



2nd Expert workshop on
lessons learned from Accuracy
Assessments in the context of REDD+

Oslo, Norway

26-28. June 2017

Proceedings

28-Aug-17

Organized by:

Global Forest Observation Initiative (GFOI)
Global Observations of Forest and Land Dynamics (GOFC-GOLD)
Norwegian Space Center (NSC)

Contents

1.	Background	3
	Objectives	4
	Sponsors.....	4
	Organizing committee leaders	5
2.	Meeting summary.....	6
	a. About providing guidance	6
	b. Actions	6
	Glossary for MGD	6
	Overview module (FAQ structure)	7
	Technical module: Statistical inference	8
	Technical module: Reference data.....	9
	Next meetings / other.....	10
3.	Agenda	11
4.	Participants	12

1. Background

UNFCCC negotiations have identified the need to establish national forest monitoring systems that use an appropriate combination of remote sensing and ground-based forest carbon inventory approaches for estimating anthropogenic forest-related greenhouse gas emissions by sources, and removals by sinks to support REDD+ implementation and assessing performance in implementing REDD+ activities. Countries and jurisdictions have been advancing in the development of their Reference Levels in order to be able to measure performance of their REDD+ activities in the context of results based payments. In order to set their Reference Levels, countries have to generate forest cover change information to estimate activity data that would facilitate combination with Emission Factors and estimate (historical) emissions. In this context, countries are producing forest cover change estimates as well as maps and are validating these through accuracy assessments. The application of Olofsson et al. (2014) is commonly used in the design of these accuracy assessments. It is becoming more and more common to replace map estimates with estimates obtained using statistically unbiased estimators based on the reference sample data typically used to produce the confusion matrices. However, the experience in the application of this procedure has revealed some challenges that are not covered with existing guidance, e.g. implications of very imprecise estimates as well as implications for monitoring.

The application of Olofsson et al. (2014) has identified multiple issues that require some additional guidance. In view of this, it is becoming timely but urgent to gather lessons learned and constraints, and to provide additional guidance. A dedicated expert meeting and technical workshop was held Monday 3 and Tuesday 4 April 2017 at FAO. The meeting brought together several experts and international partners to discuss assessments for map accuracy purposes versus assessments for unbiased estimation of areas and how the two different objectives can be addressed in country monitoring. Some summary notes of the meeting are provided in Annex A. The outcomes of the workshop resulted in a series of concrete activities to resolve practical issues in country accuracy assessment activities. It was also clear that, based on these experiences improved guidance can be provided through additional expert synthesis (workshops) while other issues require more experiences and additional research.

The GFOI R&D coordination component (lead by the GOF-C-GOLD land cover office, <http://www.gofcgold.wur.nl>) holds regular expert workshops with the aim to provide expert synthesis and guidance to countries through GFOI (see <http://www.gfoi.org/rd/>), the most recent one being the GFOI R&D and GOF-C-GOLD Science Meeting (The Hague, Netherlands, Oct. 31-Nov. 4, 2016, http://www.gofcgold.wur.nl/sites/gofcgold-gfoi_sciencemeeting2016.php). In the past, the GFOI R&D coordination and MGD Component have been successful in writing MGD modules to target issues that require better guidance to countries and examples pertaining to the development of national forest monitoring systems in tropical MRV systems. These modules have since been incorporated into MGD2 and include:

1. Inference for activity data (Section 5.1.5 in MGD2). A workshop was organized in Boston 2013 by Boston University that brought together authors (McRoberts, Næsset, Olofsson, and others) of various estimation protocols for the purpose of estimating areas of activity data. The aim of the workshop was to provide guidance for estimation of activity data. The result of this effort was a decision tree with related text that guides practitioners through various decisions related to the nature of available maps and reference observations. It also generated a longer review of estimators and sampling designs of relevance for estimating activity data, which that is currently being prepared for submission to a scientific journal.
2. Inference for emissions factors (Section 5.2.5 in MGD2). A workshop was organized in Oslo 2014 by the Norwegian Space Centre to develop similar material as described above but for

emissions factors. Göran Ståhl and Christoph Kleinn among others participated in addition to many of the authors of the activity data inference section. A decision tree and related text were the main outcomes of the workshop, which have been incorporated into MGD2.

