Uncertainty Modules in moja global

v1.0.0

The purpose of this document is to briefly describe the uncertainty modules that are included within the FLINT. This report draws heavily on the Design for Monte Carlo and Propagation of Error Uncertainty in the Full Lands Integration Tool and the FLINT Uncertainty Manual. The report was drafted with support of multiple contributors, including the support of the Government of Canada and the United Nations Framework Convention on Climate Change.

Introduction

Uncertainty is used for different purposes in land sector greenhouse gas reporting and accounting. This introductory section details the key requirements from the most common land sector greenhouse gas (GHG) reporting and accounting frameworks for national and jurisdictional levels, as well as being required under voluntary schemes.

There are two main methods of estimating uncertainty, Propagation of Error and Monte Carlo Uncertainty assessment. The IPCC guidance for these two methods is almost entirely focused on the context where emissions are calculated from activity data (AD) and emission factor (EF) information. The key difference in the use of the two methods is that the IPCC propagation of error approach assumes that the usual assumptions about the independence of variables and normally distributed errors apply, whereas the Monte Carlo approach provides methods for incorporating covariance and allows analysis of errors that may be asymmetrical. Uncertainty assessment in FLINT also needs to deal with a third situation, where assessments are based on models that may contain numerous uncertainties attributable to model input data, parameters, or algorithms. In this situation, the IPCC provides little specific guidance. This is because there are not yet general methods that universally apply, and therefore model users and developers must typically create customized solutions appropriate to their needs. FLINT provides the ability to do this and the Monte Carlo design in this document is one method for doing this. As noted in the introduction, the FLINT uncertainty module has been developed for the purpose of assessing the ‘precision’ component of uncertainty related to random error/variability. It is assumed that any systematic errors, or issues of ‘accuracy’ in IPCC terms have been identified and corrected prior to analysis in FLINT.

The Uncertainty Assessment is described here as a module, however, in practice, the Monte Carlo Uncertainty implementation will require changes and functionality enhancements in existing FLINT libraries as well as new modules. The enhancements represent both functionality and calculation aspects of FLINT modules, as the Monte Carlo needs to control simulations and aggregate outputs consistent with the functionality module definition; as well as, specifying and changing the state of input variables.

Propagation of Error

IPCC Approach 1 Propagation of Error is relatively straightforward to apply and suits applications where the simplified AD X EF approach is used. For a full explanation of the key assumptions and requirements of Approach 1 Propagation of Error, refer to section 3.2.3.1, Volume 1 Chapter 3 of the 2006 IPCC Guidelines. This approach can be used to quantify uncertainty in estimates for any year as well as uncertainty in trends.

In the Propagation of Error approach, uncertainty in GHG emissions or removals is calculated by propagating the uncertainty of the inputs. This requires estimates of the mean and the standard deviation of all inputs.

There are two main steps to the method for calculating the uncertainty for any particular year. The first is combining the uncertainties of the emission factor and the activity data by multiplication (Equation 1) and the second is combining these uncertainties through addition to arrive at an aggregate uncertainty (Equation 2).

Equation 1

$$U_{total} = \sqrt{U_1^2 + U_2^2 + ... + U_n^2}$$

Where:
U_{total} = the percentage uncertainty in the product of the quantities (half the 95 percent confidence interval divided by the total and expressed as a percentage);
U_i = the percentage uncertainties associated with each of the quantities

Equation 2

$$U_{total} = \frac{\sqrt{(U_1 \times x_1)^2 + (U_2 \times x_2)^2 + ... + (U_n \times x_n)^2}}{|x_1 + x_2 + ... + x_n|}$$

Where:
U_{total} = the percentage uncertainty in the sum of the quantities (half the 95 percent confidence interval divided by the total (i.e., mean) and expressed as a percentage). This term ‘uncertainty’ is thus based upon the 95 percent confidence interval;

X_i and U_i = the uncertain quantities and the percentage uncertainties associated with them respectively (According to the IPCC guidance the quantities X_1... X_n should be independent i.e., the covariances are assumed to be zero).

Monte-Carlo Analysis

Monte Carlo Monte Carlo analysis involves repeated model simulations where the values of model inputs are determined on a stochastic basis by randomly selecting their value from within their individual probability density functions. The outputs of these repeated model simulations are then used to calculate uncertainty statistics for the particular variables of interest (IPCC, 2006; McMurray et al., 2017).

