Tutorial of:
“Calc” from “Open Foris”

Jordi Vayreda i Lluís Comas
Free open-source solutions for environmental monitoring

Collect
- Easy and flexible survey design and data management

Collect Mobile
- Intuitive data collection and validation in the field

Collect Earth
- Innovative land assessment through freely available satellite imagery

Calc
- Efficient and collaborative data analysis and results dissemination

Geospatial Toolkit
- Powerful command-line utilities for processing geospatial data
Introduction to Calc

Open Foris Calc is a robust tool for data analysis and results calculation. The input data and metadata come from Open Foris Collect and Calc provides a flexible way to produce aggregated results which can be analyzed and visualized through the open source software Saiku. Calc allows expert users to write custom R modules to perform calculations working with a variety of sampling designs.
Calc is open source. It's hosted, developed, and maintained on GitHub.

Tutorials

1. Installation
2. Calc Home
3. Settings
4. Data
5. Calculation
6. Sanku
7. Vietnam - Use case for Calc
Installation

CALC can be installed on your computer through a CALC installer. CALC operates in synergy with other supporting software which, if not already present, should be installed. The supporting software is:

- Java Development Kit (JDK)
- Postgres: an object-relational database management system; as a database server, its primary function is to store data and retrieve it later, as requested by other software applications.
- R: a free software programming language and software environment for statistical computing and graphics.

The installation process can be carried out following these steps:

1. Download and install JDK 1.8+ [here](http://www.oracle.com/technetwork/java/javase/downloads/jdk8-downloads-2133151.html)
   Note: the required password should be set as 'postgres'.
3. Download and install R 3.1+ [here](http://cran.r-project.org/).  
   Note: for Windows users, R should be installed in a folder that contains no spaces (""), for example "C:\Opt\R\R-3.1.0"
4. Download and install OpenForis Calc  
   Note: for Windows users, CALC should be installed in a folder that contains no spaces (""), for example "C:\opt\OpenForisCalc"

   - OpenForisCalc Installer for Windows
   - OpenForisCalc Installer for Linux

Step number 4 will download the following executable file:

   `OpenForisCalc-(VERSION)-(PLATFORM)-installer.exe`

Double click to launch the CALC setup Wizard

Note:

- Administration rights are required.
- Both Windows and Linux users will be required to re-start the computer in order to update the required system settings.

- **Calc setup Wizard**
- **Calc Control Panel**

Click [here](#) to return to Tutorials Main Menu
Descargar base de datos

A test survey named Atlantis is available for download here:

- Click to download the test data-set Atlantis.

The folder test-data contains two surveys, one set up as a one-phase sampling, the other with a double sampling design.

This tutorial is based on the double sampling example which is the one that requires more settig up steps.

- Open the folder atlantis-2phases-sampling
- Download collect-backup-atlantis.zip which contains the data and metadata of the survey created in Collect.
- Open the folder calc

The files contained in the folder calc allow to experiment with two different types of set-up: automatic or manual.

- **Automatic** set-up: download atlantis-calc-workspace.zip which will be uploaded into Calc and already contains all the necessary information (files .csv).
- **Manual**: download the following csv files
  - calc-aois.csv: containing information on the "areas of interest".
  - calc-phase1-plots.csv: containing phase-1 information on the strata, cluster identification number, plot and a code indicating the area of interest.
  - calc-strata.csv: containing information on the number of strata included in the survey.
  - calc-volume-models.csv: containing the volume equations that will be used in the analysis.
Calc is open source. It's hosted, developed, and maintained on GitHub.

Tutorials

3. Settings

1. Installation
2. Calc Home
4. Data
5. Calculation
6. Seiku
7. Vietnam - Use case for Calc
Calc, menú principal
Crear un nou “workspace”
Afegir un nou “workspace”
Donar-li un nom i guardar
Carregar dades
Seleccionar el fitxer zip que conté la base de dades
Running

1. Import metadata 100%
2. Create database 100%
3. Import species 100%
4. Import data 100%

Log

Close
Estructura de l’inventari forestal
Seleccionar l’àrea d’interès
Areas of interest

The next step is to upload the information about the Areas of interest. Click on Areas of interest, then on Upload Csv and select the appropriate file. In the test example: calc-aois.csv. This file contains data on the area of reference associated to the survey: it could be the total area of a country and its subdivisions in regions, provinces etc. We can consider ‘areas of interest’ as reporting units.

