

Estimation of area for REDD+ using stratified sampling design: application with SEPAL tools

Webinar GOFC-GOLD, 2017/06/06

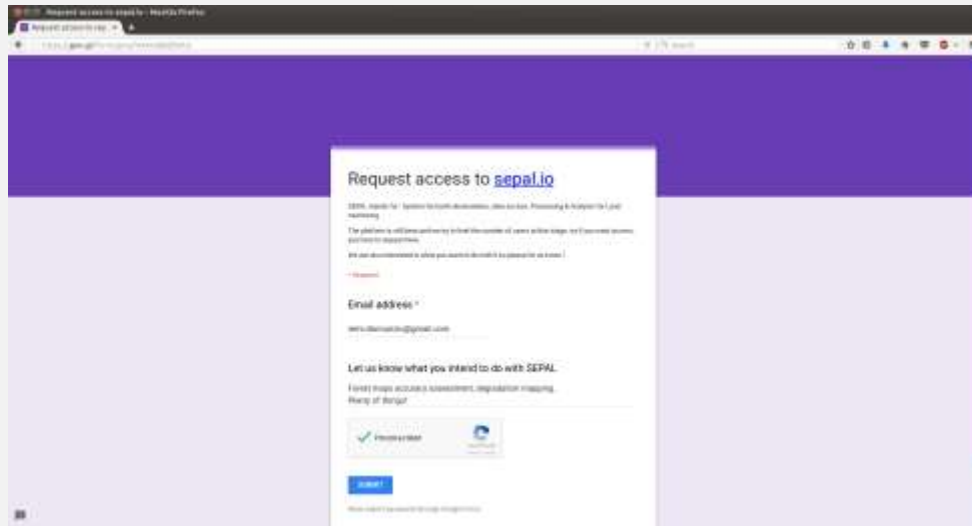
remi.dannunzio@fao.org



Request access to SEPAL

SEPAL stands for : System for Earth observations, data access, Processing & Analysis for Land monitoring.

The platform is still beta and we try to limit the number of users at that stage



The screenshot shows a Google Survey form titled "Request access to sepal.io". The form is displayed on a purple background. The text on the form includes: "SEPAL stands for : System for Earth observations, data access, Processing & Analysis for Land monitoring.", "The platform is still beta and we try to limit the number of users at that stage, we therefore require you to request access.", "We are disappointed to hear you cannot access it, please let us know!", "Email address *", "Let us know what you intend to do with SEPAL.", "I intend to use SEPAL for: accuracy assessment, map validation, mapping, threat of deforestation.", There is a "Submit" button at the bottom.

