What does IPCC say about statistics?

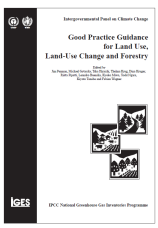
Dr. Lutz Fehrmann
Regional Course on REDD+, MRV and Monitoring
11-15 July 2011
Sokoine University of Agriculture, Morogoro, Tanzania

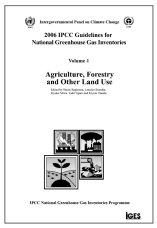
What does IPCC say about statistics?

- In more detail: What does IPCC say about statistical sampling (and reporting) of
  - **Changes in Area** (LUC), also „Activity data“ if result of human intervention,
  - **Mean carbon stocks** in different land use classes
    - As basis to estimate „Emission factors“ in case of LUC

IPCC Guidelines

- It is likely that a REDD mechanism would draw on established IPCC guidelines for national reporting on greenhouse gas emissions and land use change.





IPCC Guidelines

- 2006 IPCC Guidelines for National Greenhouse Gas Inventories (vol.1, Chap.3) provides general information on how to deal with "uncertainties" and describe statistical terminology.
- GPG 2000 (and GPG LULUCF 2003, SR LULUCF) addresses further issues regarding the quantification and reporting of errors.
- Nevertheless, there is **less guidance on how to influence uncertainty of estimates and/or how to optimize inventories (for good reason?)**.

GPG Glossary

GOOD PRACTICE

"…a set of procedures intended to ensure that greenhouse gas inventories are **accurate** in the sense that they are **systematically neither over nor underestimates** so far as can be judged, and that **uncertainties are reduced so far as possible.**"

Translated to statistical terminology: Estimates should be possibly **unbiased** and **precise as possible**.

- No precision defined!

IPCC Guidelines

- "Good practice guidance further supports the development of inventories that are **transparent**, **documented**, **consistent over time**, **complete**, **comparable**, **assessed for uncertainties**, **subject to quality control** and assurance, **efficient** in the use of resources available to inventory agencies, …".

Quantifying errors (GPG)

It is good practice to:

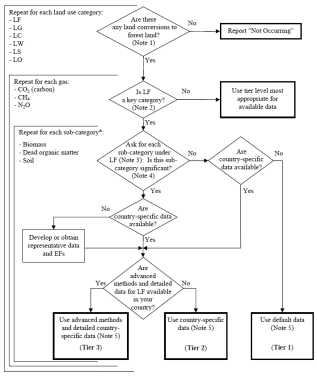
- **assess the uncertainty of the estimates strictly according to the principles of sampling theory**, taking into account what sampling design and estimator were used.
- apply standard sampling theory for deriving the uncertainty estimates, rather than simple error propagation (Tier 1).
- **assess the applicability of core models** for the target population.
- **conduct careful check assessments** on a fraction of the plots.

Tiers

"Tiers correspond to a progression from the use of simple equations with default data to country-specific data in more complex national systems"

Tiers

- Choice of tier-level:



How to assess LUC (IPCC)?

- Approach 1: BASIC LAND-USE DATA
  - Based on available data on LU in two different points in time

TABLE 2.3.1
EXAMPLE OF APPROACH 1:
AVAILABLE LAND -USE DATA WITH COMPLETE TERRITORIAL COVERAGE

Time 1			Time 2			Land-Use Change between Time 1 and Time 2		
F	=	18	F	=	19	Forest	=	+1
G	=	84	G	=	82	Grassland	=	-2
C	=	31	C	=	29	Cropland	=	-2
W	=	0	W	=	0	Wetlands	=	0
S	=	5	S	=	8	Settlements	=	+3
O	=	2	O	=	2	Other land	=	0
Sum	=	*140*	*Sum*	=	*140*	*Sum*	=	*0*

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

  - But: both area estimates are affected by errors!?
  - An error of 5.55% would lead to T1=18±1, T2=19±1

How to assess LUC (IPCC)?

- Approach 2: SURVEY OF LAND USE AND LAND-USE CHANGE

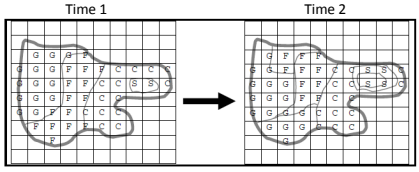
TABLE 2.3.5
SIMPLIFIED LAND-USE CHANGE MATRIX FOR EXAMPLE APPROACH 2

Land-Use Change Matrix							
Final \ Initial	F	G	C	W	S	O	*Final sum*
F	15	3	1				*19*
G	2	80					*82*
C			29				*29*
W							
S	1	1	1		5		*8*
O						2	*2*
Initial sum	*18*	*84*	*31*		*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).
There is no Wetlands in this example. Blank entry indicates no land use change.

How to assess LUC (IPCC)?

- Approach 3: GEOGRAPHICALLY EXPLICIT LAND USE DATA



Error of LUC (IPCC)

- "It is *good practice* *to present* an estimate of the accuracy of the land-use/cover map category-by-category **and a confusion matrix may be employed for this purpose** where remote sensing is used."

- Does a confusion matrix allow to estimate an accuracy or precision like defined by IPCC (bias of an estimator, Standard error of an estimate)? No!

IPCC

2.4.4.2 GROUND-BASED SURVEYS

"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. Prior to the advent of remote sensing techniques such as aerial photography and satellite imagery, ground-based surveys were the only means of generating maps."

Sampling design and stratification (IPCC)

- Sampling designs should aim for a good **compromise between simplicity and efficiency**, and this can be promoted by following three aspects of good practice:
  - Use of auxiliary data and stratification;
  - Systematic sampling;
  - Permanent sample plots and time series data.

Efficiency

- Example: If a „high quality“ estimate of LUC (say SE=<5% for „forest land converted to grassland“) is combined with an estimate of related „Emission factors“ (with SE=20%, including model errors and sampling error)

16

The role of remote sensing

- It is interesting to note that RS is regarded as important source of useful ancillary information from the viewpoint of forest inventory experts,
  - for stratification
  - for model based observations (considering the model errors!)
- Contrary, RS experts often describe field sampling as ancillary information to „calibrate“ image classification algorithms!

17