3. The use global forest cover change maps (Section 4.1.7 in MGD2). Following the writing of MGD1 was the release of a global forest cover change map at 30 m spatial resolution by the Hansen lab at University of Maryland. SilvaCarbon identified a need for guidance on the use of this dataset in national forest monitoring systems. The new leadership of the GFOI R&D component by GOFC-GOLD organized a larger workshop in Wageningen in 2015 with the aim of providing tangible guidelines to be provided to the community. A five-page document outlining recommendations and guidelines was authored and published as a MGD module until incorporated into MGD2.

At the last GFOI plenary, several items have been discussed as key priorities for future R&D expert workshop. These topics include:

- Follow up from 3/4 April expert meeting: Sampling design, stratification options for area estimation, how to decrease uncertainty (country experiences/examples, new guidance, R&D gaps)
- 3/4 April expert meeting /The Hague meeting outcomes: Emission factor uncertainty (relation to total C-emissions) and uncertainty of the trend
- Assess the potential inconsistency in MGD2/The Hague meeting outcomes: Uncertainty estimation from “fully integrated tools” (wall-to-wall output such as maps of carbon emissions and removals) for Tier 3 reporting

It was decided (at the plenary) that the next workshop should focus on a direct follow up from 3/4 April expert meeting. The remaining topics are also very important and should be subject to future expert meetings.

Objectives

Overall, we envision a similar workshop to previous ones to actively work on providing improved guidance on accuracy assessments in the context of REDD+. In particular, the workshop would aim to provide guidance to practitioners to meet the IPCC good practice criteria related to bias and uncertainty in estimates of activity data. More specifically we aim to:

- Discuss and present country examples and case studies for accuracy assessments
- Draw lessons learned and synthesize them into good practice for accuracy assessments
- Discuss how understanding of uncertainties can feed back into efforts for continuous improvements
- Scope a new MGD module capturing the new guidance
- Discuss additional related and follow-up issues for GFOI R&D synthesis (i.e. EFs) and contributions to 2019 methodological refinement of the IPCC GPG

For a summary of the first meeting on this topic view the [announcement](#).

Sponsors

Norwegian International Climate and Forest Initiative (NICFI)

Norwegian Space Center
US Silvacarbon Program
Global Forest Observations Initiative (GFOI) R&D component
Center for International Forest Research (CIFOR)
European Space Agency (ESA)
Worldbank FCPF

Organizing committee leaders

Evie Hagen (NSC)
Martin Herold and Sarah Carter (GFOI, GOFC-GOLD)

2. Meeting summary

a. About providing guidance

- To inform users: When do you need to be worried about something, and what can you do about it?
- Should ensure that things are approached systematically.
- We must build on what was done in the Boston workshop -> the decision tree, based on statistical theory is a great starting point, but now we have to deal with the 'special cases' for which more information is required.
- We should show how to use/apply the guidance (it is not about new guidance).
- We need to have examples as well as technical details, which are easier for users to grasp.
- We can also clarify issues - sometimes we are just saying this is an issue, but not providing a solution.
- It should not be premature guidance, on issues which we are not sure about.
- Should give countries confidence to move forward in their space.
- These guidance 'modules' were previously termed "rapid response modules" in response to urgent problems, which are not covered in the guidance.
- The MGD2 is only 6 months old, and the next update will not be until after the next IPCC guidelines in 2019. So the aim is not to do a revision of this yet, but to provide examples which will compliment MGD2.
- One option is to give guidance as a kind of FAQ format. This was preferred in some contexts.

b. Actions

Key action points have been outlined in detail in this section, and include creating a glossary of terms for the MGD as well as 3 new 'modules'.

