

Regional workshop for the “Capacity Development for sustainable national Greenhouse Gas Inventories – AFOLU sector” (CD-REDD II) Programme

Data collection and analysis

Daniela Romano

Institute for Environmental Protection and Research (ISPRA)

Quito, Ecuador

10-13 May 2011

Emission inventories and good practice

- Estimating emissions
 - Methods and data requirements
- Good Practice
 - National inventory system

Emission inventories and good practice

- Main phases in developing and maintaining a national inventory:
 - *planning*
 - *preparation*
 - *management*

The inventory estimation process

Planning, preparation and management of the inventory includes:

- data collection
- selection of methodologies
- activity data and other estimation parameters
- emission factors
- estimation of emissions and removals
- uncertainty assessment
- QA/QC and verification activities
- documentation and archiving

Inventory planning

- collection and processing of data from different data sources;
- selection of appropriate emissions factors and estimation methods consistent with the IPCC Guidelines
- Set priorities for action according to the importance of the sector, putting in place a strategy for accessing, collecting, and processing the data needed for inventory purposes
- Collect data at a level of detail appropriate to the method used

Methodological principles of data collection

- Collect data needed to ‘improve’ estimates of *key categories* which are the largest, have the greatest potential to change, or have the greatest uncertainty
- Choose data collection procedures that improve the quality of the inventory in line with the data quality objectives
- Put in place data collection activities (resource prioritisation)
- Collect data at a level of detail appropriate to the method used
- Review data collection activities and methodological needs on a regular basis
- Introduce agreements with data suppliers to support consistent and continuing information flows

Data collection: starting point

- National Institutes of Statistics
- Sectoral experts, stakeholder organisations
- International organisations publishing statistics e.g., UN, FAO, Eurostat or the International Energy Agency, OECD and the IMF (which maintain international activity and economic data)
- Scientific and technical articles in environmental books, journals and reports
- Universities
- International experts
- IPCC Emission Factor Database
(<http://www.ipcc-nggip.iges.or.jp/EFDB/>)
- National Inventory Reports from Parties to the UNFCCC

Data collection: national and international data

- It is preferable to use national data since national data sources are typically more up to date and provide better links to the originators of the data
- Countries are encouraged to develop and improve national sources of data to avoid being reliant on international data
- Most international datasets rely on nationally-derived data, and in some cases data from reputable international bodies may be more accessible and more applicable to the inventory
- In some cases, groups such as international trade associations or international statistical bodies will have country specific datasets for industries or other economic sectors that are not held by national organisations
- Cross-checking national data sets with any available international data can help to assess completeness and identify possible problems with either data set

Data collection: surveys and census information

- Survey and census information provide the best agricultural, production and energy statistics that can be used for GHG inventories
- Generally these data are compiled by national statistical agencies or relevant ministries for national policy purposes (energy, agriculture and forestry) or to comply with international demand for data, or other activities that are outside of the direct control of the inventory compiler although the needs of the inventory can sometimes trigger or influence surveys or censuses
- Unless they can be consistently repeated, surveys are only able to give measurements relating to one point in time

Data collection: emission factors

Literature sources

- Countries should use their own, peer-reviewed, published literature because this should provide the most accurate representation of the country's practices and activities
- If there are no country-specific studies available, the inventory compiler can use IPCC default factors and data from the Emission Factor Database, or other literature values e.g., modelled/estimated data from international bodies that reflect national circumstances, other countries' experiences

Measured emissions

- Careful consideration of their use to ensure their compatibility with un-measured parts of the inventory

Time series completeness and consistency

Gaps in data sets

For a complete and consistent time series, it is necessary to determine the availability of data for each year.

