Handbook to propagate uncertainty of measurement errors and allometric models

quantufy uncertainty program (quest) - suny

RuTH Yanai, oswaldo carrillo & craig wayson

2016

Contents

[Sources of uncertainty in the estimation of biomass 2](#_Toc463949413)

[Measurement errors 2](#_Toc463949414)

[Errors in allometric models. 3](#_Toc463949415)

[Sampling Error. 4](#_Toc463949416)

[The Montecarlo method. 5](#_Toc463949417)

[Process to propagate uncertainties in Excel 6](#_Toc463949418)

[Process to estimate the measurement error 6](#_Toc463949419)

[Process to obtain the parameters model. 6](#_Toc463949420)

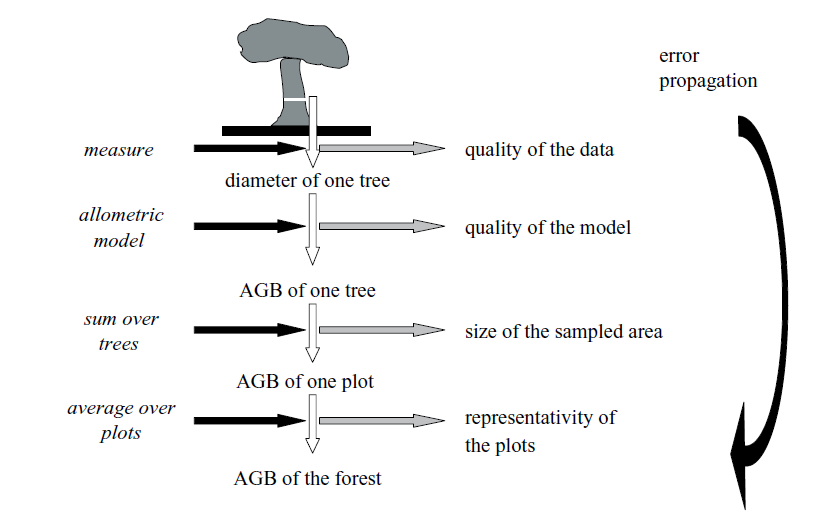
[Process to propagate uncertainties of the measurement errors and the allometric models. 11](#_Toc463949421)

[Propagation of uncertainty at a plot level. 14](#_Toc463949422)

[Process to propagate uncertainties in R 15](#_Toc463949423)

# Sources of uncertainty in the estimation of biomass

According to Chave (2004) in the estimation of biomass there are 4 main sources of uncertainties that are associated to: measurement error, the allometric model, size of the plots and sampling error.



In the context of national Green House Gases (GHG) inventory for the forest sector, the estimation of carbon stocks and carbon stock changes of Above Ground Biomass (AGB) needs a quantification of different sources of uncertainties and its correct propagation according to the Guidelines of Good Practice of the Inter-Government Panel for Climate Change (IPCC).

So, a complete uncertainties analysis for AGB needs in first place the quantification of the uncertainties of each source and then the propagation of these. For the quantification of uncertainties of measurement errors it is necessary to have a well knowledge of measurement in field; and in terms of error in allometric equations we need to know the characteristics of the model. In this way, next is explained, what do these sources of error in the estimation of biomass consist of.

## Measurement errors

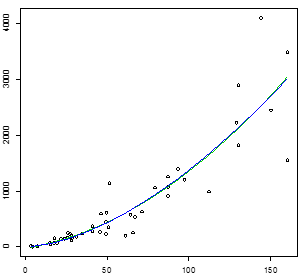
The biomass area is typically estimated indirectly through the Diameter at Breast Height (DBH). This variable (DHB) is normally obtained through direct measurement of the trees through forest inventories. While, in those forest inventories there are protocols to assure the quality of the measurements of the DBH, there measurement is subject to different sources of error like the one associated to the errors in the instrument of measurement, to the measuring process, and to the skill of the teams, among others.



In practice, the measurement of each of these sources of errors is complicated; however, there are practical techniques that can help to quantifying, like the re-measurement of trees through others teams. In other words, through a re-sampling, it is possible to measure the same trees 2 times (with different team) and then quantify the differences between both re-measurements of DBH.

