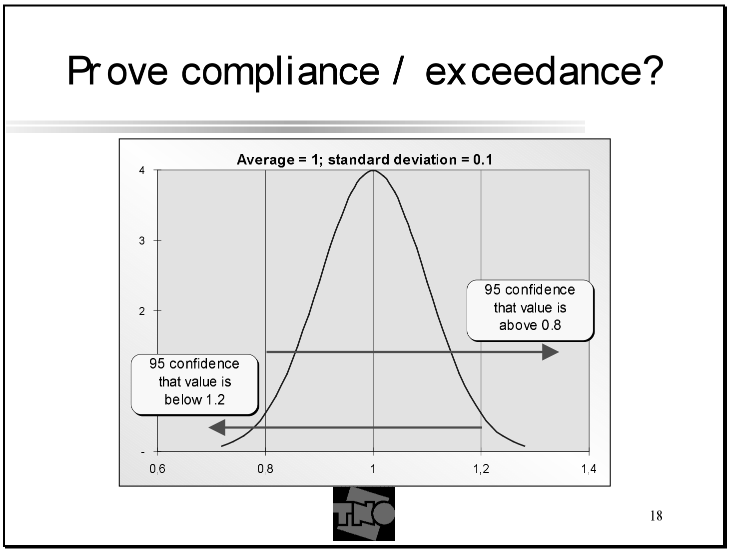


The effect of choosing probability distribution functions for parameters in a complicated emission estimation method (N_2O from agriculture in the Netherlands):

- **The mean value of all these graphs is more or less the same.**
- **This is not true for the most probable value (the mode of the distribution)**
- **The triangular distribution clearly leads to the highest estimate of uncertainty. This is probably too high.**

This analysis also points out that the interpretation of default emission factors and the uncertainty ranges as presented in the literature is not trivial. However the literature does not specify whether these defaults should be understood as mean values or as most probably values.

Slide 18



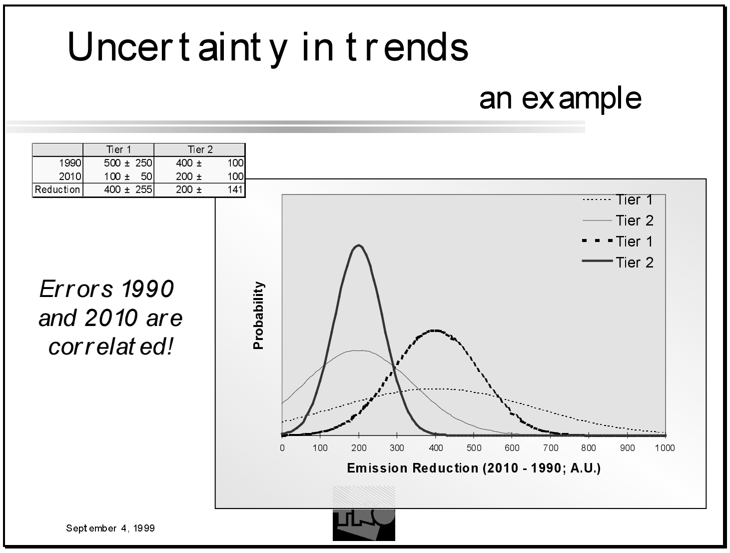
A final item on uncertainties of an emission inventory is presented in the graph above. If the result of an emission inventory shows the above probability distribution function, one should conclude that

- **the probability is 95 % that the total emission is below 1.2**
- **the probability is 95 % that the total emission is above 0.8**

So it depends upon what the user wants to prove (compliance or exceedance) what the value in the conclusion should be!

This observation might be rather relevant for the discussions on the compliance mechanism. When it is decided to report quantitatively on uncertainties, I.e. 95% confidence intervals, this problem of proving compliance versus proving exceedance is important!

