FCPF Carbon Fund Methodological Framework Discussion Paper #3:

Measurement, Reporting and Verification Options for the FCPF Carbon Fund

WORKING DRAFT SOLELY FOR INPUT INTO AND DISCUSSION BY CARBON FUND WORKING GROUP

Draft 1-23-13

by TAP Experts[1], managed and edited by Facility Management Team

Table of Contents

I. Key Questions for the Carbon Fund's Eventual Guidance on This Topic

II. Introduction

III. Measurement and Reporting:
- A. Key Considerations in designing an MR system
- B. Scope of Measurement and Reporting: Options for the Carbon Fund
- C. Scale of Measurement and Reporting: Options for the Carbon Fund

IV. Uncertainty in Measurement, Reporting and Verification Systems
- A. Sources of Error and Uncertainty and How to Reduce Uncertainty
- B. Quantifying Uncertainty of Individual Components
- C. Combining Errors from Individual Components
- D. Managing the Remaining Uncertainty

V. Verification
- A. Key Considerations in Designing a Verification System
- B. Verification Options for the Carbon Fund

VI. Recommendations
- A. Measurement and Reporting scope comparison
- B. Measurement and Reporting scale comparison
- C. Comparison for Verification Options
- D. Other Recommendations for MRV System Design

VII. Potential Next Steps : Topics on Which Further Analysis or Discussion is Needed

[1] Additional technical support provided by Forest Carbon, Markets and Communities (FCMC), a US Agency for International Development funded program.

I. Key Questions for the Carbon Fund's Eventual Guidance on This Topic

1. What is the simplest measurement approach and appropriate level of guidance that could provide key CF requirements and be reasonably cost effective and implementable by REDD+ countries, given their varying capacities, in the short timeframe of a CF ERPA (say 2015-2020)?
2. Is it possible for ER-Program MRV systems to distinguish between anthropogenic and non-anthropogenic sources of emissions?
3. What is the most promising approach to addressing the complex relationship between the national and subnational ER Program scales?
4. As a piloting step, could the CF produce a guidance requiring detailed assessment and transparent reporting of the uncertainty of each major part of the MRV system (e.g., land use activity data), but recognize that the overall uncertainty of an Emission Reductions Program might be significantly higher than the MRV system's ability to detect mitigation actions in comparison to the reference level?
5. Any there existing climate initiative guidance or requirements that would meet the needs of the CF regarding MRV guidance to ER Program REDD country proposers and FCPF review of ER proposals?
6. Does the CF need to provide guidance that requires MRV systems of ER Programs—and ER generated-- to be strictly comparable, or is more flexibility in the pilot stage reasonable?

II. Introduction

Background

Performance-based attribution accounting requires measurement, reporting and verification (MRV) of ER generated by program activities relative to an established business as usual reference emission level (REL)[2]. Credible ER require the quantification of uncertainty and an expert review of reporting results and methods. A six-year history of Clean Development Mechanism (CDM) under Kyoto, the advancement of REDD+ under the UN-REDD Programme, the emergening agreements[3] and guidance under the UNFCCC, and the availability of multiple voluntary standards all provide context and lessons to help guide the methodological framework for of the FCPF Carbon Fund.

[2] Decision 1/CP.16 requires countries participating in REDD+ activities to establish "A robust and transparent national forest monitoring system for the monitoring and reporting" of such activities (paragraph 71 (c)), and establishes that results-based actions "should be fully measured, reported and verified (paragraph 73)."

[3] For instance, decision 4/CP.15 providing methodological guidance for REDD+ activities, decision 1/CP.16 defining REDD+ activities and basic participation requirements, inter alia, and decision 12/CP.17 on modalities relating to forest reference emission levels and forest reference levels.

Objectives

The aim of this issue paper is to provide key issue on MRV and options for consideration under the Carbon Fund Methodological Framework. The FCPF Participants Committee has provided general guidance (principles) stating that methods for the Carbon require i) consistency with UNFCCC emerging guidelines; ii) payment for performance; iii) reduction of delivery risk; iv) consistency with REL methods; and v) synergy with the greater national REDD+ readiness preparations. Based on this, this paper provides a review of current practices and good guidance under national and subnational schemes. It assumes that a reference emissions level baseline is available, and that ER Program measurement and quality control and quality assurance (QA/QC) parameters and requirements for verification will be formally defined, either within the developing FCPF methodological framework or within ER Program specific methodological documents.

Monitoring of non-carbon values (e.g., biodiversity and social aspects) are important aspects of an ER-Proram but are treated as parts of different papers. Likewise, the establishment of national arrangements is beyond the scope of this paper.

III. Measurement and Reporting

This section presents key issues and options related to scope and scale of measurement and reporting. The next section addresses uncertainty and accuracy. For brevity and focus other considerations listed in the previous section are not treated in more detail.

Measurement and reporting are discussed separately from verification. Measurement and reporting, under REDD+ programs, principally combines estimates of areal changes (activity data or AD) and estimates corresponding carbon density (emission factors or EF) in the Agriculture, Forestry and other Land Use (AFOLU) sector to quantify CO_2e emissions and removals. Verification is likely to be carried out by an independent and objective body (agreements regarding this are still under discussion).

A. Key Considerations

COP guidance

Specific modalities were discussed during COP 18 but SBSTA could not reach agreement. Prior agreements provide the general shape of a UNFCCC MRV framework:

- Use a combination of remote sensing and ground-based forest carbon inventory approaches for estimating, as appropriate, anthropogenic forest-related GHG emissions by sources and removals by sinks, forest carbon stocks and forest area changes
- Provide estimates that are transparent, consistent, as far as possible accurate, and that reduce uncertainties, taking into account national capabilities and capacities
- Ensure measurement and reporting is transparent and results are available and suitable for review, as agreed by the COP

Relationship to RL/REL

While there may be an implicit link between REL/RLs and measurement and reporting, this is not yet explicit in the COP decisions. The Cancun Agreements states that national (and if appropriate subnational) RL/RELs should be developed for the REDD+ activities a country chooses to undertake, as should national (and if appropriate subnational) measurement and reporting systems.[4] SBSTA 37 at Doha could not reach an agreement on further guidance on MRV, though un-bracketed text in the chair's summary made a clearer link between RL/RELs and measuring and reporting.[5]

The methods and data used to generate the RL/REL should be consistent or at least compatible with the measurement and reporting methods. They are unlikely to be identical as RL/RELs will involve historic data sets whereas measurement technology and methods are continuing to improve.