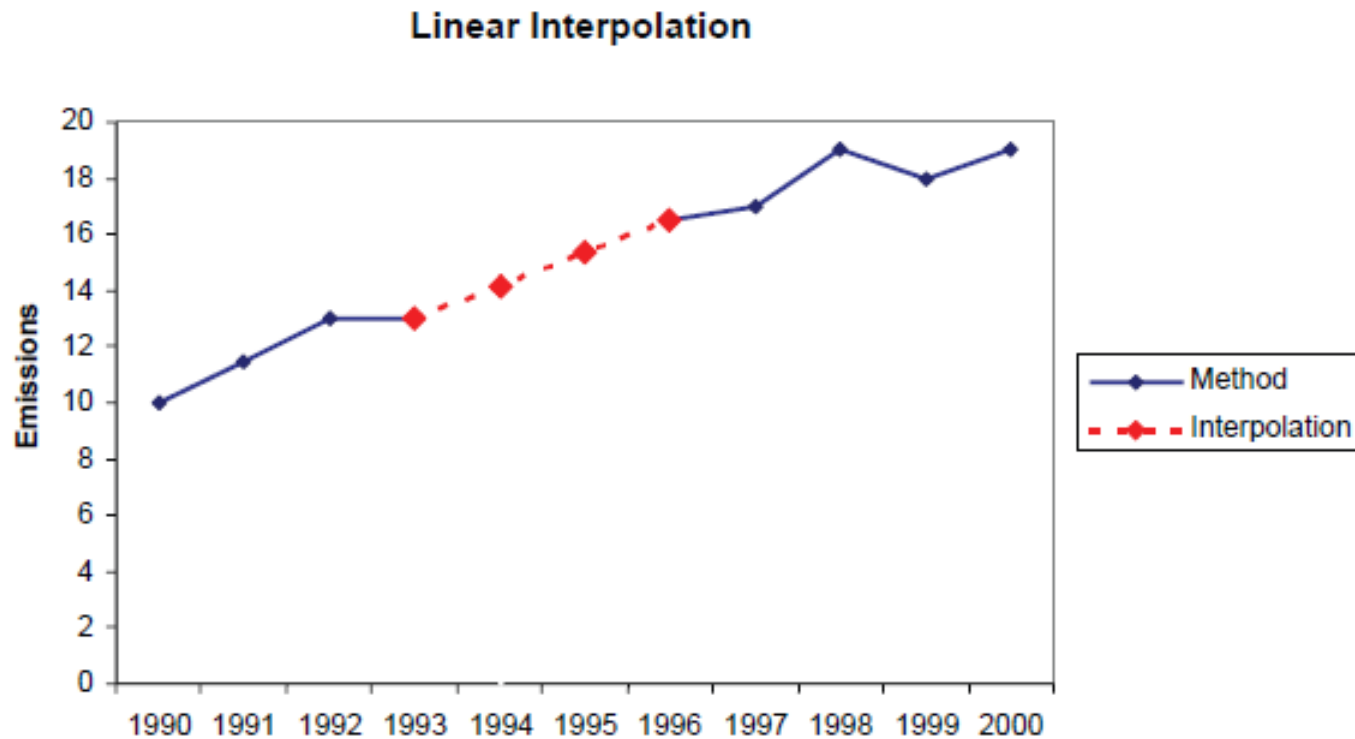
If data are missing for one or more years or the data do not represent the year or national coverage required, they have to be filled in.

For example, time consuming and expensive surveys relating to natural resources - such as national forest inventories - are compiled at intervals of every fifth or tenth year. Time series data may need to be inferred to compile a complete annual estimate for the years between surveys, and for fore- and backcasts

Time series completeness and consistency

Interpolation

Estimates for the intermediate years in the time series can be developed by interpolating between the detailed estimates