## Errors in allometric models.

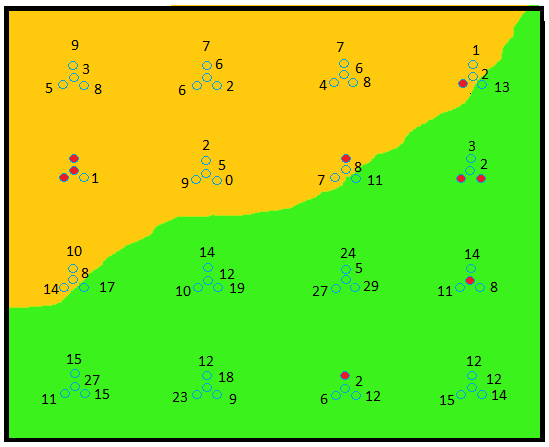
As well, models are a representation of reality and in the case of allometric models these try to estimate the AGB through a biometric variable like the DBH, Height among others.

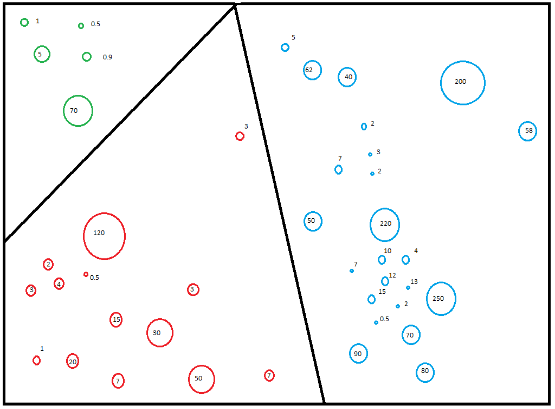


The allometric models are typically built through a sample of trees that have been cut down and weighed. As Yain mentions (2016), when changing a sample, the adjustment of the model will change slightly because the model is being built only with a sample and not with the whole population. In that case, the biomass estimations are obtained through models that are subject to model error.

## Sampling Error.

When we need to obtain an estimate of biomass through measured trees in different plots, there is a source of error associated to the natural variability that is presented between different stocks of biomass at a plot level. This source of error is quantified through sampling error and is the source uncertainty that is typically quantified. The way to estimate it depends basically on the type of sample design. For example, if we have a sample design like the one that is shown on the first figure, there are different estimators to use and depending on them, the way the variance will be obtained. Also, if the sample design changes, the estimator will also change and then the expression to obtain the variance that will be the one to obtain the sampling error.





In practice, this error is adequately estimated, so then in the rest of the presentation we will focus in showing in a general way how is it possible to obtain the measurement error and then we will concentrate especially in the estimation of uncertainties of models.

# The Montecarlo method.

There are analytical solutions to simple cases of error propagation. For example, the variance of a sum is the sum of the variances of the individual terms, if the terms can be assumed to be independent. In the case of forest biomass, however, the calculations are too complex to be solved analytically.

A Monte Carlo simulation can be used to characterize uncertainty in a complex result by simple repetition, thanks to modern computing capability.

The Monte Carlo approach to error propagation involves making a calculation many times, each time with a different random sample of the input values that reflects their uncertainty (the distribution of possible values). After many iterations, the distribution of the many results is used to characterize the uncertainty of the result due to the uncertainty in the inputs. The Monte Carlo approach is easy to implement, but it is also easy to make mistakes in representing the uncertainty in the inputs and in deciding at what level to randomly select values

In the case of above ground biomass, the propagation of uncertainties requires for the errors in measurement and the allometric model in a nested way. In other words, first some variations are generated from DBH through measurement errors and these random values of DBH are used in the allometric models in which also variations associated to model errors are included.

# Process to propagate uncertainties in Excel

## Process to estimate the measurement error

Let's suppose we have a set of trees measured by different brigades in different times slightly. Taking into account that information we proceed as follow:

* First we capture in an excel column the DBH measured by brigade 1
* Then in another excel column we capture DBH measured by brigade 2.
* After that, we get the difference between the measurements and
* Finally, we obtain the standard deviation of the differences.

Of course there are other more complicated methods to estimate the variations of the differences of the re-measurements; however, in this workshop we are more interested in describing the procedure to estimate the uncertainties of the models; so we will now concentrate on this point.

