Section Summary

Stephen Hagen[1]

Well planned sampling designs and robust approaches to estimating uncertainty are critical components of forest monitoring. The importance of uncertainty estimation increases as deforestation and degradation issues become more closely tied to financing incentives for reducing greenhouse gas emissions in the forest sector. Investors like to know risk and risk is tightly linked to uncertainty. Uncertainty assessment is also important for evaluating the implications of forest management actions, and it helps us identify and design future projects to reduce uncertainty.

Design considerations

The conservative principle for accountable carbon credits allows credit for the estimated lower 95% confidence limit of carbon stock resulting from the project minus the expected baseline carbon stock. This serves as a financial incentive to reduce uncertainties where possible. Uncertainties in forest stock carbon estimates can be reduced by a) inventory planning and sampling design optimization; b) technological development; and c) biomass allometric models.

A rational decision about an optimal design can be made only by comparing the set of alternatives using objective selection criteria that combine information on survey cost and the achievable reliability of the results. Many choices exist for optimizing sampling design. For example, the selection of plot locations can be chosen by systematic sampling, simple random sampling or stratified random sampling. The optimal plot design depends on objectives, costs, and variability of the population's attributes of interest. Inventory planning steps include 1) identifying information needs and priorities; 2) assembling and evaluating existing data to answer the questions; 3) selecting the main monitoring components; and 4) setting the precision and cost requirements to compute the optimal sample size. During optimization, one should minimize the total inventory cost subject to fixed precision requirements (confidence interval and confidence level) for key attributes of interest. After this process, the objectives and priorities should be reevaluated and the process should be repeated.

There are many technological approaches that can help reduce error and increase confidence. Some examples include, a) improving the link between remote sensing and ground sampling; b) matching the imagery pixel resolution with the field plot size; c) improving the temporal resolution of remote sensing data; and d) reducing locational error by using survey-grade GPS to improve the links between remote sensing imagery and ground sampling (but with increased cost). In terms of statistical methods, several

[1]Applied Geosolutions, 87 Packers Falls Rd., Durham, NH 03924
Corresponding author: shagen@appliedgeosolutions.com

technological developments are potentially useful including a) using regression to link modeled imagery to plots; b) using double sampling for stratification or regression; c) using stratification to create homogenous strata with respect to key attributes and costs; d) allocating plots to strata based on stratum size, variance, and/or costs; and e) making all plots permanent initially, but remeasuring based on change strata.

Improvements to biomass allometric models may include a) developing regional models to help cover the range of species, growth forms, and wood densities; b) including large trees and buttressed trees in equations; c) using Randomized Branch Sampling on a subsample of trees as a means to correct for bias in the biomass models; and d) terrestrial laser scanning for individual stem volume (which shows promise, but may be challenging to apply in the tropics).

Uncertainty

Approaches for quantifying uncertainty should be decided upon at the initiation of a project. Unfortunately, this is typically done, if at all, at the end of a project; after the final products have been created.

Uncertainties are a composite of errors arising from observations and models. Different types of errors can be quantified by their precision, accuracy, or bias. Precision is the variation about the sample mean. Accuracy is the variation about the true mean, which includes both precision and bias. Bias occurs when the sample mean differs systematically from the true mean. Uncertainties can arise from multiple sources. For instance, we will never know the true values of large forest population parameters, so we measure a subsample of all of the individuals. The inaccuracy of estimation of these parameters from sample results is termed the sampling error. The size of sampling errors can be controlled by the survey design and the size of the sample. Non-sampling errors come from all other sources, such as faulty application of definitions, classification errors, measurement errors, model application errors, calculation errors, and sampling frame errors. These errors must be addressed on a case-by-case basis, through training and other quality control measures. To calculate the error budgets of total survey error, one needs to systematically identify and define all potential sources of error, Assess importance of each error and develop cost functions for reducing them, and balance the risk of mistakes due to the uncertainties with the cost of reducing the magnitude of the uncertainties.

The most common approach to combining uncertainties from multiple sources is to "sum by quadrature". In this approach, given uncertainties from multiple sources, total uncertainty is calculated as the square root of the sum of the squared uncertainty from all individual sources. This approach requires assumptions about the data, such as having a normal distribution, that are often incorrect.

A Monte Carlo framework with bootstrapping allows one to combine uncertainties from many sources and does not require assumptions about data distributions. It provides a means of handling non-linear models and data with complicating characteristics such as leptokurtosis (peakedness) and heteroscedasticity (non-constant variance).

The basic steps of a Monte Carlo framework include a) fit a model to the existing data; b) calculate the model residuals; c) create a new realization of the existing data by adding a random draw of residuals back to the original model predictions; d) fit a new model to this new realization to estimate or predict; and e) repeat *n* times. This approach provides *n* estimates for each new data point and the distribution created by the *n* estimates can be used to quantify uncertainty. This distribution of estimates can be used as input into another model or otherwise combined with other observations to link uncertainty from multiple sources. Estimates from this approach can be combined efficiently to allow for spatial aggregation of uncertainty estimates, which often require dubious assumptions when done analytically. The downside to this approach is that it is data and computationally intensive.

Conclusion

Inventory planning is an optimization problem. Sampling error can be minimized for a given cost or the cost can be minimized for a desired level of maximum acceptable error. Monitoring costs need to be smaller than potential financial benefits. It is critical, however, to identify and manage all sources of error from early in the project. The IPCC should consider revisiting the 95% confidence interval due the potentially high cost of a monitoring system to meet this requirement.

Several issues must be addressed during implementation of an uncertainty accounting. Complete accounting of uncertainty is important, but agreement is needed as to what the project boundaries are (i.e. which sources of uncertainty should be included). Collaborative effort is required for complete uncertainty accounting. Sharing models and data is required among project members.