Test files box

calc-aois.csv - this file contains information on the “areas of interest”. Hierarchical levels are indicated as consecutive numbers (level_1; level_2; etc.) each one with an identification code and a label. In this example, the total area “Atlantis” is divided into two sub areas: “Timaeus” and “Critias”, each one with an area magnitude indication.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>level_1_code</td>
<td>level_1_label</td>
<td>level_2_code</td>
<td>level_2_label</td>
<td>level_2_area</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>Atlantis</td>
<td>1</td>
<td>Timaeus</td>
<td>650000</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>Atlantis</td>
<td>2</td>
<td>Critias</td>
<td>350000</td>
</tr>
</tbody>
</table>
calc-aois_menfri.csv
Running

1. Calculate sampling unit weight  
   Progress: 100%

2. Assign area of interest columns  
   Progress: 100%

3. Calculate expansion factor  
   Progress: 100%

Log

Close
Carregar equacions de cubicació
External equations

The last step allows the user to upload **External equations** for the calculation of volume according to tree species or any other condition. This is done by uploading a csv file.

Click on **External equations**, then on **Upload Csv**. In the test example, the file to upload is: **calc-volume-models.csv**

---

**Test files box**

**calc-volume-models.csv** - contains the volume equations that will be used in the analysis. Required fields include: *code* (indicating the species), the actual volume equations, and, where needed, a condition. The last equation listed acts a default equation, to be applied where none of the previous conditions apply.

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>code</td>
<td>equation</td>
</tr>
<tr>
<td>2</td>
<td>EUC/GRA 0.000065 * dbh^1.633 * h^1.137</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>PIN/PAT 0.00002117 * dbh^1.8644 * h^1.3246</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>TCT/GRA 0.0001 * dbh^1.91 * h^0.75</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>DAL/MEL 0.00023 * dbh^2.231</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.0001 * dbh^2.032 * h^0.66 round(as.numeric(vegetation_type)/100) == 2</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.5 * pi * (0.01 * dbh / 2)^2 * h</td>
<td></td>
</tr>
</tbody>
</table>
Donem un nom i guardem

Name: volume

Import
<table>
<thead>
<tr>
<th>Code</th>
<th>Equation</th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>EUC/GRA</td>
<td>0.000065 * dbh^1.633 * h^1.137</td>
<td></td>
</tr>
<tr>
<td>PIN/PAT</td>
<td>0.00002117 * dbh^1.8644 * h^1.3246</td>
<td></td>
</tr>
<tr>
<td>TCT/GRA</td>
<td>0.0001 * dbh^1.91 * h^0.75</td>
<td></td>
</tr>
<tr>
<td>DAL/MEL</td>
<td>0.00023 * dbh^2.231</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0001 * dbh^2.032 * h^0.66</td>
<td>round(as.numeric(vegetation_type)/100) -- 2</td>
</tr>
<tr>
<td></td>
<td>0.5 * pi * (0.01 * dbh / 2)^2 * h</td>
<td></td>
</tr>
</tbody>
</table>
Inventory

In order to produce aggregated results based on a sampling design, one must be defined. Calc handles point sampling surveys based on a variety of sampling designs: single or double sampling, cluster, random, systematic and stratified sampling.

The hierarchical structure defined in Collect is converted into a relational database: the entities are converted into tables, attributes into columns, and tables are linked by logical joins.

A representation of the data tables and their relational joins under different sampling design is presented below. [Click on Data tables and relational joins to expand it].

Sampling Design

The sampling design section has a question/answer user interface. Choices are made by clicking the appropriate button which, by turning green, indicates that a selection has been made.

The following steps will guide you to define the sampling design in Calc.