Glossary for MGD

- Should be aligned with IPCC definitions, but should not assume that IPCC definitions are OK, but best to base our work on this.
- Pontus has compiled a list of words and definitions from published sources, and will send to Carly
- Should not be like a text book, but to show what words are commonly used in our documents
- One option is to provide a definition 'for our purposes', but McRoberts says we should not perpetuate the erroneous use of words, so better to choose definitions from literature
- Use a hyperlink on words in the documents which links to a glossary
- Some definitions can be graphical (i.e. can use figures in addition to text).
- Words to include:
 - o Assessment unit
 - o Estimate
 - o Gross emissions
 - o Net emissions
 - o Monitoring (we mean monitoring change usually)
 - o Trend
 - o Sampling unit
 - o Population unit - this is already in Annex B of the MGD
 - o Design based inference
 - o Model based inference
 - o Probability sample
 - o Unbiased estimate
 - Consider changing the name from "unbiased estimate" to "estimate obtained using an unbiased estimator, which has a confidence interval" (comment from McRoberts):
 - "Bias is a property of an estimator (a procedure or a formula), not an estimate, a map, or a sample. Bias is assessed by comparing the true value to the mean of the estimates over all possible samples. Thus, bias cannot be assessed if the true value is not known or if only a single sample is available. The best that can be asserted is that an estimate was obtained using an unbiased estimator, but even then the estimate obtained with any particular sample may deviate substantially from the true value. This is why confidence intervals are important."

- There is not a correct understanding on this, so people think that it is the true estimate (although it contains the word 'estimate', which should inform people it is as such) – this is a problem for users (not for the statistics community).
 - It depends on the context, the IPCC use the phrasing “you must neither over, nor overestimate the ..”, and the IPCC use this because it is precise and simple.
 - Degradation
 - On the one hand best not to define it, but to just measure some forest properties which you can define, and report on what is happening
 - On the other hand, we can't estimate it unless we can define it
 - There is a lot of confusion on this term, and some countries are labelling any reduction in biomass as degradation, which is not always the case. What is and is not degradation has implications for policy making.
 - One question which needs to be answered for each country is: is a temporarily unstocked forest an example of degradation?
 - Deforestation

Timing: To be agreed.

Overview module (FAQ structure)

This FAQ (frequently asked questions) structure will make a valuable contribution to the existing MGD Rapid Response Module format.

This will give a summary of country experiences and critical issues related to stratified area estimation

The purpose is to highlight the issues considered when generating activity data (this document should be written from the country perspective and linked to MGD guidance)

Assumption: Design should be for a national forest monitoring system.

List of terms – Pontus, Carly

1. Overview/Introduction - Carly
 - a. What products do you have?
 - b. What is useful for generating activity data for the purpose of generating estimates? – link to MGD guidance (are their gaps?)
 - i. How maps are used? – examples and consequences (good and bad)
Relative efficiency of using maps for estimates
 - c. Generation of activity data with confidence intervals (should I use my map or not?)

2. Sampling design – can be based on Steve's presentation, example and last workshop report:

- a. What are the objectives? – Steve, Erik L, Andres, Maria
 - i. National and nested jurisdictional design – examples DRC (relationship to GHGI)
- b. Assessment unit – Steve, Andres, Pontus
- c. Type of inference – Pontus

NOTE: there is only one type of inference in the examples – remind them there is another type that can be used – are there any consequences – examples from other land uses. Should aim for design based but there are other alternatives.