Time series completeness and consistency

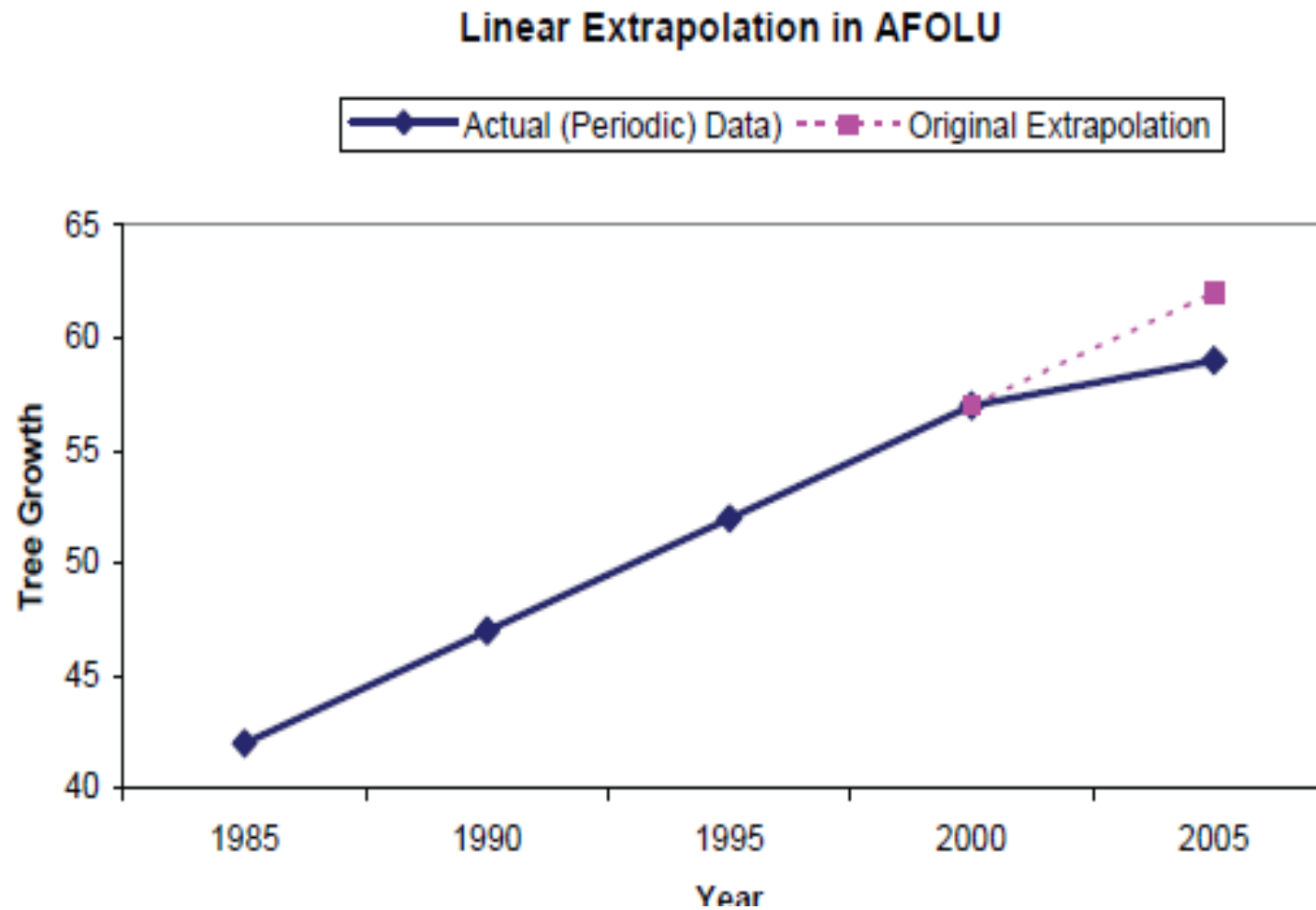
Extrapolation

When detailed estimates have not been prepared for the base year or the most recent year, it may be necessary to extrapolate from the closest detailed estimates

Trend extrapolation simply assumes that the observed trend remains constant over the period of extrapolation.

So the method should not be used if the change in trend is not constant over time or over long periods of time without detailed checks to confirm the continued validity of the trend

Time series completeness and consistency



Time series completeness and consistency

Use of surrogate data

Emissions or removals can be related to underlying activity or other indicative data, then changes in these data are used to simulate the trend

(e.g. road transport emissions may be related to trends in vehicle distances travelled, emissions from domestic wastewater to population)

If information on the general trends or underlying parameters is available, then the surrogate method is preferable

Expert judgement

Ultimately the basis to inventory estimation.

Expert judgment should be elicited using an appropriate protocol

Time series recalculations

The availability of data is a critical determinant of the appropriate method, and thus changes in data may lead to changes or refinements in the estimation process.

Extrapolated figures can be recalculated when new data become available.

The ability of countries to collect data generally improves over time. For example direct sampling and measurement programs can be introduced, new method can be developed to take advantages of new technologies or scientific information (e.g remote sensing), new categories can be added.

The addition to the inventory of a new category or subcategory or new information requires the calculation of an entire time series, and estimates should be included in the inventory from the year emissions or removals start to occur.

Time series recalculations

A country should make every effort to use the same method and data sets for each year

The data may not necessarily be suitable for earlier years as the new program cannot measure past conditions

Sometimes this can be addressed if the new data are sufficiently detailed (e.g., if emission factors for modern abated plant can be distinguished from those of older unabated plant) and the historic activity data can be stratified using expert judgement or surrogate data.

Time series completeness and consistency

Overlap

The overlap technique can be used when a new method is introduced but data are not available to apply it to the earlier years.

The time series can be constructed by assuming that there is a consistent relationship between the results of the previously used and new method. The emission or removal estimates for those years when the new method cannot be used directly are developed by proportionally adjusting the previously developed estimates, based on the relationship observed.

Uncertainty evaluation

- An essential part of an inventory
 - Helps prioritise efforts to improve accuracy
 - Difficult or impossible to quantify and completely characterise all inventory uncertainties
 - Simple approach - Use best available data and expert judgement
- Need uncertainties in all parameters used
 - activity data and emission factors
 - Mean and 95% confidence interval (preferably need pdf as well)

Sources of uncertainty

Activity data

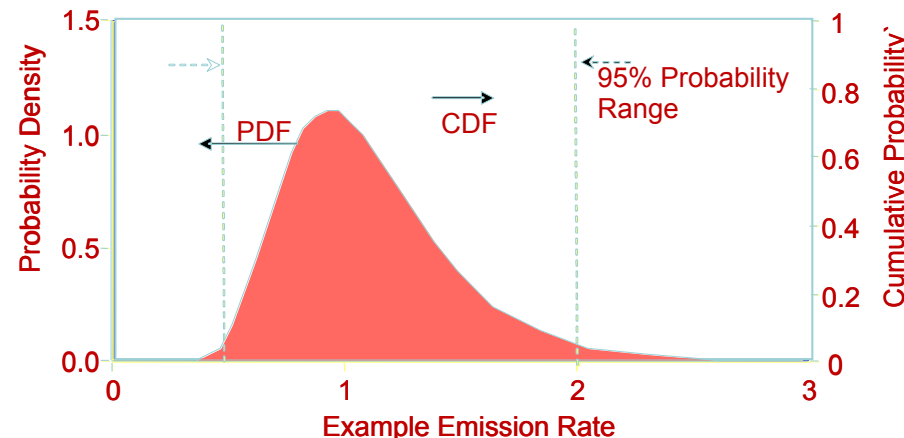
- Gaps in time series
- Use of surrogate or proxy variables
- Lack of references (calculation or estimation methods, representativeness at local or national level)
- Aggregation of data