The conceptual process for Monte Carlo is relatively straight forward:

1. Specify probability density functions for model inputs and any correlation between inputs
2. Run repeated simulations
    a. Select input values randomly from within their Probability Density Functions (PDF)
    b. Run model using the selected input values
    c. Repeat specified number of times
3. Combine inputs from repeated simulations and calculate Uncertainty Statistics

In the context of FLINT for greenhouse gas estimation, we use the term model here in its broad sense and therefore model inputs can range from inputs for basic ‘models’ such as emission factors through to inputs for more advanced empirical and process based ‘models’ such as decomposition rates, growth curve parameters,and turnover fractions etc.

Uncertainty for FLINT

With support from the United Nations Framework Convention on Climate Change, a monte-carlo functionality has been added to the FLINT framework. Monte-Carlo analysis is more suited to the FLINT framework, as it allows for representation of spatial and non-spatial variables. The uncertainty functionality is a relatively simple process to describe, although there is complexity in the implementation of this functionality especially for a spatial modelling system, and to allow for flexibility for the user.

Monte-Carlo Implementation

In general terms, when the FLINT is running a simulation, a module will call for specific variables required to operate. These may be a single value, an object like variable or a table. A provider will return the relevant variables to the module. As the variables are returned to the module, the uncertainty functionality will ‘intercept’ them, and replace them with a random selection from a defined probability distribution based on the user input. The FLINT will then run multiple iterations of the simulation (e.g. 1000 times), returning randomly sampled values from these distributions.

The Probability Density Functions available for the user are: Normal, Triangular, or Manual

These PDFs can be applied to any variable, as determined by the user (See the FLINT Uncertainty Manual for details). The outputs from the uncertainty analysis will be written to a table, and will include :

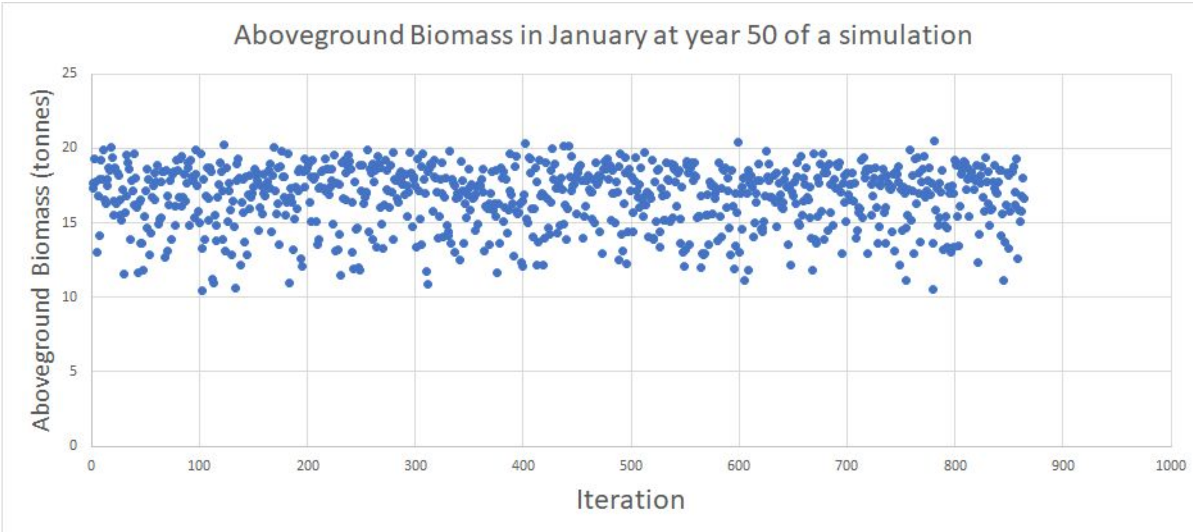
- Year of the flux (year)
- Location ID (localdomain_id)
- Flux source pool (src pool)
- Flux sink pool (sink_pool)
- Mean flux from all iterations (mean)
- Standard deviation from all iterations (S.D.)
- Margin of error (margin of error)
- The lower limit of the estimate (90% limit low)
- The upper limit of the estimate (90% limit high)
- Optional raw values

This will allow for further analysis of the outputs in Excel or other statistical packages (Table 1, Figure 1).