- d. Use of auxiliary information:
 - i. Number and type of strata (use of buffers, pre and post stratification) – Erik L, Pontus
 - ii. Country experience showed that MGD allocation of samples works well – Inge, Andres, Steve
- e. Sample size (check formula in MGD and its purpose – total sample size) –
 - i. Focus on presenting examples - Erik L, Steve, Andres
- f. Repeated/long term monitoring considerations (Guyana example) – Martin, Ron, Erik L to provide examples

3. Response design:
 - a. Assessment unit: provide a list of issues to consider and point at examples – point / polygon / pixel / contextual
 - i. Forest definition MMU vs. reference sample interpretation unit - Andres, Christophe, Erik L (explain Collect Earth), Danny Donaghue, Erik N
Should contextual information be used or not? With the examples explore the variation in the methods and any impacts. Link to discussion on land use vs land cover.
 - b. Source for reference data (greater quality!) (refer to Ron module, Erik N, Christophe, Erik L)
 - c. Labelling protocol (transparent/consistent; land use vs land cover, cover degradation challenges; examples) – Maria, Erik L, Andres, Inge
 - i. Is the classification system fit for purpose/the objective (include degradation examples)
 - ii. Quality of reference interpretation, create a “manual of examples” for representative reference data interpretation,
 - iii. QC, assign confidence to interpretation – guidance on how to identify and check critical points (refer to Ron Module)
4. Analysis (specific on results, tied closely to design) – Maria, Christophe (MAR),
 - a. Check MGD content - Carly
 - b. Draw on country experiences to identify common mistakes – Inge, Martin
 - c. Insights into random and systematic uncertainty, its relevance (drawing on examples) and use of the information for continuous improvement – Martin, Inge
5. Transparency / completeness (structure focus) – Andres, Maria, Inge
 - a. Examples of where recalculation was achieved
 - b. Clear communication on decisions on distinguishing degradation

Timing:

- August – Material provided
- September – collation into first draft
- October finalise

Technical module: Statistical inference

- This module is very specific, and is on the problem of having a high omission error, which occurs when a small class has one incorrect .
- It must give advice to countries which are experiencing a particular problem, and we provide some potential solutions, based on what has been done in the past.
- We first need to give a perspective, why this is happening – it is not a mistake / error in the analysis. They are not doing anything wrong – this has to be clarified.
- We need to reassure practitioners, in that this is not a serious problem, but that it can be dealt with.
- Examples, can be real or fictitious (maybe USA example).
- This issue is not new or unique... in USA, forest non-forest classifications. In non-forest they don't sample intensively. If you do a buffer around the edge of strata (because this is where the misclassifications are likely to be), and sample in the buffer at the same intensity as in the forest class, this can solve the issue (Reference: McRoberts, R.E., Wendt, D.G., Nelson, M.D., Hansen, M.H., 2002. Using a land cover classification based on satellite imagery to improve the precision of forest inventory area estimates. Remote Sensing of Environment 81, 36–44.). This should capture missed deforestation, and reduce the CI.
- More sample units (or increasing the sample size) is one way of reducing the CI.
- The main way to deal with it is to buffer, so how big should the buffer be etc. should be clarified.
- There can be scripts which are developed and put into SEPAL when they are ready.
- Currently in SEPAL, a standard report is generated for stratified area estimation containing all relevant choices (transparency).

Timing: To be agreed.

Technical module: Reference data

The plan for the "Reference data error" topic is as follows:

- (1) Steve Stehman (in collaboration with R. McRoberts and E. Naesset) will complete and publish a paper that partially address the issue of estimating the effects of reference data error.
- (2) As second study (in collaboration with S. Stehman) will be conducted using Monte Carlo techniques to estimate the effects for a variety of conditions including the accuracy of the underlying classification used as the basis for stratification, possibly the number of interpreters and their probability of agreement, and the proportional size of the land use class of interest. This study will also assess the effects of reference data error using the multi-resolution data that Erik Næsset presented at the Oslo meeting. This will include data from Tanzania provided by Norway. A paper based on this study will also be submitted to a journal.
- (3) The MGD module will include the following: (i) a more detailed statement of the criteria for reference data, (ii) a generic description and discussion of the consequences of reference data error, (iii) a brief review of any literature we find on the topic and of the two studies mentioned previously, and (iv) to the degree possible, recommendations for avoiding and/or circumventing the effects of reference data error (see draft outline below).