## Process to obtain the parameters model.

First, let’s suppose we have data of a destructive sampling of 48 trees cut down and weighed. In this way, in the first column, starting from row 14, we show the Ids of 48 trees that were cut down; in column B the DBH of each of the cut down trees and in column C we can see the measured biomass from each of the trees (kg).

|  |  |  |
| --- | --- | --- |
| **A** | **B** | **C** |
| Id\_tree | DBH | Biomass\_kg |
| 1 | 19.09 | 68.0 |
| 2 | 31 | 179.4 |
| 3 | 21.64 | 139.2 |
| 4 | 40.87 | 369.4 |
| 5 | 28.17 | 213.3 |
| 6 | 27.11 | 224.1 |
| 7 | 17.18 | 69.6 |
| 8 | 15.27 | 46.6 |
| 9 | 28.32 | 152.0 |
| 10 | 34.37 | 240.8 |
| 11 | 15.08 | 46.6 |
| 12 | 17.5 | 153.4 |
| 13 | 4.2 | 2.0 |
| 14 | 4.13 | 5.8 |
| 15 | 28.01 | 120.0 |
| 16 | 49.33 | 608.1 |
| 17 | 3.18 | 11.7 |
| 18 | 25.46 | 160.9 |
| 19 | 51.24 | 1141.7 |
| 20 | 14.96 | 78.7 |
| 21 | 23.87 | 148.0 |
| 22 | 7.32 | 8.6 |
| 23 | 25.94 | 246.8 |
| 24 | 45.6 | 262.6 |
| 25 | 46 | 598.1 |
| 26 | 49 | 443.2 |
| 27 | 50.5 | 356.4 |
| 28 | 65.6 | 256.5 |
| 29 | 49.0 | 225.4 |
| 30 | 41.0 | 277.6 |
| 31 | 61.0 | 198.9 |
| 32 | 64.0 | 576.6 |
| 33 | 87.0 | 1253.4 |
| 34 | 71.0 | 632.2 |
| 35 | 79.0 | 1057.7 |
| 36 | 87.0 | 1068.4 |
| 37 | 97.0 | 1205.9 |
| 38 | 93.0 | 1387.3 |
| 39 | 87.0 | 909.9 |
| 40 | 67.0 | 528.2 |
| 41 | 112.0 | 992.6 |
| 42 | 129.0 | 2219.1 |
| 43 | 144.0 | 4106.2 |
| 44 | 160.0 | 3480.6 |
| 45 | 160.0 | 1554.3 |
| 46 | 130.0 | 1821.3 |
| 47 | 150.0 | 2443.8 |
| 48 | 130.0 | 2898.4 |

With this information, we proceed to run a regression in the way:

B=exp^(a + b\*ln(DBH))

However, we can simplify this model in a linear form:

ln(B)= a + b\*ln(DBH)