Emission Factors

- Scarcity of quantitative information (measurements, sample representativeness) as compared to qualitative information (experts judgement)
- Statistical Random Sampling Error
- Measurement error

Uncertainty and variability

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



- **Variability:** Heterogeneity of a variable over time, space or members of a population. Variability may arise, for example, due to differences in design from one emitter to another (inter-plant or spatial variability) and in operating conditions from one time to another at a given emitter

Uncertainty analysis – Approach 1

- Information provided in the IPCC Guidelines as well as expert judgement can be used
- Standard deviations can also be considered when measurements available
- General approach:
set values within a range low, medium and high according to the confidence the expert has on the value

Uncertainty analysis – Tier 1

Activity data

- low uncertainty (e.g. 3-5%) to activity data derived from the energy balance and statistical yearbooks
- medium-high uncertainty (20-50%) to the data not directly or only partially derived from census or sample surveys or estimated data

Emission factors

- IPCC uncertainty values are used when the emission factor is a default value or when appropriate
- low values are used for measured data
- otherwise uncertainty values are high

Archive and Documentation

- Rationale for choice of methods
- Assumptions and criteria for selection of activity data and emission factors
 - Emission factors used, including references
 - Activity data or sufficient information to it to be traced to the referenced source
 - Information on the uncertainty associated with activity data and emission factors
- Changes in data inputs or methods from previous years
- Databases or software used and information about their use
- Worksheets, calculations, aggregated estimates and any recalculations of previous estimates
- QA/QC plans and outcomes of QA/QC procedures

Conclusions

- It is important to maintain supply of inventory data
- Engage data suppliers in the process of inventory compilation and improvement by involving them in activities such as:
- Offering an initial estimate for the category, pointing out the potentially high uncertainties and inviting potential data suppliers to collaborate in improving estimates
- Scientific or statistical workshops on the inventory inputs and outputs
- Specific contracts or agreements for regular data supply
- Regular/annual informal updates on the methods that use their data
- Establishment of terms of reference or memoranda of understanding for government and/or trade organisations providing data to clarify what is needed for the inventory, how it is derived and provided to the inventory compiler and when

Conclusions

- Maintaining good inventory management principles will ensure the efficient and timely delivery of high quality inventory data
- An inventory management system needs to be established and should include:
- a **clear inventory process** so that key activities and resources can be focused towards delivery deadlines and delivery quality;
- **institutional arrangements:** clearly defined roles and responsibilities for delivering the inventory to specified time and quality standards;
- a **quality framework** to ensure that the data is fit for inventory purpose

Regional workshop for the “Capacity Development for sustainable national Greenhouse Gas Inventories – AFOLU sector” (CD-REDD II) Programme

Thank you

Quito, Ecuador

10-13 May 2011

Regional workshop for the “Capacity Development for sustainable national Greenhouse Gas Inventories – AFOLU sector” (CD-REDD II) Programme

Examples of data collection and analysis

Land-use

- The *classification* of national territory into the six *land-use categories* (*forest land, cropland, grassland, wetlands, settlements, other land*) is the basis of estimating and reporting greenhouse gas emissions and removals from land use and land-use conversions.
- The categories may be further *stratified* (e.g., by climate or ecological zones). The categories (and sub-categories) are intended to be identified through use of Approaches for representing land-use area data before described.
- Care needs to be taken in inferring *land use* from the *land cover* characteristics and vice versa.
- Countries will use their own *definitions of these categories*, which may or may not refer to internationally accepted definitions, such as those by FAO, Ramsar, etc.
- Countries should describe the methods and definitions used to determine areas of *managed and unmanaged lands*.
- All land definitions and classifications should be specified at the national level, described in a transparent manner, and be applied consistently over time.

Land representation

Three Approaches may be used to represent areas of land use categories.

The Approaches are not presented as hierarchical tiers and do not imply any increase or decrease in accuracy. The Approaches are not mutually exclusive, and the mix of Approaches selected by a country should reflect emissions estimation needs and national circumstances. In all cases, countries should characterize and account for all relevant land areas in a country consistently and as transparently as possible.