On the literature review, there is considerable literature on the topic of reference data error, but very little on the topic of multi-interpreter disagreement. A team of interpreters is currently producing a multi-interpreter dataset to facilitate a real world analysis using the Hansen/U of Maryland dataset.

Outline:

Context

- Role of reference data
- Criteria for reference data
- Sources of reference data
- Errors in reference data (examples)

Consequences of reference data errors (R&D study)

- Estimating the effects of reference data errors (hybrid inference)
- Activity data estimation method: stratified vs model-assisted
- Forest type: dry forest (Tanzania) vs humid forest (Gabon)

Dealing with the effects of errors

- Conditions characterizing negligible effects
- Number of interpreters
- Resolution

Improving reference data

- We need to put an emphasis on training interpreters
- This requires an investment – in time etc.
- These problems are common for anyone who does anything to do with remote sensing – this is a general problem – we cannot ignore it
 - o Already in MGD is some guidance, and we can build on this,
- There needs to be clear guidelines which can minimize the uncertainty in the reference data.
- Based on our existing knowledge, put together examples of how to do visual interpretation (i.e. what is 'better data', and which spectral bands in Landsat are most effective, how to look at time series, and what to look for in images)
- What guidance can we provide about training interpreters?
- **Action point:** Check with CNES about the availability of historical SPOT data
- Response design guidance should cover: assessment unit and labelling (could use Stehman 2011, RSE as a base) – better guidance on collecting reference, – provide a list of issues to consider and point at examples, labelling protocol should be provided, Forest definition MMU vs. reference sample interpretation unit, reference data sources (higher quality!), interpretation vs. quality of reference interpretation, create a "manual of examples" for representative reference data interpretation, assign confidence to interpretation – guidance on how to check critical points, QC.

About reference data:

- Assumptions – error free and fixed quantities (there should be something in the MGD on this)

- Confidence of observers – Erik N's example: What does this tell us about the uncertainty of the estimate?
- What is happening on the ground – there are some countries which are very centralized, and some (Indonesia) it is more regional in their approach to reference data?
- The bias issue is the largest concern with this rather than the inconsistencies.
- What will this look like for change – where there are different resolutions of data etc...
- This is very dependent on the type of forest, for example dense moist forest would be much clearer

Timing: Work undertaken in Fall / Winter 2017, or as time permits.

Next meetings / other

Links/considerations for capacity building (CB):

- GFOI CB summit:
 - Presentation of these efforts, results, inform others, solicit feedback (Pontus)
 - Perhaps organize a breakout group – in particular to add examples
 - Get feedback in what formats the CB component is expecting input from MGD/R&D
- Planning targeted capacity development (specific regional workshops with few countries working on it?), dissemination to community, link to specific tools (i.e. REDD Compass, CEO, SEPAL)
- Once completed – organize Webinar (also invite TAP (Andres) and UNFCCC roster of experts (Maria))
- FCPF: maybe training for the TAP

Other meetings:

- Globbiomass meeting in September (11-13th)
- Bern ICI –monitoring community – in November

Next workshop on uncertainties of EF and biomass maps

- Topics:
 - Several components requiring attention (i.e. biomass, allometry, BCEF, root/shoot)
 - MGD: some sections have been drafted (on EF uncertainty)
 - Cover stock change, and losses and gains (removal factors)
 - Draw from previous meetings (Oslo 2017, De Hague 2016)
 - Propagation of uncertainty of EFs (involves several elements), for emission (reduction) estimation, and trends
 - Uncertainties of the IPCC default factors
 - Uncertainty of biomass maps
 - In particular for countries that don't have a designed-based/probabilistic sample NFI and for countries that don't have any data
 - Approaches for model-based estimation of uncertainty, Monte-Carlo examples (i.e. Mexico)
- Date: Dec./Jan. 2017/18
- Location: Bilbao?, Maria offer to host and to help with local logistics
- Funding: GFOI office, Silvacarbon, ask NICFI/NSC, WB-FCPF
- Duration: 2.5-3 days
- Key issues:
 - What are factors determining the uncertainty of EFs?
 - How can current biomass maps be used to increase precision in national estimation?
 - How to construct a new biomass map to improve IPCC-compliant estimation?