In that case, to be able to run this model, its necessary to obtain the logarithm of biomass and the DBH shown in columns D and E.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **A** | **B** | **C** | **D** | **E** |
| Id\_tree | DBH | Biomass\_kg | ln(Biomass\_kg) | ln(DBH) |
| 1 | 19.09 | 68.0 | 4.2 | 2.95 |
| 2 | 31 | 179.4 | 5.2 | 3.43 |
| 3 | 21.64 | 139.2 | 4.9 | 3.07 |
| 4 | 40.87 | 369.4 | 5.9 | 3.71 |
| 5 | 28.17 | 213.3 | 5.4 | 3.34 |
| 6 | 27.11 | 224.1 | 5.4 | 3.30 |
| 7 | 17.18 | 69.6 | 4.2 | 2.84 |
| 8 | 15.27 | 46.6 | 3.8 | 2.73 |
| 9 | 28.32 | 152.0 | 5.0 | 3.34 |
| 10 | 34.37 | 240.8 | 5.5 | 3.54 |
| 11 | 15.08 | 46.6 | 3.8 | 2.71 |
| 12 | 17.5 | 153.4 | 5.0 | 2.86 |
| 13 | 4.2 | 2.0 | 0.7 | 1.44 |
| 14 | 4.13 | 5.8 | 1.8 | 1.42 |
| 15 | 28.01 | 120.0 | 4.8 | 3.33 |
| 16 | 49.33 | 608.1 | 6.4 | 3.90 |
| 17 | 3.18 | 11.7 | 2.5 | 1.16 |
| 18 | 25.46 | 160.9 | 5.1 | 3.24 |
| 19 | 51.24 | 1141.7 | 7.0 | 3.94 |
| 20 | 14.96 | 78.7 | 4.4 | 2.71 |
| 21 | 23.87 | 148.0 | 5.0 | 3.17 |
| 22 | 7.32 | 8.6 | 2.2 | 1.99 |
| 23 | 25.94 | 246.8 | 5.5 | 3.26 |
| 24 | 45.6 | 262.6 | 5.6 | 3.82 |
| 25 | 46 | 598.1 | 6.4 | 3.83 |
| 26 | 49 | 443.2 | 6.1 | 3.89 |
| 27 | 50.5 | 356.4 | 5.9 | 3.92 |
| 28 | 65.6 | 256.5 | 5.5 | 4.18 |
| 29 | 49.0 | 225.4 | 5.4 | 3.89 |
| 30 | 41.0 | 277.6 | 5.6 | 3.71 |
| 31 | 61.0 | 198.9 | 5.3 | 4.11 |
| 32 | 64.0 | 576.6 | 6.4 | 4.16 |
| 33 | 87.0 | 1253.4 | 7.1 | 4.47 |
| 34 | 71.0 | 632.2 | 6.4 | 4.26 |
| 35 | 79.0 | 1057.7 | 7.0 | 4.37 |
| 36 | 87.0 | 1068.4 | 7.0 | 4.47 |
| 37 | 97.0 | 1205.9 | 7.1 | 4.57 |
| 38 | 93.0 | 1387.3 | 7.2 | 4.53 |
| 39 | 87.0 | 909.9 | 6.8 | 4.47 |
| 40 | 67.0 | 528.2 | 6.3 | 4.20 |
| 41 | 112.0 | 992.6 | 6.9 | 4.72 |
| 42 | 129.0 | 2219.1 | 7.7 | 4.86 |
| 43 | 144.0 | 4106.2 | 8.3 | 4.97 |
| 44 | 160.0 | 3480.6 | 8.2 | 5.08 |
| 45 | 160.0 | 1554.3 | 7.3 | 5.08 |
| 46 | 130.0 | 1821.3 | 7.5 | 4.87 |
| 47 | 150.0 | 2443.8 | 7.8 | 5.01 |
| 48 | 130.0 | 2898.4 | 8.0 | 4.87 |

With the information from columns D and E and, using the statistical software R, we adjusted a linear model:

ln(B)= -0.46213 + 1.65874\*ln(DBH) Eq 1

So then, the resulting parameters were:

a= -0.46213

b= 1.65874

After that, to quantify the model error, we proceed to obtain the Mean Square Error (MSE) according to the next expression.

Eq 2

where

Yi = any value of Y used in constructing Eq. 1

= the corresponding predicted value obtained from Eq. 1

n = number of observations in the sample used to estimate the parameters in Eq. 1.

In terms of the model we are developing, the MSE that can be written the next way.

So, to obtain the MSE we proceed to obtain predictions from the biomass using the regression model that we previously mentioned, like its shown in column F and after, we obtain the square of the difference between the natural logarithm of the real biomass and the estimated one like it is shown in column G; these square differences are added to the cell "G62" and this is the numerator to obtain the MSE. Finally, "n" is the sample size, that as it can be observed in column "A" 48 trees were cut down. The results of this operation are shown in cell “B9”.