Click on Sampling Design and then on the Edit button.
Double sampling, Cluster, Stratified

1. Sampling unit table
   - Column A: cluster
   - Column B: plot
   - Column C: aoi_code

2. Phase-1 table
   - Column A: stratum
   - Column B: cluster
   - Column C: plot
   - Column D: aoi_code

3. Stratum table
   - Column A: stratum
   - Column B: stratum_code

4. Areas of interest Table
   - Column A: level_1_code
   - Column B: level_1_label
   - Column C: level_2_code
   - Column D: level_2_label
   - Column E: level_2_area

<table>
<thead>
<tr>
<th>level_1_code</th>
<th>level_1_label</th>
<th>level_2_code</th>
<th>level_2_label</th>
<th>level_2_area</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Atlantis</td>
<td>1</td>
<td>Timaeus</td>
<td>650000</td>
</tr>
<tr>
<td>0</td>
<td>Atlantis</td>
<td>2</td>
<td>Critias</td>
<td>350000</td>
</tr>
</tbody>
</table>
Disseny del mostreig
Sampling Design

The sampling design section has a question/answer user interface. Choices are made by clicking the appropriate button which, by turning green, indicates that a selection has been made.

The following steps will guide you to define the sampling design in Calc:

Click on **Sampling Design** and then on the **Edit** button.

1. Select the entity that represents the sampling unit (table 1). For the test example select plot.

Click on the right-arrow to proceed to the next step.
3. Select whether the survey is designed with double sampling (2 phases). For the test example select **2-phases**.
calc-phase1-plots.csv contains information on the strata (in this case 3), cluster identification number, plot (in this case ranging from 1 to 10) and a code indicating the area of interest (in this case 1 or 2). See image below.

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
</tbody>
</table>

Example data in the file:
<table>
<thead>
<tr>
<th>Column</th>
<th>Import</th>
<th>Data type</th>
</tr>
</thead>
<tbody>
<tr>
<td>stratum</td>
<td></td>
<td>Integer</td>
</tr>
<tr>
<td>cluster</td>
<td></td>
<td>String</td>
</tr>
<tr>
<td>plot</td>
<td></td>
<td>String</td>
</tr>
<tr>
<td>aoi_code</td>
<td></td>
<td>String</td>
</tr>
</tbody>
</table>

Import button
Running
1. Import csv into database table 100%

Log

plot

String

aoi_code

String

Import

Close
Especificar les relacions
Carregar el fitxers amb els estrats de vegetació
Especificar el camp “estrat”
Especificar si és un disseny en cluster

6. Si la encuesta tiene un diseño de clúster, seleccione Cluster.

El usuario debe definir la columna que representa el código de clúster [en caso de muestreo doble, la columna debe estar presente en la tabla de fase 1 (Tabla 2), de lo contrario en la tabla de unidad de muestreo (Tabla 1)].

Para el ejemplo de prueba, seleccione cluster. (vea la columna de cluster en la sección de Data tables and relational joins).
Definir l’àrea d’interès

7. Define the area of interest. The system requires the users to indicate the column that represents the lowest level of the administrative unit hierarchy previously imported (table 4). [In case of double sampling the column has to be present in the phase 1 table (table 2), otherwise in the sampling unit table (table 1)].

For the test example select aoi_code. (see the join between table 2 and table 4 in the Data tables and relational joins section)

Then move to the next step.
Especificar el camp que especifica l’àrea d’interès
8. The last step requires the user to write an R script to calculate the **weight** of each record of the sampling unit table. The script will assign a weight to each record (sampling unit) by adding a column (named weight) to the sampling unit table (table 1).

For the test example the following R script will be entered:

```r
plot$weight = ifelse ( is.na ( plot$subplot ) | plot$subplot == 'A' , 1 , 0 )
```
I finalment el resultat del disseny del mostreig
Agregació de la informació

Next step is to define the **Aggregation** function which represents the formula of the plot area for the entities you wish to aggregate following the sampling design. (e.g. trees / dead wood etc.)

Click on **Aggregation** to select the Entity. For the test example select **Tree**.
Ex: disseny de mostreig amb parcel·les circulars aniuades

tree$plot_radius <- with(tree, 
    ifelse(dbh < 5, 1, 
        ifelse(dbh < 10, 5, 
            ifelse(dbh < 20, 10, 15) 
        )
    )
);

tree$plot_area <- with(tree, pi * plot_radius^2 * share / 100);
#convert plot area factor from m2 to ha

tree$plot_area <- tree$plot_area * 0.0001;
Pantalla inicial per visualitzar les dades
Filtrar per variables contínuas
Filtre per variables categòriques
Filtrar i exportar a csv

Table view of attributes dbh and total_height with filter set on total_height = IS NOT NULL.
Fer gràfiques
Càlculs derivats d’informació de camp

Calculation

CALC allows to build calculation steps in three ways:

- Writing a script in R
- Using external equations previously added in the Settings section
- Introducing a new Categorical variable

To create a new Calculation step click on the “plus” icon.