Scope of Measurement and Reporting

REDD+ activities. REDD+ under UNFCCC encompasses reducing emission from deforestation, forest degradation and enhancement of forest carbon stocks. Deforestation is the conversion from a forest to a non-forest land cover class and standard methods and good practice examples exist. However, degradation and enhancement present significantly greater challenges. Degradation has not been formally defined, but guidance includes degradation from timber harvesting, fuelwood collection, human-induced fires and over grazing. Degradation emission factors can be obtained through quantification of carbon stock changes from forest transitions (i.e. from one defined strata to another, or from emission factor updates through biomass sampling).

Carbon Pools. COP 17 states that significant carbon pools should not be excluded from setting forest reference and reference emission levels (RL/REL). To the extent that RL/RELs are linked to measurement and reporting the pools included in a RL/REL will also need to be monitored.

Scale of Measurement and Reporting

Scale of measurement and reporting relates to both the geographic extent and the spatial resolution of the smallest area of land captured by measurement (or minimum mapping unit). The geographic extent is dictated by the spatial distribution of ER Program activities and the potential requirement of reference and leakage areas outside of the ER Program boundaries. Under the COP 16 Cancun decision, participating nations must develop a national scale monitoring system, though a subnational system can be developed as an interim measure. A national system may not generate sufficient information to verify ERs at the scale of the Program, in which case additional or more detailed measurements for the ER Program may be required.

---

[4] Decision 1/CP.16, paragraph 71 (b),(c)

[5] FCCC/SBSTA/2012/L.31, Methodological guidance for activities relating to reducing emissions from deforestation and forest degradation and the role of conservation, sustainable management of forests and enhancement of forest carbon stocks in developing countries, Draft conclusions proposed by the Chair, Annex, paragraphs 9 – 12.

Frequency of Measurement and Reporting

The frequency of measurement and reporting should be set such that it balances the ability to detect threats to ER Program success, cost of **measurement**, technical limitations and reporting requirements. Due to the sensitivity of remote sensing to seasonal environmental conditions such as phenology and cloud cover, qualitative rapid assessment through remote sensing can be ongoing and far more frequent, limited only by the available and costs of imagery. If data availability makes matching seasons too difficult, adaptive methods to account for region specific seasonality challenges should be applied.
Parties agreed at COP 17 in Durban that financing for emission reductions should be results-based. Therefore, reporting frequency, at a minimum, should be done at intervals set by ER Program terms for performance-based payment for verified ERs. Mandatory frequency of GHG inventory reporting for non-Annex 1 countries is not currently set, but biennial reporting[6] is required commencing December 2014.

Uncertainty, Accuracy and Error associated with Measurement and Reporting

There is a large potential uncertainty associated with wide area assessments of land cover, both in the establishment of the spatial distribution of carbon pools and in the accurate recording of land cover change. The estimates of areal change (activity data) and emission factors rely on data inputs from multiple sources, each with their own level of uncertainty. Error arising from mistakes is controlled through the establishment of data management protocols. These protocols are codified in standard operation procedures which embed quality control and assurance measures to reduce mistakes and to track uncertainty. Methods to track and quantify uncertainty from both inputs and to report accuracy of results are required at all stages of M & MRV and are specific to what is measured (see section IV).

Capacity

Lack of in-country technical capacity in remote sensing, GIS, forestry and quantitative modeling are frequently cited as an obstacle towards developing MRV systems. An ER program must identify knowledge gaps prior to initiating the program and ensure that the required technical staff are available and capable of fulfilling all reporting requirements. UNFCCC guidelines allow for flexibility of standards for national MRV efforts dependent on national capacity, however standards are still required to be sufficient to bound uncertainty relative to ER measurements.

Data Availability and Quality

For activity data, an important consideration is the availability and quality of imagery for analysis. Persistent cloud cover limits the utility of multispectral imagery. Cloud penetrating radar technology is becoming available but requires more technical capacity and promising new approaches and sensors are largely at the stage of research and development at this time (as opposed to optical methods that are well established). Also satellite systems can experience technical failures can and do occur and render a particular data source unavailable. Likewise, quality emission factor data for various forest-type strata

[6] Per UNFCCC Decision 4/CP.15, paragraph 1(d)

are also often not available to the degree necessary for the variance of each carbon pool to meet high quality standards.

B. Scope: Major Considerations for the Carbon Fund (Which Activities and Pools?)

1. MR Scope to Follow Scope of RL/REL

The FCPF mandate, to be consistent under the UNFCCC, requires that ER programs are patterned after UNFCCC decisions regarding the scope, scale and frequency of MRV activities. As noted above, while there may be an implicit link between REL/RLs and the scope of measurement and reporting, this is not yet explicit in the COP decisions. Under this option, the scope of the measurement and reporting would follow the activities and pools included in the RL/REL, plus any additional activities or pools that need to be tracked and accounted for as part of the ER Program's leakage requirements. For example, if the RL/REL of an ER Program only covered deforestation and above ground biomass, and did not extend to degradation or enhancements (or include all carbon pools), then at a minimum the scope of measurement and reporting must also include deforestation and above ground biomass along with any additional measurement required for leakage accounting purposes.

2. **Follow IPCC Guidance**

In the IPCC Good Practice Guidance (IPCC 2003) the most common simple methodological approach is to combine activity data (AD - information on the extent to which a human activity takes place) with emission factors (EF - coefficients which quantify the emissions or removals per unit of activity):

$$\text{Emissions} = \text{AD} \times \text{EF}.$$

Regarding the AD in general, the IPCC indicates that countries should accurately and completely represent and report all land areas in a country where human activities take place (land-use classes). The IPCC Guidance requires that land be separated into 5 minimum classes (forest, grasslands, cropland, wetlands, settlements and other lands).

<u>Land use classes</u>

It is possible for the measurement and reporting system to cover IPCC land use classes. Increasing the number of monitored strata beyond the minimum six IPCC defined land use classes reduces the variability of EF within monitored classes, thus allowing for higher levels of certainty in reported results. Additionally, increased strata allow for the inclusion of IFM and ecosystem restoration activities in which the carbon state of stratum increases through direct management or avoided degradation. General remote sensing techniques that have proven relatively successful for measuring a general forest class would need to be enhanced with high-resolution imagery and/or the inclusion of ancillary data. Generally, the higher resolution a dataset is the better it is at identifying land cover, but typically too expensive for the comprehensive evaluation of large areas. In order to minimize costs measurement activities can prioritize areas at higher risk, using various criteria to determine which areas are at higher risk.[7] Areas identified as high risk may be subject to additional remote sensing or ground based measurement.