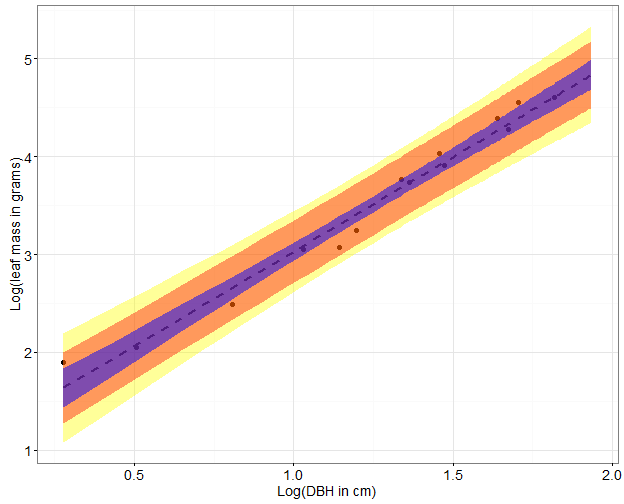
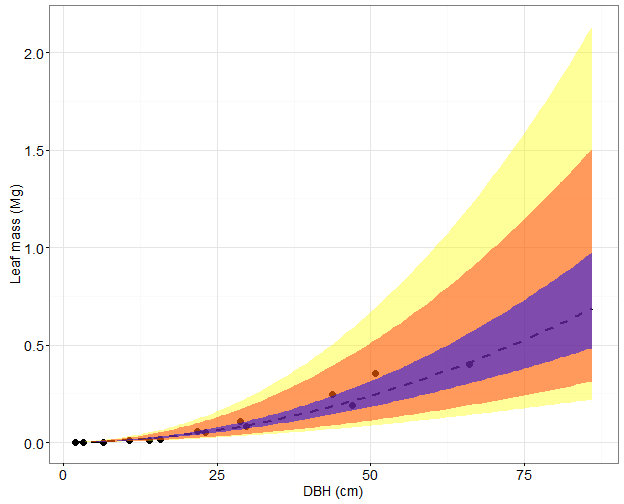
Now we have all the necessary information to get the:

Error in prediction of the mean ():

(3)

Error in prediction of an individual (:

(4)



We now have the y ; so now we only need to know:

which is the average DBH of the sample used to develop the model

which is the sum of the squared deviations of x.

It is worth mentioning, that to obtain we simply need to obtain the average of the DBH from column "B" and this value is stored in the cell B10.

Finally, to obtain the sum of squared deviations we proceed to create a new column “H”, in which we obtain the square from the differences of each DBH and the medium .The sum of squared deviations of x is stored in the cell "B11"