Slide 19

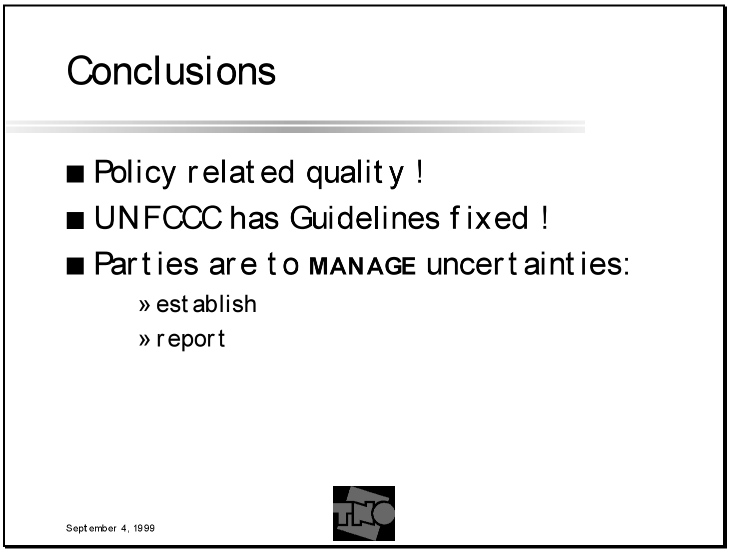


The graph above shows a result of estimating the difference between 1990 and 2010 and the associated uncertainty. The table provides the (assumed) numbers.

If we apply simple error propagation theory, and calculate the uncertainty in the difference between both base years is calculated we find 400 ± 255 and 200 ± 141 for Tier 1 and Tier 2 respectively. These are plotted in the thin lines. In that case both results are compatible. Tier 2 provides a lower reduction with a smaller 95% confidence interval, but both are not significantly different from each other.

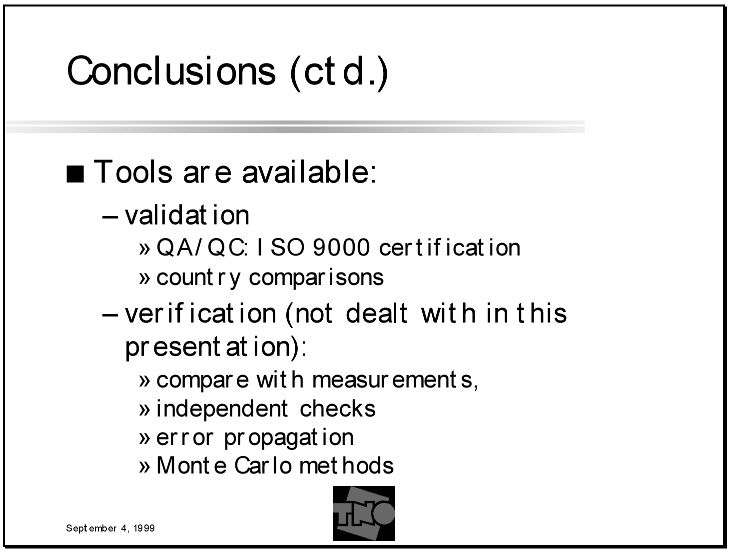
However, it is not unlikely that the errors in the estimates for both base years are correlated, meaning that if the 1990 emission is underestimated, probability is high that the 2010 is also underestimated. So assuming that about 80 % of the errors in both base years are correlated, the thick lines result. Now the probability that the Tier 1 and Tier 2 estimates are different increases. In this hypothetical example however, the probability that both Tiers are compatible still is quite high.

Slide 20



- **While compiling data for reporting to UNFCCC, the policy perspective on data quality is to be used.**
- **It is not expected that UNFCCC will accept any changes in the IPCC Guidelines in the near future. However the reporting format will be changed to a set of much more detailed tables than has been published in national communications until now.**
- **Parties need to manage the uncertainties, not necessarily decreasing them. Uncertainties need to be established and reported.**

Slide 21



Tools for management of uncertainties are partly available:

- **Validation**
  - ✓ Quality Control / Quality Assurance systems will help parties to properly perform and document their inventory, assuring transparency and compliance with the Guidelines and reporting format; this also will include independent auditing ("in depth reviews"?)
  - ✓ Once data are available at the UNFCCC secretariat country comparisons will be used to validate the inventories against each other ("comparability")

- **Verification**
  - ✓ Inventories can be compared with budget studies, air quality measurements and models (forward and inverse modelling) etc.
  - ✓ Independent checks on the total or part of the inventory can be performed
  - ✓ The propagation of uncertainties in the inventory, given the uncertainties in inputs, can be established by error propagation studies or Monte Carlo simulations. To enable this some issues regarding the probability distribution functions and the interpretation of "default values" need to be solved.