Tiers

To estimate emission factors IPCC accepts the use of three different "Tiers". Tier 1 and Tier 2 methods are not spatially explicit and require only the use of IPCC methods and published data, default IPCC values for Tier 1 and country specific data for Tier 2. Tier 3 involves the greatest level of analysis and requires spatially explicit monitoring and measurement. EF for Tier 3 key classes are to be derived from direct field measurements from site specific research, ER program data collection activities or from a NFI. Tier 3 pools are identified via the key category prioritization. As such, the scope and depth of analysis of a MRV system must balance resource capacity with the needs of producing verifiable results-based ERs. Remote sensing, biomass plots and GIS are the primary tools of Tier 3 measurement. To efficiently allocate limited resources, the level of detail in assessing the emission factors of the various carbon pools can be prioritized. The prioritization is accomplished by defining key categories. A key category is a GHG source or sink for which variation due to uncertainty would have a substantial influence on national GHG reporting net results.

IPCC Key Category Tiers

Tier	Details
Tier 1	IPCC methods and IPCC default values (no data collection needed)
Tier 2	IPCC methods and country specific data for key factors (including more detailed country specific strata)
Tier 3	Country specific methods or models, national inventory of key carbon stocks, repeated measurements of permanent plots to directly measure changes in forest biomass

[7] These could include, for example, distance to roads, rivers, human settlements, other disturbances etc.

Approaches

To estimate emission factors IPCC accepts the use of three different "Approaches". Approach 1 uses basic land-use data; Approach 2 uses a survey of land use and land-use change; and Approach 3 uses geographically explicit land use data.

Approach 3 would require a Satellite Land Monitoring System or extremely intense field sampling to monitor AD carbon stock and changes of carbon pools in the six broad IPCC land use categories (forest land, cropland, grazing land, wetlands, settlements, and other land). While UNFCCC standards will possibly include degradation and enhancement, currently developed national monitoring systems are only capable of measuring some land-use change, deforestation, afforestation and reforestation activities. Suitable platforms for national monitoring include medium resolution multispectral and radar satellite imagery.

Summary of remote sensing data types

Type	Description	Advantages	Disadvantages
Coarse resolution multispectral	Multiple bandwidth datasets with a pixel resolution of 100 m or greater.	• Rapid revisit rate (often daily) • inexpensive or free • often have a wide spectral range and long archival history, • generally has global coverage.	• Pixel resolution too course to reliably identify small parcel LULUC • subject to atmospheric disturbances and clouds.
Medium resolution multispectral	Multiple bandwidth datasets with a pixel resolution greater than 15 m and less than or equal to 30 m.	• Generally inexpensive or free, • often have a wide spectral range, • sufficient resolution to measure most deforestation, • long archival history.	• Pixel resolution too course to reliably identify degradation, • subject to atmospheric disturbances and clouds, • significant gaps in global historic coverage.
High resolution multispectral	Multiple bandwidth datasets with a pixel resolution less than 15 m.	• Allows visual confirmation of many forms of degradation, • allows identification of a wider range of strata.	• Very expensive, • limited coverage • limitedrepeat visit intervals, • subject to atmospheric disturbances and clouds, • limited historic

			archive.
Aerial photography	Typical black and white or color photography, sometime including near infrared. Resolution typically less than 3 m.	• Allows visual confirmation of many forms of degradation, • allows identification of a wider range of strata.	• Very expensive to collect, • limited band range, • limited historic archive.
Radar (SAR)	Active sensor recording scattered energy returned from different bandwidths.	• Many sources are inexpensive, • detects structural information of land cover, penetrate clouds and atmospheric disturbances.	• Technically complex to process, • visual interpretation not intuitive, • prone to errors in topographically complex environments.
LiDAR	Active sensor recording returns from rapid laser pulses.	• Best source of structural data.	• Technically complex to process, • very expensive to collect.

C. Scale: Options for the Carbon Fund

The UNFCCC allows for subnational RL/RELs and monitoring systems as an interim step towards establishing a national monitoring system, noting that subnational systems should include measurement and reporting on displacement that may occur at the national level and a means to be integrated into a national system.[8] Economies of scale are frequently cited as incentives for establish national scale monitoring systems. However, project based REDD experience has shown that monitoring requirements vary greatly depending on eco-region composition and complexity, topology, seasonality and phenology. Reliable measurement at finer scales requires detailed local knowledge and the development of area-specific techniques involving ancillary data inputs. For example, the ER PIN submitted by Costa Rica in September 2012 involves activities occurring within forest parcels, mostly under 50 ha. For a performance-based system to properly measure and report on these parcels it must be of sufficient resolution and accuracy to identify emissions and removals above the noise of uncertainty. COP decisions have a built in sensitivity to national conditions, which takes into account the capacity and abilities of participating non-Annex 1 countries. As a result, there will be no unifying standard under which all national monitoring systems must adhere. Uncertainty of submitted GHG emission reports will vary and can easily exceed the magnitude of ER gains and losses under an ER program.

1. National level Measurement, Reporting

The current stage of development, accuracy and resolution of national MRV systems varies widely among REDD countries. The spatial resolution of a national system may be insufficient to properly

[8] Decision 1/CP.16, paragraph 71 (c)

quantify emissions and removals of small parcel-based activities. ER Program activities need to be evaluated against the scale, resolution and accuracy of the national monitoring system to determine if uncertainty exceeds the ER Program activities performance. To avoid the duplication of effort, ER Program requirements could be set to monitor the national scale scope of carbon pools only, and program resources could assist in the development of the national MR system.

Uncertainty can be addressed using discount factors (see next section on Error and Uncertainty). Due to the inherent variability in ER Program types and the resolution required to measure them, decisions regarding scale could utilize a decision tree to aid in assessing the suitability of a national scale system. Spatial monitoring and measurement systems allow for the generalization of smaller geographic areas into a larger extent. If national scale measurement is not currently feasible, sub-national scale efforts can later be integrated into the larger system as long as the scope of measurement components is consistent.