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **A** | **B** | **C** | **D** | **E** | **F** | **G** | **H** |
| **Id\_tree** | **DBH** | **Biomass\_kg** | **ln(Biomass\_kg)** | **ln(DBH)** | **Estimated Biomass  (Yi\_est)** | **(Yi-Yi\_est)^2** | **(Xo-Xmean)^2** |
| 1 | 19.09 | 68.0 | 4.2 | 2.95 | 4.43 | 0.04439 | 2.64 |
| 2 | 31 | 179.4 | 5.2 | 3.43 | 5.23 | 0.00195 | 1.30 |
| 3 | 21.64 | 139.2 | 4.9 | 3.07 | 4.64 | 0.08871 | 2.25 |
| 4 | 40.87 | 369.4 | 5.9 | 3.71 | 5.69 | 0.04814 | 0.75 |
| 5 | 28.17 | 213.3 | 5.4 | 3.34 | 5.08 | 0.08261 | 1.53 |
| 6 | 27.11 | 224.1 | 5.4 | 3.30 | 5.01 | 0.16060 | 1.63 |
| 7 | 17.18 | 69.6 | 4.2 | 2.84 | 4.25 | 0.00015 | 3.00 |
| 8 | 15.27 | 46.6 | 3.8 | 2.73 | 4.06 | 0.04791 | 3.42 |
| 9 | 28.32 | 152.0 | 5.0 | 3.34 | 5.08 | 0.00361 | 1.52 |
| 10 | 34.37 | 240.8 | 5.5 | 3.54 | 5.41 | 0.00620 | 1.08 |
| 11 | 15.08 | 46.6 | 3.8 | 2.71 | 4.04 | 0.03919 | 3.46 |
| 12 | 17.5 | 153.4 | 5.0 | 2.86 | 4.29 | 0.55857 | 2.93 |
| 13 | 4.2 | 2.0 | 0.7 | 1.44 | 1.92 | 1.56034 | 9.86 |
| 14 | 4.13 | 5.8 | 1.8 | 1.42 | 1.89 | 0.01844 | 9.96 |
| 15 | 28.01 | 120.0 | 4.8 | 3.33 | 5.07 | 0.07753 | 1.54 |
| 16 | 49.33 | 608.1 | 6.4 | 3.90 | 6.00 | 0.16472 | 0.46 |
| 17 | 3.18 | 11.7 | 2.5 | 1.16 | 1.46 | 1.00490 | 11.68 |
| 18 | 25.46 | 160.9 | 5.1 | 3.24 | 4.91 | 0.03001 | 1.79 |
| 19 | 51.24 | 1141.7 | 7.0 | 3.94 | 6.07 | 0.94627 | 0.41 |
| 20 | 14.96 | 78.7 | 4.4 | 2.71 | 4.03 | 0.11566 | 3.49 |
| 21 | 23.87 | 148.0 | 5.0 | 3.17 | 4.80 | 0.03880 | 1.97 |
| 22 | 7.32 | 8.6 | 2.2 | 1.99 | 2.84 | 0.47126 | 6.68 |
| 23 | 25.94 | 246.8 | 5.5 | 3.26 | 4.94 | 0.32492 | 1.74 |
| 24 | 45.6 | 262.6 | 5.6 | 3.82 | 5.87 | 0.09210 | 0.57 |
| 25 | 46 | 598.1 | 6.4 | 3.83 | 5.89 | 0.25512 | 0.56 |
| 26 | 49 | 443.2 | 6.1 | 3.89 | 5.99 | 0.01012 | 0.47 |
| 27 | 50.5 | 356.4 | 5.9 | 3.92 | 6.04 | 0.02803 | 0.43 |
| 28 | 65.6 | 256.5 | 5.5 | 4.18 | 6.48 | 0.86548 | 0.15 |
| 29 | 49.0 | 225.4 | 5.4 | 3.89 | 5.99 | 0.33109 | 0.47 |
| 30 | 41.0 | 277.6 | 5.6 | 3.71 | 5.70 | 0.00509 | 0.74 |
| 31 | 61.0 | 198.9 | 5.3 | 4.11 | 6.36 | 1.13182 | 0.22 |
| 32 | 64.0 | 576.6 | 6.4 | 4.16 | 6.44 | 0.00628 | 0.17 |
| 33 | 87.0 | 1253.4 | 7.1 | 4.47 | 6.95 | 0.03534 | 0.01 |
| 34 | 71.0 | 632.2 | 6.4 | 4.26 | 6.61 | 0.02539 | 0.10 |
| 35 | 79.0 | 1057.7 | 7.0 | 4.37 | 6.79 | 0.03175 | 0.04 |
| 36 | 87.0 | 1068.4 | 7.0 | 4.47 | 6.95 | 0.00080 | 0.01 |
| 37 | 97.0 | 1205.9 | 7.1 | 4.57 | 7.13 | 0.00097 | 0.00 |
| 38 | 93.0 | 1387.3 | 7.2 | 4.53 | 7.06 | 0.03200 | 0.00 |
| 39 | 87.0 | 909.9 | 6.8 | 4.47 | 6.95 | 0.01752 | 0.01 |
| 40 | 67.0 | 528.2 | 6.3 | 4.20 | 6.51 | 0.05900 | 0.14 |
| 41 | 112.0 | 992.6 | 6.9 | 4.72 | 7.36 | 0.21555 | 0.02 |
| 42 | 129.0 | 2219.1 | 7.7 | 4.86 | 7.60 | 0.01120 | 0.08 |
| 43 | 144.0 | 4106.2 | 8.3 | 4.97 | 7.78 | 0.29026 | 0.16 |
| 44 | 160.0 | 3480.6 | 8.2 | 5.08 | 7.96 | 0.03948 | 0.25 |
| 45 | 160.0 | 1554.3 | 7.3 | 5.08 | 7.96 | 0.36903 | 0.25 |
| 46 | 130.0 | 1821.3 | 7.5 | 4.87 | 7.61 | 0.01093 | 0.09 |
| 47 | 150.0 | 2443.8 | 7.8 | 5.01 | 7.85 | 0.00230 | 0.19 |
| 48 | 130.0 | 2898.4 | 8.0 | 4.87 | 7.61 | 0.12966 | 0.09 |
|  | 58.08 |  |  | 3.70 |  | 9.83116 | 80.28 |

## Process to propagate uncertainties of the measurement errors and the allometric models.

Once we have the error in measurement and the model parameters, these are sent to call on the tab "Biomass calculation" to propagate the uncertainties of measurement error and the allometric models.