Approach 1

- identifies the total change in area for each individual land-use category within a country, but does not provide information on the nature and area of conversions between land uses;
- represents land-use area totals within a defined spatial unit, which is often defined by administrative boundaries. Only the net changes in land-use area can be tracked through time.
- The land-use area data may come originally from periodic sample survey data, maps or censuses, but will probably not be spatially explicit. The sum of all land-use category areas may or may not equal the total area of the country or region under consideration. The final result of this Approach is a table of land use at given points in time.

Land representation

Approach 2

- introduces tracking of land-use conversions between categories (but is not spatially explicit);
- provides an assessment of both the net losses or gains in the area of specific land-use categories and what these conversions represent (i.e., changes both from and to a category).
- The final result of this Approach can be presented as a non spatially-explicit **land-use conversion matrix**. The matrix form is a compact format for representing the areas that have come under different conversions between all possible land-use categories.

NET Land-use Conversion Matrix							
Final \ Initial	F	G	C	W	S	O	Final sum
F	15	3	1				19
G	2	80					82
C			29				29
W				0			0
S	1	1	1		5		8
O						2	2
Initial sum	18	84	31	0	5	2	140

Note:
F = Forest land, G = Grassland, C = Cropland, W = Wetlands,
S = Settlements, O = Other land
Numbers represent area units (Mha in this example).

Land representation

Approach 3

- extends Approach 2 by allowing land use conversions to be tracked on a spatially explicit basis;
- is characterized by spatially-explicit observations of land-use categories and land-use conversions, often tracking patterns at specific point locations and/or using gridded map products, such as derived from remote sensing imagery. The data may be obtained by various sampling, wall-to-wall mapping techniques, or combination of the two methods;
- analysis tools such as Geographic Information Systems can be used to link multiple spatially Explicit data sets (such as those used for stratification) and describe in detail the conditions on a particular piece of land prior to and after a land-use conversion. This analytical capacity can improve emissions estimates by better aligning land-use categories (and conversions) with strata mapped for classification of carbon stocks and emission factors by soil type, vegetation type.

Land-use databases

There are three broad sources of data for the land-use databases needed for greenhouse gas inventories:

1. databases prepared for other purposes

- national databases;

typical sources of data include forest inventories, agricultural census and other surveys, censuses for urban and natural land, land registry data and maps.

- international databases;

Several projects have been undertaken to develop international land-use and land cover datasets at regional to global scales. Almost all of these datasets are stored as raster data generated using different kinds of satellite remote sensing imagery, complemented by ground reference data obtained by field survey or comparison with existing statistics/maps. These datasets can be used to estimate spatial distribution of land-use categories and assess reliability of the existing land-use datasets.

2. collection by sampling

- Sampling techniques for estimating areas and area changes are applied in situations where direct measurements in the field or assessments by remote sensing techniques are not feasible or would provide inaccurate results.
- Sampling usually involves a set of sampling units that are located on a regular grid within the inventory area. Where sampling for areas is repeated at successive occasions, area changes over time can be derived to construct land-use conversion matrices.

Land-use databases

3. complete land inventory

- A complete inventory of land use of all areas in a country will entail obtaining maps of land use throughout the country at regular intervals. This can be achieved by using remote sensing techniques.
- A complete inventory can also be achieved by surveying all landowners. Inherent problems in the method include obtaining data at scales smaller than the size of the owner's land as well as difficulties with ensuring complete coverage with no overlaps.

Tools for Data Collection

→ Remote sensing techniques

A complete remote sensing system (aerial photographs, satellite imagery using visible and/or near-infrared bands, satellite or airborne radar imagery and lidar) for tracking land use conversions can include many sensor and data type combinations at a variety of resolutions.

→ Ground-based surveys

may be used to gather and record information on land use, and for use as independent ground-truth data for remote sensing classification.

Reporting under Kyoto Protocol rules

Land related information

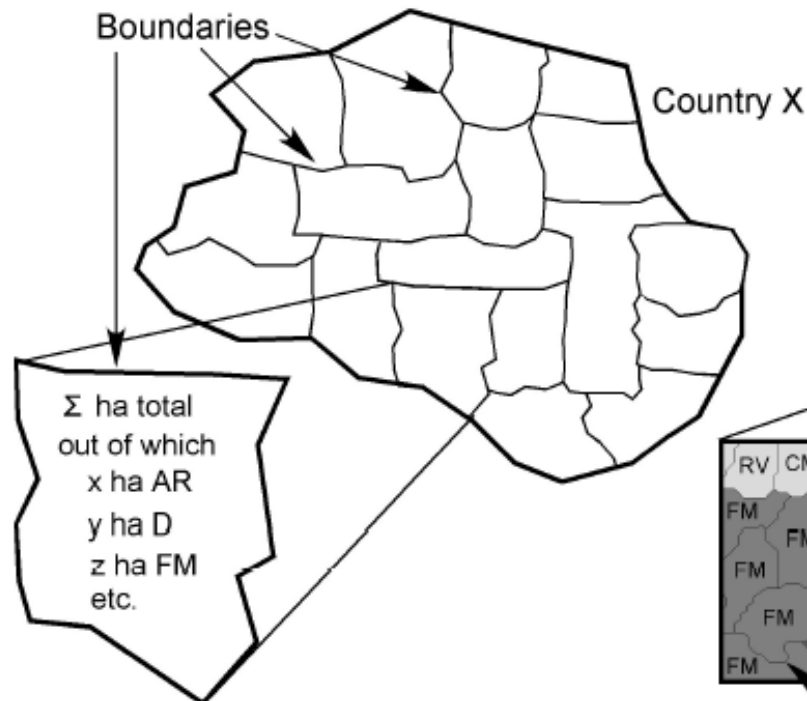
- Maps and/or database to identify the geographical locations, and the system of identification codes for the geographical locations:

A clear description of used reporting method (1 or 2) has to be provided;

Areas of lands subject to Art. 3.3 and 3.4 activities need to be identifiable, adequately reported and tracked in the future.

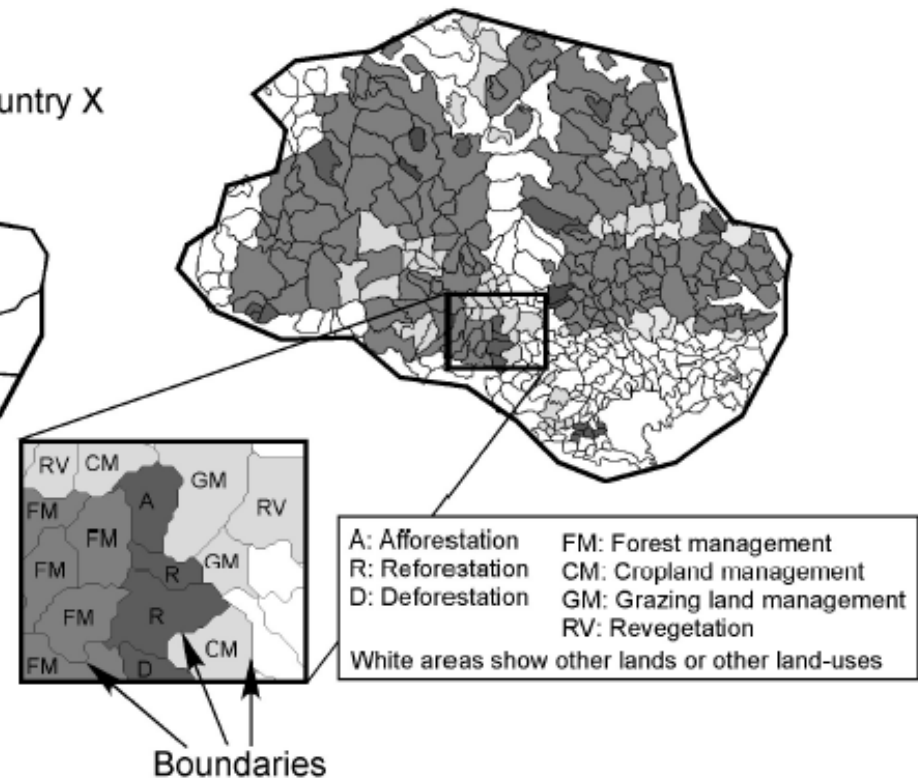
Reporting Method 1

A geographic boundary encompasses units of land or land subject to multiple activities.



Reporting Method 2

A geographic boundary encompasses units of land or land only subject to a single activity.



Emissions from Livestock and Manure management

Livestock production can result in methane (CH₄) emissions from enteric fermentation and both CH₄ and nitrous oxide (N₂O) emissions from livestock manure management systems.

The methods for estimating CH₄ and N₂O emissions from livestock require definitions of livestock subcategories, annual populations and, for higher Tier methods, feed intake and characterization.

Livestock population and feed characterization

- Good practice is to identify the appropriate method for estimating emissions for each category, and then base the characterization on the most detailed requirements identified for each livestock species.
- The livestock characterization used by a country will probably undergo iterations as the needs of each source category are assessed during the emissions estimation process.

Livestock population and feed characterization

Livestock species and categories:

A complete list of all livestock populations that have default emission factor values must be developed (e.g., dairy cows, other cattle, buffalo, sheep, goats, camels, llamas, alpacas, deer, horses, rabbits, mules and asses, swine, and poultry) if these categories are relevant to the country.

Identify the most detailed characterization required for each livestock species

On the basis of the assessments for each species under each category, the most detailed characterization required to support each emission estimate for each species should be identified.

The 'Basic' characterization can be used across all relevant source categories if the enteric fermentation and manure sources are both estimated with their Tier 1 methods. An 'Enhanced' characterization should be used to estimate emissions across all the relevant sources if the Tier 2 method is used for either enteric fermentation or manure.