If the national monitoring system never allows for the spatial resolution required by ER Program activity measurement requirements, ER Program monitoring and measurement must continue as a separate but connected activity. While course resolution monitoring and measurement cannot be downscaled, finer resolution measurement can be generalized for inclusion in a national effort outside of the ER Program specific activities.

2. Emission Reduction Program Level Measurement and Reporting

Where the ER Program is smaller than the entire country, requirements for an ER Program specific monitoring system can be set at a higher standard than national monitoring systems. A program-specific structure can be built to ensure that ERs generated are performance-based and are measured in a credible and transparent manner. Advantages of a jurisdictional scale monitoring system are the smaller area may allow for the inclusion of higher resolution imagery, greater density of EF plots and ground truth and the ability to tune measurement methods to a smaller regional range of conditions. All of these advantages lead to greater certainty and less bias in measurement results, allowing for the inclusion of strict standards that enables highly credible ER's. Furthermore, ER Program scale measurement can allow for attribution via location analysis for transparent benefit sharing mechanisms and evaluation of individual program activity performance.

The disadvantage is that the burden of the system would be exclusive to the ER Program. The cost of developing an ER Program-scale system would be tied to the financial viability of the ER Program itself. Another disadvantage could be the lack of capacity and expertise at more local levels. Jurisdictional participants should work under the context of the national measurement effort, both for technology transfer and to ensure that the subnational MRV systems can later be linked to the national scale.

3. Combined Measurement and Reporting

If some, but not all, components in a national system are of sufficient certainty to monitor ER Program activities, an ER Program scale system can be employed only to the extent required to establish credible ERs under ER Program activities, for which the national scale MR system cannot reliably monitor. For example, a national system may have a robust database of cover-type EF and suitable accuracy in non-forest LULC classes. These AD and EF factors could be used to fulfill those ER Program measurement requirements. This would leave only the detailed measurement requirement of forest stratum or fine scale activates to a dedicated ER Program monitoring and measurement system.

IV. Uncertainty in Measurement, Reporting and Verification Systems

UNFCCC Key Category Analysis (KCA) states that IPCC default values and unqualified country-specific data on carbon pool changes (Tier 1 and Tier 2) are reserved for non-key categories of overall limited influence on national GHG emissions. For Key Categories, country-specific Tier 2 data inputs should only be used if an uncertainty analysis is included in the data documentation or if the data provides sufficient transparency for an assessment of uncertainty. All new program-specific data collection activities, e.g., Tier 3, need to be designed such that QA/QC minimizes mechanical error. Sampling frequency must also be sufficient for uncertainty analysis, as UNFCCC default requires a 95 percent confidence interval. Errors associated with each multiple input needs to be propagated.

The accuracy of a measurement is the proximity of measurements to the actual (true) value. The precision of a measurement, also called “reproducibility” or “repeatability,” is the degree of closeness of the measurements to one another. If measurements are highly precise but not very accurate, they are biased. If a measurement is unbiased, taking more measurements will lead to a reduction in uncertainty. This is not the case for biased measurements.

A. Sources of Error and Uncertainty and How to Reduce Uncertainty

The total uncertainty around emission or emission reduction data is dependent on the uncertainty of individual measurements and components that were used to calculate the final value of emissions and emission reductions. Because emission or emission reductions are calculated by multiplying activity data with emission levels, it is essential to investigate the error associated with these components. Before mechanisms to manage uncertainty are discussed, this section outlines the individual components and the sources of uncertainty.

Stratification

One important mechanism to manage variability and uncertainty is stratification. In the context of REDD+, stratification is the division of land into areas that are assumed to be more or less homogeneous. Within each stratum, one value of AD— e.g., deforestation rates and emission levels primarily impacted by biomass stock densities—is calculated. If no stratification were to be employed, the uncertainty associated with either biomass stock densities or deforestation rates of the entire land can be expected to be relatively high, due to the heterogeneity of most forest systems. Tropical forest systems are especially heterogeneous. Stratification increases accuracy of calculations for the same sampling intensity compared to a non-stratified approach, or reduces the cost to achieve a set accuracy. Calculations, sampling and measurements are usually done separately for each stratum

Activity Data

Key Category AD is monitored through remote sensing and LULC mapping. LULC classification at the provincial or national scale requires the use of remote sensing imagery and ancillary data and is interpreted using computer algorithms. Areas of uncertainty arise in spatial alignment of input layers, integrity of the input layers (i.e., sensor noise or data gaps), errors and uncertainty in the training dataset used to calibrate the computer algorithm, spurious seasonal changes in phenology and overall algorithm classification performance.

- **Spatial alignment of input layers**
- **Integrity of input layers**
- **Errors in reference data**
- **Algorithm classification performance**

Emission Levels

Emission level coefficients are created through statistical sampling through field and laboratory measurements. Mechanical error associated with recording and transcription should be minimized through the implementation of sufficient QA/QC. Sampling uncertainty is bounded by variability within each of the data parameters measured. Main sources of error are the use of empirical equations such as biomass expansion factors and allometric equations, inherent variability of biomass and carbon stock density.

- **Allometric equations and biomass expansion factors** masses.
- **Inherent variability of biomass and carbon stock density**

B. Quantifying Uncertainty of Individual Components

The total uncertainty around emission or emission reduction data is dependent on the uncertainty of individual measurements and components that were used to get to the final value of emissions and emission reductions. First, the quantification of the errors of these individual components is discussed before we detail how the uncertainty from individual components can be combined to yield the uncertainty of the emission (reductions).

Error Matrix to Quantify Errors in Activity Data

Uncertainty in LULC classification is quantified through the use of an error matrix which matches ground-truthed cover identification from various sources (i.e., field visit, aerial survey) to the cover values predicted by the classification algorithm. The ground-truthed reference data represents the actual or true values. As a consequence, the comparison of LULC classes gathered from field visits with LULC classes predicted during classification is an appropriate way to quantify accuracy. The error matrix, also known as a confusion matrix, provides statistics for overall accuracy, which can be further refined to errors of commission and errors of omission. It is reported in a simple percentage of correct algorithm predictions for each of the ground truth points. Uncertainty can further be quantified by providing the Kappa statistic which, based on the number of classes being predicted, takes into account the chance of randomly achieving a correct classification. It is essential that the reference data against which the LULC map is evaluated is truly independent and not used during training of the classification algorithm.