This way, in the cell A7 we find the measurement error that we have just calculated, in the cells A13 and A18 the statistics of the adjusted model are indicated.

In one way, we create the switched shown in the blue cells with which source of uncertainty can be turned on or off by changing the numbers in the blue boxes to " 1 " or " 0". This means, if the cell "D7" is deactivated placing a 0, then the uncertainty of the measurement error will not be propagated. The same will happen with cells D19 and D23 with the: “Error in prediction of an individual ( PI)” and “Error in prediction of the mean (CI )” respectively.



So then, with all this information, it is possible to estimate the errors "PI" and "CI", however, to get this, first it is necessary to create the column "C" where the DBH with the error in measurement is included. Afterwards, this variable is transformed applying the natural logarithm (as we can see in column “D”). so now we can apply the formulas for the estimation of the errors of PI and CI.

|  |  |  |  |
| --- | --- | --- | --- |
| **A** | **B** | **C** | **D** |
| Plot | Tree diameter | DBH with measurement error | ln(DBH) |
| 1.00 | 24.90 | 24.9 | 3.2 |
| 1.00 | 9.20 | 10.0 | 2.3 |
| 1.00 | 11.00 | 11.0 | 2.4 |
| 1.00 | 11.70 | 11.1 | 2.4 |
| 1.00 | 23.30 | 23.3 | 3.1 |

This way, the column "E" we will get the error in prediction of the mean, so in this cell we will generate a random number of normal distribution with medium 0 and variance.

To generate a random number of normal distribution, Excel has a function called "DISTR.NORM.INV" which needs as arguments: The probability, the medium and standard deviation. The probability is obtained generating a random value between 0 and 1; the medium is given values, so then its only necessary to obtain the standard deviation . To get it, we proceed to program the expression , for which, as a variance we take the ECM and the size of the sample which the model was built with. X0 is the DBH for each tree in the inventory. is the mean of DBH of the sample used to construct the model and is the sum of the squared deviations of x.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **A** | **B** | **C** | **D** | **E** |
| Plot | Tree diameter | DBH with measurement error | ln(DBH) | Error in prediction of the mean |
| 1.00 | 24.90 | 24.9 | 3.2 | 0.0 |
| 1.00 | 9.20 | 9.2 | 2.2 | 0.1 |
| 1.00 | 11.00 | 11.8 | 2.5 | 0.1 |
| 1.00 | 11.70 | 11.3 | 2.4 | 0.2 |
| 1.00 | 23.30 | 23.4 | 3.2 | 0.0 |

To obtain the “Error in prediction of an individual” we proceed from an analog manner shown in the cell "E" with the difference is that in the formula of the standard deviation we will add a 1 like it shown in the formula. This source of error is programing in column F.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **A** | **B** | **C** | **D** | **E** | **F** |
| Plot | Tree diameter | DBH with measurement error | ln(DBH) | Error in prediction of the mean | Error in prediction of an indivual |
| 1.00 | 24.90 | 24.9 | 3.2 | 0.0 | 0.2 |
| 1.00 | 9.20 | 9.2 | 2.2 | 0.1 | 0.0 |
| 1.00 | 11.00 | 11.8 | 2.5 | 0.1 | 0.1 |
| 1.00 | 11.70 | 11.3 | 2.4 | 0.2 | -0.6 |
| 1.00 | 23.30 | 23.4 | 3.2 | 0.0 | 0.4 |

Once both sources of error are estimated, in column “G” we proceed to estimate the natural logarithm of the biomass including the 2 sources of error and conditioning them to the switched indicated in the cells D19 and D23. To include the variation associated to measurement error, we evaluate the allometric model in the DBH variable that has already included this source of error. This way, three sources of error are included in the estimation of biomass and each source of error can be can be turned on or off by changing the numbers in the blue boxes to " 1 " or " 0".