Then select the type of calculation you wish to create by clicking one of the three options:

- R script
- External equation
- Category
Càlcul via script de R

Calculation writing a script in R

The following is an example of how to create a calculation step using R script to assign a height value to each tree record according to a linear fit model.

Note: useful sources for learning about R: Quick-R, Codeschool - Try R
Estimar els valors d’altura fent un model estadístic amb dades de camp

```r
sample_trees <- tree[ !is.na( tree$total_height ) , ]
sample_trees <- sample_trees[ sample_trees$total_height > 0 , ];
height_model <- with( sample_trees, lm(total_height ~ dbh + I(dbh ^2) + I(dbh ^3)) );
tree$est_height <- predict( height_model, newdata = tree[ , c('dbh','total_height')] );
tree$est_height <- ifelse( is.na(tree$total_height), tree$est_height, tree$total_height) ;
```
Al tutorial hi ha més exemples

**Examples of R scripts**
The following are some examples of R scripts used to perform Calculations

---

**Calculation name:** Tree - Basal area  
**Calculation type:** R script  
**Script:**
```r
tree$basal_area <- with(tree, pi * (0.01*dbh/2)^2)
```

---

**Calculation name:** Tree - Volume  
**Calculation type:** R script  
**Script:**
```r
# Basic form factor volume model
ff <- 0.515;
tree$volume <- with( tree, (0.1291+1.5984 * ff) * pi * (0.01 * dbh / 2)^2 * est_height^0.764 );
```

---

**Calculation name:** Stand - IPCC class  
**Calculation type:** Category  
**Script:**
```r
# '1' Forest land, '2' Grass land, '3' Cropland, '4' Settlements, '5' Wetland, '-1' NA
stand$ipcc_class = with ( stand,  
  ifelse(forest_status ==160 | forest_status == 630, 5,  
  ifelse(forest_status < 440, 1,  
  ifelse(forest_status == 440, 2,  
  ifelse(forest_status < 600, 3, 4))))
```
Calculation name: Tree - AG Biomass
Calculation type: R script
Script:
Tree - AG Biomass
BEF_pinus <- 1.3;
tree$genus_code <- substr( tree$species_code, 1, 3 ); # compute AGB in kg
tree$aboveground_biomass <- with ( tree,
  ifelse( genus_code=="PIN", BEF_pinus * volume * 500, # Pinus, wood density = 500 kg/m3
    269.63396 * (((dbh/100)^2*est_height)^0.95193) # Evergreen forest
  )
)
# convert kg -> tons
tree$aboveground_biomass <- tree$aboveground_biomass / 1000

Calculation name: Tree - BG Biomass
Calculation type: R script
Script:
# conversion factor source:
tree$belowground_biomass <- tree$aboveground_biomass * 0.265;
Calculation name: Tree - Total biomass
Calculation type: R script
Script:
```r
tree$total_biomass <- tree$aboveground_biomass + tree$belowground_biomass
```

Calculation name: Stump - Count
Calculation type: R script
Script:
```r
stump$quantity[is.na(stump$quantity)] <- 1
stump$quantity[stump$quantity =="" ] <- 1
stump$count_stump <- stump$quantity
```
Càlcul usant equacions externes

Calculation using external equations

The following is an example of calculating tree volume using a list of volume models that was previously uploaded into Calc during the Settings phase.
<table>
<thead>
<tr>
<th>Type</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caption</td>
<td>Tree volume</td>
</tr>
<tr>
<td>Entity</td>
<td>tree</td>
</tr>
<tr>
<td>Variable</td>
<td>volume</td>
</tr>
<tr>
<td>Equation list</td>
<td>volume</td>
</tr>
<tr>
<td>Code variable</td>
<td>species_code</td>
</tr>
<tr>
<td>Variable 'vegetation_type'</td>
<td>vegetation_type</td>
</tr>
<tr>
<td>Variable 'dbh'</td>
<td>dbh</td>
</tr>
<tr>
<td>Variable 'h'</td>
<td>est_height</td>
</tr>
</tbody>
</table>

Aggregate function:
- sum
- min
- max
- avg
- count
- distinct-count

Available only for those entities that won't be aggregated using the sampling design.
Sum will be applied by default if none of the functions is selected.