Choice of method: Tier 1

Annual population: use population data from official national statistics or industry sources. FAO data can be used if national data are unavailable.

Seasonal births or slaughters may cause the population size to expand or contract at different times of the year, which will require the population numbers to be adjusted accordingly. Most animals in growing populations (e.g., meat animals, such as broilers, turkeys, beef cattle, and market swine) are alive for only part of a complete year.

The following relation estimates the **annual average of livestock population**

$$AAP = Days_alive \cdot \frac{NAPA}{365}$$

where:

AAP = *annual average population*

$NAPA$ = *number of animals produced annually*

Choice of method: Tier 1

Dairy Cows and Milk Production:

The dairy cow population is estimated separately from other cattle. Dairy cows are defined as mature cows that are producing milk in commercial quantities for human consumption. This definition corresponds to the dairy cow population reported in the FAO Production Yearbook. Dairy buffalo may be categorized in a similar manner.

Data on the average milk production of dairy cows is also required. Country-specific data sources are preferred, but FAO data may also be used. If two or more dairy cow categories are defined, the average milk production per cow is required for each category.

Choice of method: Tier 2

Tier 2: Enhanced characterization for livestock populations

The Tier 2 characterization methodology seeks to define animals, animal productivity, diet quality and management circumstances to support a more accurate estimate of feed intake for use in estimating methane production from enteric fermentation.

The Tier 2 livestock characterization requires detailed information on:

- Definitions for livestock sub-categories;
- Livestock population by sub-category, with consideration for estimation of annual population as per Tier 1;
- Feed intake estimates for the typical animal in each sub-category.

The livestock population subcategories are defined to create relatively homogenous sub groupings of animals. By dividing the population into these subcategories, country specific variations in age structure and animal performance within the overall livestock population can be reflected.

Livestock population: example

For each category identified, the time series of average population has to be developed

- a) National Statistics. For example, poultry population may be subdivided according to the different types of growing cycles and maintenance system (dry or wet system, confinement, free range).
- b) FAO data can be used in the assessment of animal population; milk production can also be derived from FAO database
- c) In case of multiple source of data (national statistics, FAO, industry sources) attention has to be paid to the different times of the year which data refer to. For example, if national statistics are available only for few years at an identified disaggregation level, these data may be used to disaggregate non national statistics available at aggregated level.
- d) For some livestock categories (poultry, rabbits), meat production may be used as a proxy variable to infer the animal population
- e) Default values like weight, feed digestibility or EFs to assess emissions from livestock and manure management are also available in the 2006 IPCC guidelines.

Solid waste disposal on land

- Parameter values used in the landfill emissions model are:
 - 1) total amount of waste disposed;
 - 2) fraction of Degradable Organic Carbon (DOC);
 - 3) fraction of DOC dissimilated (DOC_F);
 - 4) fraction of methane in landfill gas (F);
 - 5) oxidation factor (O_X);
 - 6) methane correction factor (MCF);
 - 7) methane generation rate constant (k);
 - 8) landfill gas recovered (R).
- **Tier 1:** estimations are based on the IPCC FOD method using mainly default activity data and default parameters.
- **Tier 2:** use the IPCC FOD method and some default parameters, but require good quality country-specific activity data on current and historical waste disposal at SWDS. Historical waste disposal data for 10 years or more should be based on country-specific statistics, surveys or other similar sources
- **Tier 3:** use of good quality country-specific activity data and the use of either the FOD method with (1) nationally developed key parameters, or (2) measurement derived country-specific parameters. Key parameters should include the half-life, and either methane generation potential (L_0) or DOC content in waste and the fraction of DOC which decomposes (DOC_f).

Solid waste disposal on land: example

Municipal solid waste

Default data

Activity data consist of the waste generation for bulk waste or by waste component and the fraction of waste disposed to SWDS. Waste generation is the product of the per capita waste generation rate (tonnes/capita/yr) for each component and population (capita). Regional default values for MSW can be found in the IPCC guidelines for the generation rate and the fraction disposed in SWDS, and for the waste composition. Changes in waste management practices (e.g., site covering/capping, leachate drainage improvement, compacting, and prohibition of hazardous waste disposal together with MSW) should also be taken into account when compiling historical data.

National data

Country-specific MSW generation, composition and management practices may be obtained from waste statistics, surveys (municipal or other relevant administration, waste management companies, waste association organisations, other) and research projects (World Bank, OECD, ADB, JICA, U.S.EPA, IIASA, EEA, etc.).

Solid waste disposal on land: example

The FOD methods require data on solid waste disposal (amounts and composition) collected by default for 50 years or more. Countries that do not have historical statistical data for the whole period will need to estimate these data using surrogates (extrapolation with population, economic or other drivers). The historical data could be proportional to economic indicators, or combinations of population and economic indicators. Waste management policies to reduce waste generation and to promote alternatives to solid waste disposal should be taken into account in the analyses.