Variance Statistics to Quantify Errors in Emission Factors

While the actual values for land use classes and land use transitions are relatively easy to observe in the field and a measure of true accuracy can be obtained, actual or true biomass stock densities are more difficult to obtain. Often, only the standard deviation or variance of sampling plots within each stratum is used as a measure of the uncertainty. This is problematic as it detects precision and not necessarily accuracy. Significant bias may be present if staff is not trained well. One solution is to randomly select a subset of biomass plots and have those remeasured by an independent and well-trained team of experts. The accuracy can then be quantified by comparing the original carbon stock densities by the ones measured by the professional team. However, employing independent field teams is expensive. Re-measurement of a proportion of plots by multiple field crews may be a less costly alternative.

C. *Combining Errors from Individual Components*

Error propagation

The error propagation method calculates compound errors through a set of simple rules. For example, the uncertainty for the sum of components is the simple square root of the sum of the squares of the component errors. Error propagation is simple to use and verify and does not require specialized computer software. However, it is based on simplifications and will yield biased results in cases when errors are extremely large, when errors are not distributed symmetrically or when components are related to each other. In these cases it is better to use a Monte Carlo approach.

Monte Carlo Approach

Often, a Monte Carlo approach is used to simulate the impact of variability of individual components of an equation onto the final value of the calculation. In a Monte Carlo approach, a large number of calculations of emissions or emission reductions are executed using the individual components that are slightly different in each iteration. For example, if the uncertainty of the biomass of an individual tree as introduced by an allometric equation is known, a Monte Carlo approach would simulate the overall uncertainty by adding a random error term to each individual tree. A Monte Carlo approach allows for specific probability density functions that describe the distribution of the values of the individual components. So, for example, if the biomass density is not symmetrically or normally distributed, but rather log-normally (which is often the case) a Monte Carlo approach can take this into account. In addition, relations between components can be included. For example, forest strata that have high aboveground biomass will likely have high soil carbon as well. In calculating the total carbon stock density and associated EF, a Monte Carlo approach can take into account such relations. However, the probability density functions of the different components are often challenging to determine. Similarly, to robustly estimate correlations among individual components, a sufficient number of observations are required.

D. Managing the Remaining Uncertainty

Good practice guidance codified in Standard Operation Procedures

The first and foremost method to manage uncertainty is to ensure that all procedures and their good practice guidance are codified in a consistent standard operating procedure (SOP). Due to the large number of participants required for quantifying ERs, all data collection and recording protocols need to be defined through a SOP designed with QA/QC safeguards to ensure conformity and catch mechanical errors from measurement, recording and transcription. SOPs serve not only to minimize errors, but also to ensure consistency among field teams and over time in case of operational definitions. It is important to allow flexibility in the procedures and measurements. Where accurate measurement equipment or procedures are not possible at reasonable costs, SOPs could specify alternative approaches for measurement that have lower accuracy but with greater uncertainty limits.

Uncertainty Thresholds

The first and most blunt approach to manage uncertainty is to require a maximal uncertainty for different components measured and/or the compound uncertainty. In case measurements do not meet this threshold, emission reductions are not eligible. In this case, measurement procedures must be revised and/or more measurements must be required. To allow flexibility, it is advised to set uncertainty thresholds such that a significant amount of uncertainty may still remain that can be managed with an uncertainty discounting approach. This approach would eliminate the worst cases while incentivizing challenging cases to manage uncertainty through improved measurement. In setting such a threshold it would be important for the CF to take into account Annex I country capcities and current reported uncertainties as compared to REDD+ country capacities. For example, Australia reports 30% uncertainty for emissions related to forest remaining forest (degradation) and 10% uncertainty for Forest converted to non-forest; Germany reports 21,8% uncertainty globally for the LULUCF sector bu 59,4% uncertainty for the specific forest remaining forest category.

Discount factors

If the uncertainties are large and cannot be eliminated through good practices, approaches can be followed to ensure the resulting estimate of emission reduction remains conservative. If implemented correctly, a deduction for uncertainty is an elegant solution to this end—it not only normalizes the volume of offsets across REDD+ activities or ER Programs with different circumstances and inherent uncertainties, it also incentivizes better measurement and provides some flexibility when the field measurement budget is small. Even with great care and effort, there is some likelihood that the calculated emission reductions are not conservative, i.e., the calculated emission reductions are greater than the true emission reductions. To ensure consistency across projects and across project types, one must standardize calculations according to the compound uncertainty. A deduction based on the compound uncertainty is a logical way of standardizing across projects and project types. Such a deduction is a deliberate bias to ensure that the calculated volume of emission reductions remains conservative. However, it must be noted that any deduction comes at the expense of the total number of credits, which will make carbon projects less cost-effective. Often, the half-width of the confidence interval around the mean emission reductions is used to quantify the magnitude of the discount.