Save
Afegir una nova variable categòrica: classe diamètrica

[In the test example we wish to create four dbh classes with codes ranging from 1 to 4 and dbh classes of <10, 10-20, 20-30, 30+.}
NA,'1' <10,'2' <20,'3' <30,'4' <40,'5' >=40

tree$dbh_class <- ifelse( tree$dbh < 10 , 1 ,
    ifelse( tree$dbh < 20 , 1 ,
        ifelse( tree$dbh < 30 ,
            ifelse( tree$dbh < 40 , 4 , 5 )
            )
    )
);

Veure el resultat del càlcul

<table>
<thead>
<tr>
<th>Row #</th>
<th>dbh</th>
<th>dbh_class</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.700</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1.400</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1.200</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1.700</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>1.300</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>6.900</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>8.200</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>7.300</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>19.70</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td>41.50</td>
<td>5</td>
</tr>
<tr>
<td>12</td>
<td>26.50</td>
<td>3</td>
</tr>
<tr>
<td>13</td>
<td>21.50</td>
<td>3</td>
</tr>
<tr>
<td>14</td>
<td>13.70</td>
<td>2</td>
</tr>
<tr>
<td>15</td>
<td>33.50</td>
<td>4</td>
</tr>
<tr>
<td>16</td>
<td>26.60</td>
<td>3</td>
</tr>
<tr>
<td>17</td>
<td>22</td>
<td>3</td>
</tr>
<tr>
<td>18</td>
<td>34</td>
<td>4</td>
</tr>
<tr>
<td>19</td>
<td>2.200</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>
Saiku

• Saiku és un sistema d’analisi de la informació que permet als usuaris de forma ràpida i senzilla analitzar les dades de manera còmode i senzilla i crear i compartir informes ràpidament.

• Hi ha un manual de l’usuari a:

1- executar totes les funcions de càlcul
Executar Saiku
Usuari=calc, password=calc
Menús i funcionalitats

3.2. Toolbars

[Diagram with toolbars and functionalities described in text]

Mode:

Table  Chart
<table>
<thead>
<tr>
<th>Measures Level</th>
<th>Secondary forest</th>
<th>Plantation</th>
<th>Bushes, grassland</th>
<th>Agricultural land</th>
<th>Other land</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basal_area (Mean)</td>
<td>12.2262</td>
<td>7.0539</td>
<td>0.5411</td>
<td>1.3579</td>
<td>0.6843</td>
</tr>
</tbody>
</table>
Mostrar fileres buides (sense informació)

Showing of empty / non-empty records, click ‘Non-empty’ button.
Filtre informació de sortida

2) Filter, Limit, ...

3) Filter, Sort, ...

Agricultural land | 1 | 12.2282 | 718.4783 | 78.9456
Bushes, grassland | 0 | 718.4783 | 78.9456
Other land | 0 | 78.9456
Plantation | 7 | 78.9456
Secondary forest | | | |
Ordenar per una variable
Mostrar estadística bàsica

3.8. Statistics

Saiku allows also showing some statistics about columns values.

Click the summary button.
<table>
<thead>
<tr>
<th>Ownership</th>
<th>Measured level</th>
<th>Production forest</th>
<th>Protection forest</th>
<th>Wildlife reserve</th>
<th>Shifting cultivation</th>
<th>Agriculture</th>
<th>Grazing land</th>
<th>Other land</th>
</tr>
</thead>
<tbody>
<tr>
<td>NIA</td>
<td>Area</td>
<td>1,111,111</td>
<td>1,111,111</td>
<td>220,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Central government land</td>
<td>Area</td>
<td>47,777,777</td>
<td>144,444,444</td>
<td>58,888,888</td>
<td>1,111,111</td>
<td>10,000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local government land</td>
<td>Area</td>
<td>27,777,777</td>
<td>22,222,222</td>
<td>12,222,222</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Village land</td>
<td>Area</td>
<td>324,444,444</td>
<td>58,888,888</td>
<td>1,111,111</td>
<td>15,888,888</td>
<td>20,000</td>
<td>11,111,111</td>
<td></td>
</tr>
<tr>
<td>Private land</td>
<td>Area</td>
<td>8,889,889</td>
<td>3,333,333</td>
<td>7,777,777</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General land</td>
<td>Area</td>
<td>50,000</td>
<td>5,555,555</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>