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **A** | **B** | **C** | **D** | **E** | **F** | **G** |
| Plot | Tree diameter | DBH with measurement error | ln(DBH) | Error in prediction of the mean | Error in prediction of an indivual | ln(biomass1\_2) |
| 1.00 | 24.90 | 24.9 | 3.2 | 0.0 | 0.2 | 5.1 |
| 1.00 | 9.20 | 9.2 | 2.2 | 0.1 | 0.0 | 3.3 |
| 1.00 | 11.00 | 11.8 | 2.5 | 0.1 | 0.1 | 3.9 |
| 1.00 | 11.70 | 11.3 | 2.4 | 0.2 | -0.6 | 3.2 |
| 1.00 | 23.30 | 23.4 | 3.2 | 0.0 | 0.4 | 5.2 |

Finally, due to all these estimations have been obtained in the logarithmic scale, now it is only necessary to transform to the real scale. To do that, in column “H” we apply the exponent to the ln(biomass). This way, three sources of error are included in the estimation of the biomass.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **A** | **B** | **C** | **D** | **E** | **F** | **G** | **H** |
| Plot | Tree diameter | DBH with measurement error | ln(DBH) | Error in prediction of the mean | Error in prediction of an indivual | ln(biomass1\_2) | biomass (kg) |
| 1.00 | 24.90 | 24.9 | 3.2 | 0.0 | 0.2 | 5.1 | 157.7 |
| 1.00 | 9.20 | 9.2 | 2.2 | 0.1 | 0.0 | 3.3 | 26.9 |
| 1.00 | 11.00 | 11.8 | 2.5 | 0.1 | 0.1 | 3.9 | 47.1 |
| 1.00 | 11.70 | 11.3 | 2.4 | 0.2 | -0.6 | 3.2 | 23.5 |
| 1.00 | 23.30 | 23.4 | 3.2 | 0.0 | 0.4 | 5.2 | 180.0 |

Changing the switches, it is possible then to propagate the uncertainties of the different sources and in an automatic way and we will see the effect on the estimation of biomass.

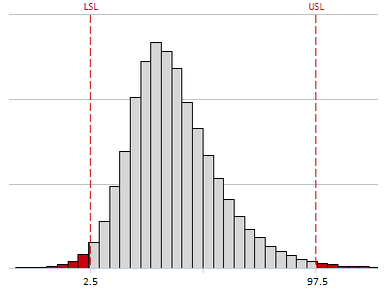
These estimations are a first realization, so to propagates uncertainties at plot level its necessary to develop thousands of simulations with which we could obtain confidence intervals including both the CI and the PI. While this is possible in Excel, this process is very slow, so this exercise will be made in R.

In the meantime, to exemplify the propagation of uncertainties at a plot level, in page “Iterations” is showed a general process to follow.

## Propagation of uncertainty at a plot level.

First, in page “Biomass calculation” we will copy cells B37 to B46, and we will paste them on row 3 of page “Iterations”. The values copy are the densities of the biomass at a plot level in terms of ton/ha and after that, these values are being related with the previous slide and will be updated in an automatic manner so one can copy and paste them as values in each of simulation. Each simulation will be copy from raw 9 in ahead.

After doing these simulation, we proceed to obtain the confidence intervals of these simulations and this is realized obtaining percentiles 2.5% and 97.5% To express confidence intervals as percentage, we proceed as follow:



# Process to propagate uncertainties in R

The process shown before was programed in R (a statistical software) because these tool is very powerful to do simulations. So, taking into account the procedure shown before, now is explained the steps to propagate uncertainties in R:

1. First, it is necessary to load library “doBy”.
2. Then it is necessary to indicate the address folder
3. In “Module 2” the parameters model are read
4. In “Module 3” are estimated the measurement error, Error in prediction of the mean, Error in prediction of an induvial and uncertainties are propagated at plot level. To do that, script in R do the follow:
   1. The number of simulations is defined
   2. Tree diameter with measurement error is estimated
   3. Logarithm of DBH (included measurement error) is obtained
   4. Error in prediction of the mean is estimated
   5. Error in prediction of an individual is estimated
   6. Logarithm of "Biomass estimation with sources of error" is estimated
   7. The logarithm of estimated biomass is transformed to its real scale
   8. Estimations of uncertainties at plot level are obtained.
5. Finally, results are shown and saved.