Impact of Scale and Aggregation

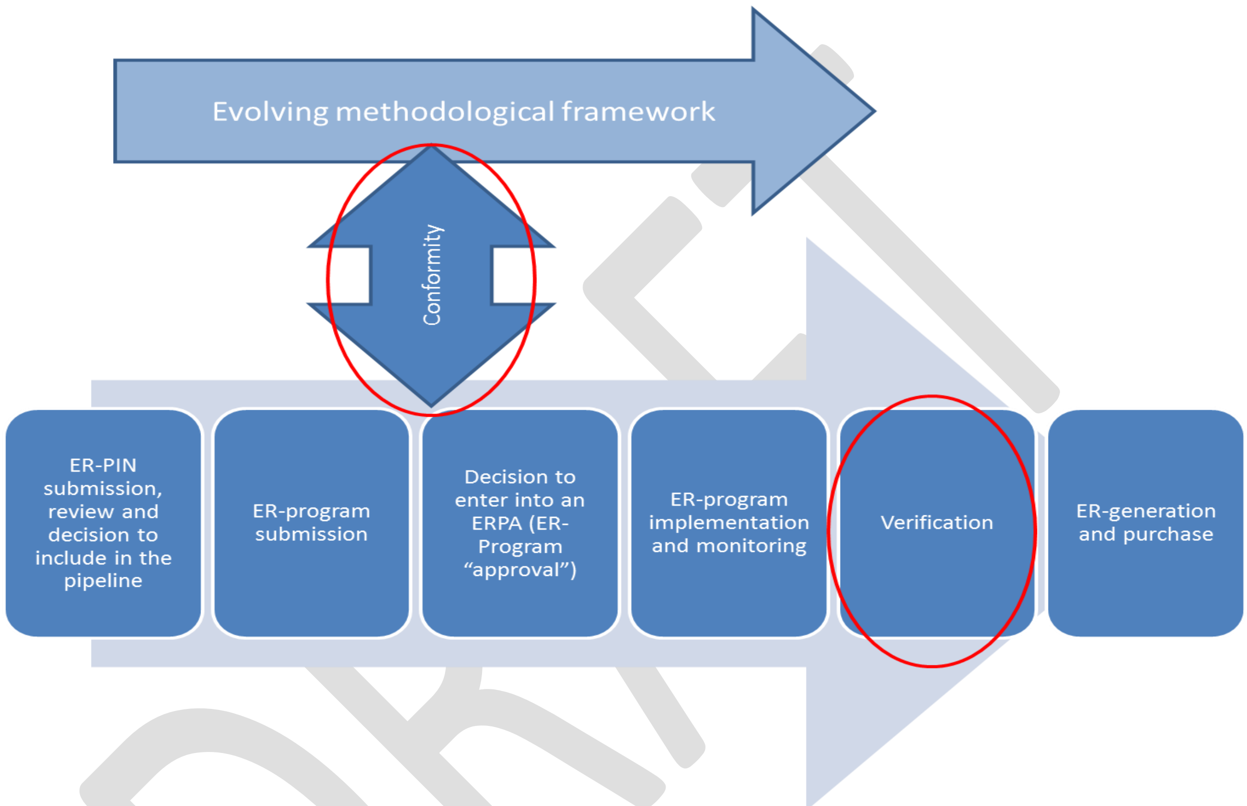
The question remains whether uncertainty thresholds, such as the CDM's proposed 15 percent at a 95 percent confidence level, should be calculated based on the sum of all emission reductions generated within the ER Program and across REDD+ activities, or whether the thresholds should be enforced for each REDD+ activity individually. The relative uncertainty of the sum of emission reductions across a number of REDD+ activities across the ER Program is smaller than the relative uncertainty of the emission reductions of only one REDD+ activity within the ER Program. If no bias exists, estimates of emission reductions are equally likely to be above or below the actual value. Hence, as more estimates from individual REDD+ activities are aggregated, they are likely to offset each individual point's deviation from the actual value. As more and more emission reductions are generated, the relative uncertainty will be reduced.

Importance of Stratification

Uncertainty can be significantly reduced by iteratively updating the stratification design. Often, increasing the number of strata after an initial stratification will reduce the compound uncertainty. Stratification is, therefore, often an iterative process in which the uncertainty within a stratum is used to decide whether re-stratification is necessary.

V. Verification

Verification is an integral part of the ER-program business process under the FCPF CF. In the context of the FCPF CF, verification is the process in which the ER-program document is checked for adherence to requirements established by the evolving *methodological framework* and ER-Program reporting is checked for accuracy, scientific validity and adherence to the requirements established by the ER-Program document (potentially integrating changes to the evolving methodological framework).



- *Verification* requires the establishment of benchmarks against which ER Program results can be assessed. The *methodological framework* should therefore set the scope, scale, frequency as well as requirements to deal with uncertainty of ER Program reporting. The *methodological framework* should contain sufficient QA/QC parameters.

- *The ER-program document* should provide the program-specific activity and measurement metrics from which a verifier or verifying body can ensure that all the requirements of the methodological framework are followed. ER-Program programs will undergo such a review process which is commonly called "validation" although in the context of the FCPF CF this is not a formal administrative step: the "validation" of an ER-Program will be implied by the decision of participants to negotiate an ER Program.

- *The Monitoring Report* will report ER results at intervals set by the ER-Program and in accordance with the methodological framework. To ensure credibility and adherence to methodology requirements, these results will require verification.

A. Key Considerations in Designing a Verification System

Key considerations in developing verification options include:

- *National Circumstances:* UNFCCC verification requirements under debate at the Doha COP 18 illustrate national concerns over the level of verification required, presumably due to the technical capacity of participating nations. Verification modalities under UNFCCC remain to be negotiated. ER credibility and fungibility will require a minimum standard of verification.
- *Costs/benefits:* The design of verification requirements has strong implications on the implementation costs and credibility of resulting ERs. There is a trade-off between the "thoroughness" of verification and the costs of increasing verification effort. The optimal cost/benefits balance is hard to strike at the onset; typically mechanisms aiming for strong integrity will invest heavily in verification at the beginning and may scale down over time as the understand how to better focus the verification effort.
- *Who does the Verification?* The independence of the verifier has strong implication for the credibility of resulting ERs. Issues include the type of validation/verification body (VVB), the number of VVB reviews required, the technical capacity of the VVB and the rigor applied in the verification process.
- *What is Verified?* The scope of verification also has implications on cost and credibility of ERs generated under an ER Program. Verification scope ranges from a high-level desk review of adherence to methodology requirements to detailed re-measurement of field data. The addition of technically complex methods in the scope of pools evaluated, for example LiDAR analysis for degradation, will require verifiers with a commensurate level of technical capacity.

B. Verification Options for the Carbon Fund

Transparent verification of project performance is a requirement of a credible results-based payment system. This assessment assumes that the monitoring report is subject to verification. The entity conducting the verification could use:

1. A desk review verifies the monitoring report against metrics established under the ER-program document in compliance with the methodological framework. Verification requires a validated template established at the onset of the project activity. In practice, under voluntary compliance schemes and

CDM these templates are a two-step process of a validated methodology which contains a quantitative framework for calculating emissions and removals under a generalized scenario and a project document which contains parameters and metrics specific to the project activities themselves.

- The verification process reviews measurement methods against the validated methodology and specific data values against those detailed in the project document.
- In the context of the FCPF CF, results could also be c*ross-checked with other sources* that have made similar assessments *when* available.
- The verification process is typically iterative in which verifiers provide an initial report of itemized discrepancies in the monitoring report allowing for the project implementer to respond and make adjustments as necessary.