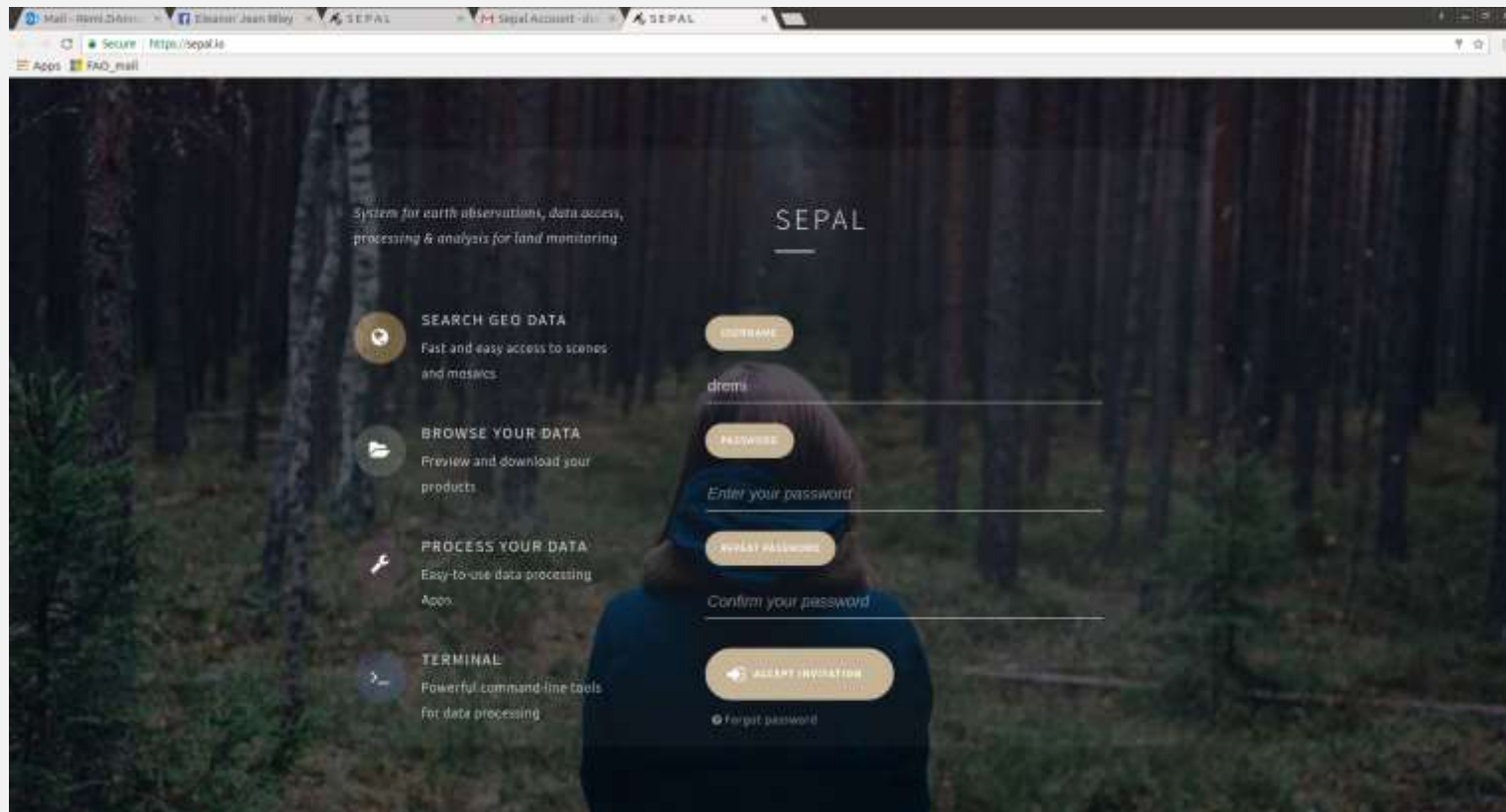
If you want access, you have to request [in this Google Survey](#)

You will then receive an email with a link to activate your access.

Remember to check in your **SPAMS**, it might be there.