- Historical data on population and GDP (or other economic indicators) can be obtained from national statistics or from UN international databases

Population data (1950 onwards with five-year intervals) (<http://esa.un.org/unpp/>).

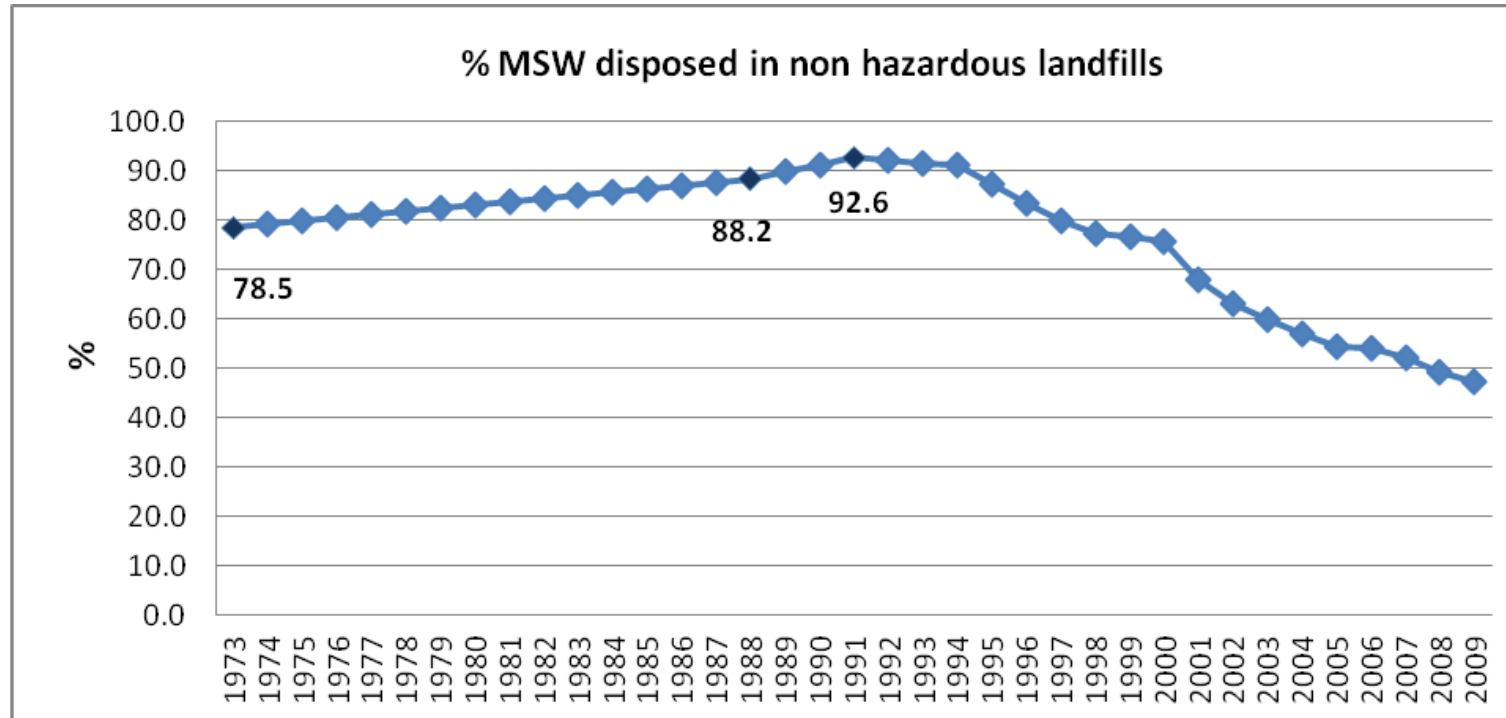
GDP data (1970 onwards, annual data at current prices in national currency)

(<http://unstats.un.org/unsd/snaama/selectionbasicFast.asp>).

Solid waste disposal on land: Italian example

- Since waste production data are not available before 1975, they have been reconstructed on the basis of proxy variables. Gross Domestic Product data have been collected from 1950 (ISTAT, several years [a]) and a correlation function between GDP and waste production has been derived from 1975; the exponential equation has been applied from 1975 back to 1950.
- The amount of waste disposed into landfills has been estimated, assuming that from 1975 backwards the percentage of waste landfilled is constant and equal to 80%; this percentage has been derived from the analysis of available data. As reported in the Figure 8.1, in the period 1973 – 1991 data are available for specific years (available data are reported in dark blue, whereas estimated data are reported in light blue). From 1973 to 1991 waste disposal has increased, because the most common practice in waste management; from early nineties, thanks to a change in national policies, waste disposal in landfill has started to decrease, in favour of other waste treatments.

Solid waste disposal on land: Italian example



Comparison among different Tiers: Italian example

Comparison of CF₄ emissions time series following Tier 1 and Tier 1+Tier 2 and different source of EFs