2. Field visits ensure that reported values meet on-the-ground reality. Typically, a verifier will visit areas reported in land use classification maps to ensure that ground conditions meet those provided in the monitoring report. Likewise, the location of a random selection of field plots will also be visited and re-measured to verify the accuracy of reported results. Field verification of measurements increase ER credibility, but adds to the time and cost of the verification process.

Comparison of verification methods

Option	Advantages	Disadvantages
1.Desk review	The simplest verification design resulting in reduced costs and time.	• Increases the risk of inconsistent verification results between ER Program's. • Not as thorough as 2. and 3.
2. Cross verification with other sources[9]	• Cost effective. • Reveals potential biases and or inconsistencies in MRV methods used	Not always available
3. Field visits	Ensures a degree of rigor and thoroughness in the review and approval process.	More costly and time consuming.

[9] Cross validation could imply checking results across different studies that used comparable data sets. An example of cross verification/validation is the data on forest cover change in the DRC between 1990 and 2000 which has been assessed both by the Joint Research Center of the European Union with the Support of Universite Catholique de Louvain and by Satellite Observatory of Central Africa (OSFAC) with the support of South Dakota State University and Maryland University. The teams used different approaches (sampling vs wall to wall) to reach similar results.

Comparison of verification division of labor

Description	Advantages	Disadvantages
1.Verification requires assessment by one third party verifier	The simplest verification design resulting in reduced costs and time.	• Increases the risk of inconsistent verification results between ER Program's. • Not as thorough as 2 and 3.
2.Verification requires assessment by two third party verifiers	Ensures a degree of rigor and thoroughness in the review and approval process.	• More costly and time consuming. • Significant delays in approval if the auditors take different views.
3.FCPF assembles and independent expert panel and one third party verifier	• High level of rigor. • Increased transparency of the verification process.	Most costly and time consuming.
4.Country itself conducts verification	• More legitimate for REDD+ countries • Potentially lower cost (burden for verification is transferred to country)	Conflicts of interest between those developing ER-Program and those verifying will be more difficult to avoid

The options presented above are not mutually exclusive, and the design of a FCPF methodological framework should allow for flexible requirements based on proposed ER Program activities, country capacities and evolving UNFCCC requirements. Further, the question of who bares the costs of verification between the FCPF CF participants and the REDD+ countries will undoubtedly be an issue regardless of the options chosen.

VI. Recommendations

UNFCCC rules, modalities and guidance provide the general framework for MRV. There are numerous good guidance resources to assist in the design of the system, as well as specific technical methodologies addressing specific carbon pool quantification techniques. Key options related to scale and scope of measurement and reporting as well as verification as summarizes in the tables below.

A. Measurement and Reporting scope comparison

Description	Advantages	Disadvantages
1. Follow scope of RL/REL	• Simple to follow • Promotes consistency between the RL/REL and MR components and estimation of ERs • More cost effective than broader monitoring	• Ability to monitor and account for leakage tied to robustness of leakage accounting requirements
2. Follow scope of RL/REL + risk indicators	• Same as 1 • Provides metrics for establishing risk and focusing limited resources in priority areas.	• Same as 1 • Risk mediation may reduce the quantity of ERs available for sale. • Requires the establishment of risk assessment protocols.
3. IPCC Classes only	• Simple protocol to follow. • Likely to require only medium resolution imagery and can include Tier 2 EF. • Consistent with emerging UNFCCC modalities for national monitoring. • Builds on current MRV capacity efforts.	• Potentially contains a high degree of variability within classes. • Variability could reduce both the quantity and credibility of measured ER's.
4. IPCC Classes and risk indicators	• May reduce variability compared to 3. • Provides metrics for establishing risk and focusing limited resources in priority areas.	• Risk mediation may reduce the quantity of ERs available for sale. • Requires the establishment of risk assessment protocols.
5. Detailed strata	• The addition of a detailed stratification allows for reduced variation of EF factors within single strata and allows for more ways of registering ER from forest enhancement or land use change. • Reduced uncertainty and a greater selection of mechanisms for removals can result in more ERs.	• Requires a greater degree of technical capacity. • High resolution imagery requirements add substantial to measurement costs. • May only be feasible at limited scales.
6. Detailed strata + risk indicators	• Same advantages as 5. • Provides metrics for establishing risk and focusing limited resources in priority areas.	• Disadvantages of 5 • Risk mediation may reduce the quantity of ERs available for sale. • Requires the establishment of risk assessment protocols.

B. Measurement and Reporting scale comparison

Description	Advantages	Disadvantages
1. National MR	• Currently undergoing development. • Multi-source financing available. • Efficient use of limited resources. • Target of numerous capacity building initiatives.	• May have insufficient accuracy or precision to measure credible ERs for small scale activities. • Dependent on national MRV development timetable with anticipated 2014 completion date.
2. ER Program Level MR	• Tuned to specific ER Program measurement requirements. • Allows for higher resolution analysis and inclusion of detailed strata.	• Costs may make ER Program efforts financially unfeasible. • In-country capacity may be insufficient.
3. Combined MR	• Avoids unnecessary duplication of effort. • May ultimately improve national scale measurement.	• Additional costs above national measurement effort. • May require extensive interagency collaboration. • May highlight accuracy issues of a national scale system.

C. Managing remaining uncertainty

Description	Advantages	Disadvantages
Apply uncertainty thresholds	• Eliminate worst cases • Incentivizes challenging cases to manage uncertainty through improved measurement	• Threshold establishment subjective? • Incentivizes “meddling” with numbers rather than honest assessment
Discount for all uncertainty (apply conservativeness principle)	• Strong environmental integrity • Promotes fungibility of ERs with carbon markets • Incentivizes improvements to reduce discount	• Impossible for most ER-Program to generate ERs (results fall inside the noise) • Promotes RL inflation to avoid producing results inside the uncertainty bounds
Apply discount proportional to uncertainty (e.g. X% of uncertainty = y% discount on ER volume)	• Fair (discount are applied in a standardized way across countries) • Incentivizes improvements in measurement to reduce discounts	• Reduced environmental integrity
No discount for uncertainty	• Promotes honesty and full transparency in assessments	• Unfair (countries with high uncertainty generate same volumes as ER-programs with lower uncertainty) • Reduced environmental integrity

D. Comparison of verification options

Description	Advantages	Disadvantages
Verification requires assessment by one third party verifier	• The simplest verification design resulting in reduced costs and time.	• Increases the risk of inconsistent verification results between ER Programs. • Not as thorough as 2. and 3.
Verification requires assessment by two third party verifiers	• Ensures a degree of rigor and thoroughness in the review and approval process.	• More costly and time consuming.
FCPF assembles and independent expert panel and one third party verifier	• High level of rigor. • Increased transparency of the verification process.	• Most costly and time consuming.