3. Agenda

Day 1: Monday 26 June

10.30	Opening	Henrik Fliflet
10.45	Background and objectives	<u>Martin Herold</u>
11.00	Summary of first workshop on this topic: What has been agreed? What requires further discussion at this workshop?	<u>Andres Espejo</u>
12.30	Lunch break	
13.30	Presentation of case studies and experiences: <ul style="list-style-type: none">• What are good practices for countries to assess the uncertainty of accuracy for area estimates /activity data?• What are the lessons learned from your country examples?• How methodological advancements and new data can improve estimation (i.e. from Sentinels) and how should that effect the AA procedures?	<u>Andres Espejo</u> , Kay Kallweit, Evan Notmann, <u>Christoph Sannier</u>

Day 2: Tuesday 27 June

09:00	Further discussions on case studies and experiences: <ul style="list-style-type: none">• What type of guidance can be provided with respect to operational methods/examples?• What issues remain open issue and require further discussions to develop solutions?• What issues cannot be solved because we lack basic R&D?	<u>Erik Naasset</u> , <u>Ronald McRoberts</u> , <u>Christoph Sannier</u>
12.30	Lunch break	
13:30	Further discussions on case studies and experiences continued	<u>Pontus Olofsson</u>
14:00	New guidance plans Working groups on synthesizing and developing new guidance: <ul style="list-style-type: none">• What new points need to be included in the guidance?• Is any current guidance obsolete?	<u>Carly Green</u> Discussion facilitated by Maria Sanchez
19.00	Workshop dinner	

Day 3: Wednesday 28 June

08:30	SEPAL demonstration	<u>Erik Lindquist</u>
09:00	Working group discussions and presentations on scoping of new MGD module: <ul style="list-style-type: none">• What guidance is needed for the MGD?• How can we formulate this?• What new modules are required?	<u>Ronald McRoberts on sampling design</u> Discussion facilitated by Carly Green
12.30	Lunch break	

13:30	Next steps:	Discussion facilitated by Martin Herold
	<ul style="list-style-type: none"> • Addressing additional issues which need further (i.e. uncertainties of trends?) • Next expert workshops on uncertainties related to EFs and biomass estimates 	
16.30	End meeting	

4. Participants

Name	Entity	Email
Andres Espejo	FCPF	aespejo@worldbank.org
Carly Green	GFOI MGD	carly.green@enviroaccounts.com
Christophe Sannier	SIRS	christophe.sannier@sirs-fr.com
Erik Lindquist	FAO	Erik.Lindquist@fao.org
Erik Naeset	UMB	erik.naeset@nmbu.no
Evan Notman	USAID/Silvacarbon/GFOI	enotman@usaid.gov
Evie Hagen	NSC	evie.merethe.hagen@spacecentre.no
Henrik Fliflet	NICFI	henrik.fliflet@md.dep.no
Inge Jonckheere	FAO	Inge.Jonckheere@fao.org
Kay Kallweit	GiZ	kay.kallweit@giz.de
Maria Sanz Sanchez	GFOI MGD	mj.sanz@bc3research.org
Martin Herold	GFOI R&D/WUR	martin.herold@wur.nl
Pontus Olofsson	Boston university	olofsson@bu.edu
Ronald McRoberts	USFS	rmcroberts@fs.fed.us
Sarah Carter	GFOI R&D / GOFC-GOLD	sarah.carter@wur.nl
Stephen Stehman	SUNY	svstehma@syr.edu
Tom Harvey	GFOI	Thomas.harvey@fao.org