Table 1 - Example output of the Uncertainty functionality, with distribution characteristics for flux values.

year	**localdomain_id**	**src pool**	**sink pool**	**mean**	**S.D.**	**margin of error**	**90% limit low**	**90% limit high**
2001	1	atmosphereCM	aboveGroundCM	0.032	0.008	0.001	0.031	0.033
2001	1	atmosphereCM	belowGroundCM	0.007	0.002	0.000	0.007	0.008
2002	1	atmosphereCM	aboveGroundCM	0.091	0.024	0.004	0.087	0.095
2002	1	atmosphereCM	belowGroundCM	0.021	0.005	0.001	0.020	0.022
2003	1	atmosphereCM	aboveGroundCM	0.144	0.037	0.006	0.138	0.150
2003	1	atmosphereCM	belowGroundCM	0.033	0.008	0.001	0.032	0.035
2004	1	atmosphereCM	aboveGroundCM	0.191	0.047	0.008	0.183	0.199
2004	1	atmosphereCM	belowGroundCM	0.044	0.011	0.002	0.042	0.046
2005	1	atmosphereCM	aboveGroundCM	0.232	0.056	0.009	0.223	0.241

2005	1	atmosphereCM	belowGroundCM	0.053	0.013	0.002	0.051	0.056



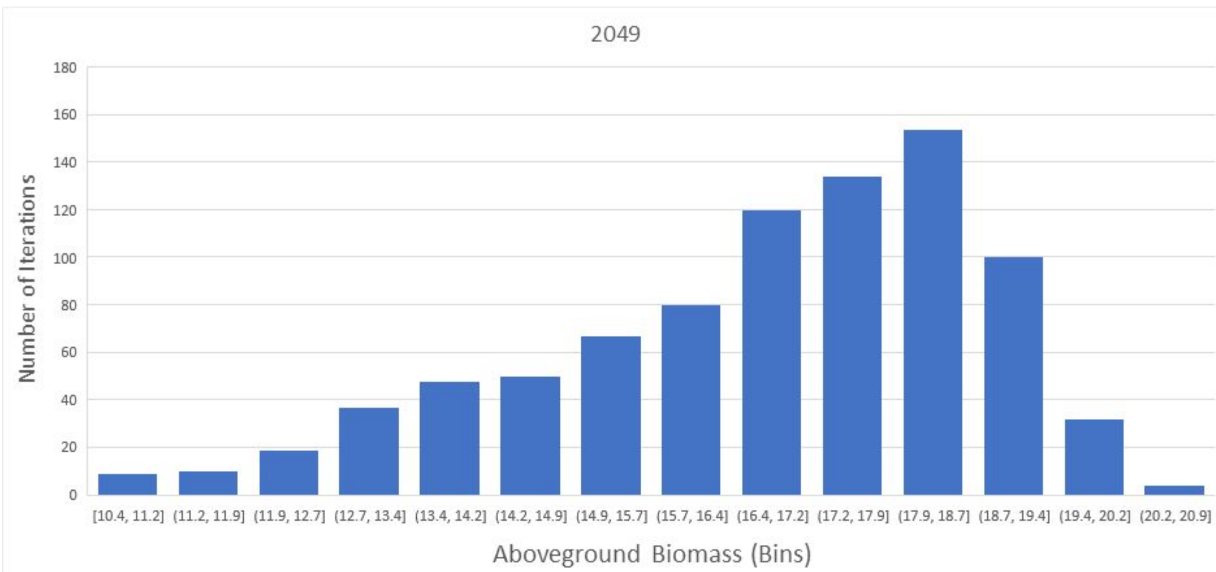


Figure 1 - Raw output for a test simulation completed with a simple forest growth curve (Chapman-Richards Curve) for one month of a simulation. Graphs show the raw outputs for each iteration, and a histogram of outputs. The non-normal distribution is caused by one of the three variables being modified using a triangular PDF.

References

IPCC, 2006. 2006 IPCC guidelines for national greenhouse gas inventories.

McMurray, A., Pearson, T., Casarim, F., 2017. Guidance on applying the Monte Carlo approach to uncertainty analyses in forestry and greenhouse gas accounting. World Resources Institute.