E. Other Recommendations for MRV System Design

The options presented above are not mutually exclusive, and the design of a FCPF methodological framework should allow for flexible requirements based on proposed ER Program activities, country capacities and evolving UNFCCC requirements. The design of an ER Program MR system will be determined by UNFCCC modalities and accuracy enhancements required for bounding uncertainty for producing credible ERs. ER Program protocols should also allow for periodic updates to include new UNFCCC modalities, advances in measurement methods and increased ER Program country capacity.

1. List of Detailed Activity Data and Emission Factors Measurement Data Sources

Data availability and suitability for measurement of activities are key issues for a successful monitoring and measurement system. ER Program activities proposed under an ER Program require available and suitable data sources. These sources should specifically be identified and matched to each proposed ER Program activity. They should be assessed for suitable resolution, availability and cost. A detailed MRV plan should be submitted with the ER Program, allowing for the assessment of the feasibility of monitoring and measurement program activities. This should include:

2. Review of Technical Capacity

Program personnel or consultant pool should be technically proficient at monitoring and measurement techniques required to report and verify ER Program activity results. ER Program participants should be required to demonstrate their technical capacity of conducting the required monitoring and measurement of each proposed ER Program activity.

3. Uncertainty Analysis of Monitoring and Measurement, Reporting and Verification Components

Performance-based measurement of ERs requires the integration of error prevention via SOPs, tracking of uncertainty for all inputs to ER calculations and a spatially explicit monitoring system capable of identifying activity specific results. In the absence of an international REDD+ compliance scheme, REDD+, under UNFCCC guidance, is currently focused on readiness and has not set modalities for protocols ensuring fungibility. The FCPF can draw on the experience and methodologies of existing compliance standards and voluntary market REDD+ efforts to identify mechanisms of ensuring credibility and reducing delivery risk under a transparent monitoring system. The FCPF should require an uncertainty analysis of each component that is measured and reported. Options for managing remaining uncertainty require further analysis.

4. **Anthropogenic vs. non-anthropogenic sources of emissions.**

Dinstinguishing between anthropogenic and non-anthropogenic sources of emissions does not appear feasible in the context of the Carbon Fund.

5. **Independent Verification**

Under the principle of objectivity and independence, the requirement of an independent verifying body limits bias and adds to the credibility of the ERs. The verifying body should have the technical capacity commensurate to the monitoring and measurement methods used.

6. Decision Trees to Address the Wide Range of ER Program Options

The development of task specific decision trees for the range of allowable emission reduction programs under the FCPF Carbon Fund would provide FCPF partner countries valuable guidance for submitting their ER Programs and supporting materials. Decision trees should assist in assessing technical capacity, designing QA/QC and uncertainty reporting systems, setting the scope and scale of monitoring and measurement, and identification of data and technical resources.

VII. Potential Next Steps : Topics on Which Further Analysis or Discussion is Needed

[to be elaborated after the Forum and Working Group discussion]

Appendix: Overview of the UN-REDD National Forest Monitoring System

The UN-REDD program has developed a National Forest Monitoring Systems structured under three pillars.

i) Pillar 1 = Activity Data: A Satellite Land Monitoring System (SLMS) to collect and assess over time AD related to forest land
ii) Pillar 2 = Emission Factors: National Forest Inventory (NFI) to collect information relevant for estimating emissions and removals and provide emissions factors, i.e. forest carbon stocks and forest carbon stock changes
iii) Pillar 3 = Emissions/Removals: Provide the bases for a national GHG inventory as a tool for reporting on anthropogenic forest-related GHG emissions by sources and removals by sinks to the UNFCCC Secretariat

UN-REDD pilot programs, along with UNFCCC national communications, provide insight into the current state of monitoring and reporting capacity in many of the FCPF participating non-Annex 1 nations. The emerging status of a REDD+ agreement under UNFCCC, along with the preliminary of national scale monitoring and reporting systems, do not currently allow for the creation of credible and fungible, open-market ERs due to the non-existence of working monitoring systems for most FCPF participating nations and high levels of uncertainty reported by those nations who do have a monitoring system in place. At this time, these systems are too preliminary to meet ER Program measurement needs.

To efficiently allocate limited resources, the UNFCCC advocates prioritizing the level of detail in assessing the emission factors of the various carbon pools. The prioritization is accomplished by defining key categories. A key category is a GHG source or sink for which variation due to uncertainty would have a substantial influence on national GHG reporting net results. Tier 1 and Tier 2 methods are not spatially explicit and require only the use of IPCC methods and published data, default IPCC values for Tier 1 and country specific data for Tier 2. Tier 3 involves the greatest level of analysis and requires spatially explicit monitoring and measurement. Tier 3 pools are identified via the key category prioritization. As such, the scope and depth of analysis of a MRV system must balance resource capacity with the needs of producing verifiable results-based ERs. Remote sensing, biomass plots and GIS are the primary tools of Tier 3 monitoring.

IPCC Key Category Tiers

Tier	Details
Tier 1	IPCC methods and IPCC default values (no data collection needed)
Tier 2	IPCC methods and country specific data for key factors (including more detailed country specific strata)
Tier 3	Country specific methods or models, national inventory of key carbon stocks, repeated measurements of permanent plots to directly measure changes in forest biomass

A UN-REDD modeled system would require, at a minimum, that the SLMS monitor AD carbon stock and changes of carbon pools in the six broad IPCC land use categories (forest land, cropland, grazing land, wetlands, settlements, and other land). While UNFCCC standards will possibly include degradation and enhancement, currently developed national monitoring systems are only capable of monitoring some land-use change, deforestation, afforestation and reforestation activities. EF factors are to be derived from NFI. Suitable platforms for national monitoring include medium resolution multispectral and radar satellite imagery. Emissions and removals are calculated by multiplying AD and EF for each carbon pool. Furthermore, UNFCCC provides for adjustments allowing for national circumstances in regards to technical capacity. If you discount for uncertainty under these guidelines, ER measurement uncertainty is likely to greatly reduce or even negate all quantifiable ERs.