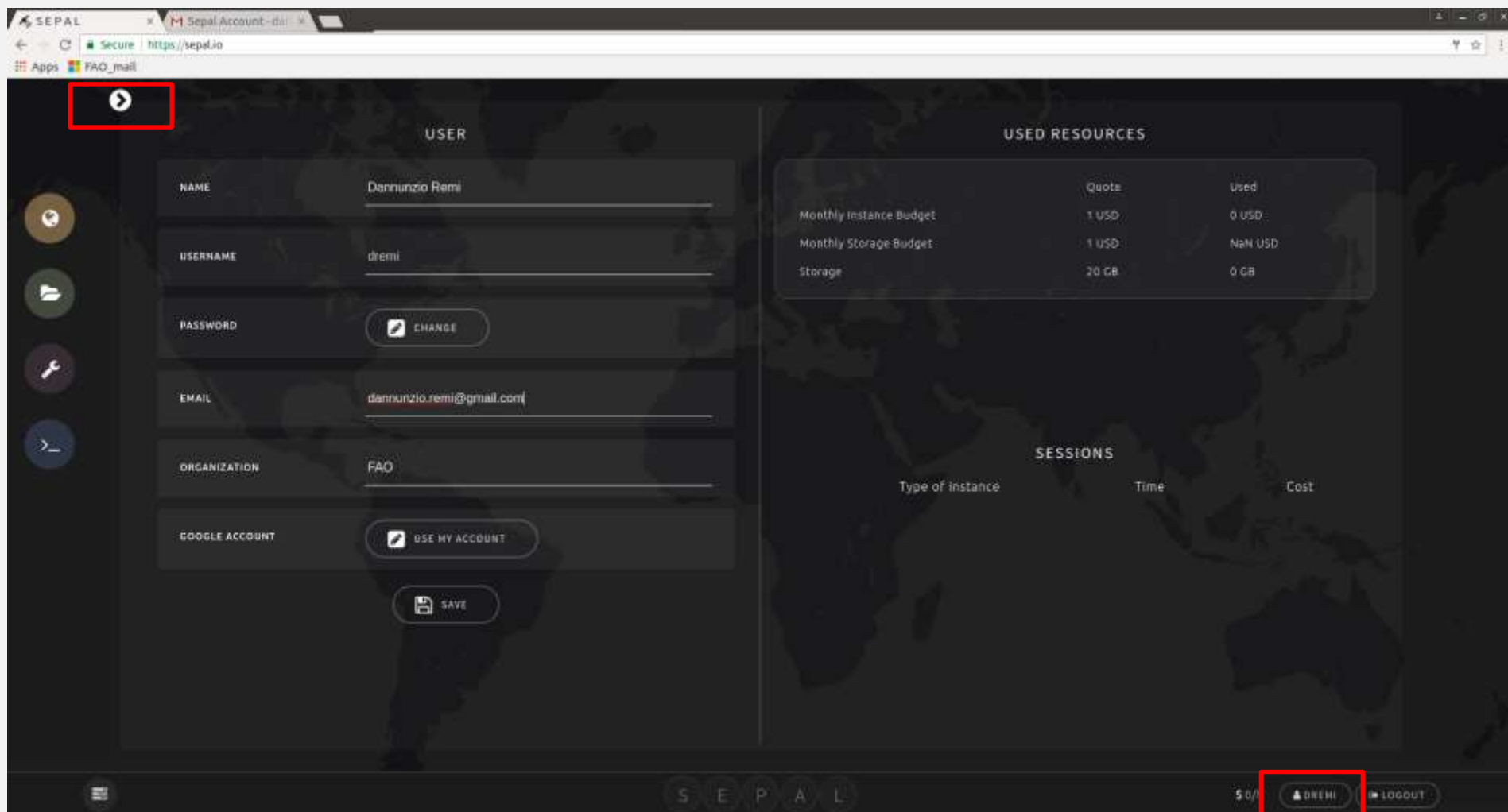
Setup your password in SEPAL

It needs at least one upper case, one lower case and one number.
The minimum size is 6 digits.



You can always reach the platform at <https://sepal.io>

Check your status, budget, parameters



The screenshot displays the SEPAL web application interface. The browser address bar shows the URL <https://sepal.io>. The interface is divided into two main sections: 'USER' and 'USED RESOURCES'.

USER Section:

- NAME:** Dannunzio Remi
- USERNAME:** dremi
- PASSWORD:** Includes a 'CHANGE' button.
- EMAIL:** dannunzio.remi@gmail.com
- ORGANIZATION:** FAO
- GOOGLE ACCOUNT:** Includes a 'USE MY ACCOUNT' button.
- A 'SAVE' button is located at the bottom of the user profile section.

USED RESOURCES Section:

	Quote	Used
Monthly Instance Budget	1 USD	0 USD
Monthly Storage Budget	1 USD	NaN USD
Storage	20 GB	0 GB

SESSIONS Section:

Type of instance	Time	Cost
------------------	------	------

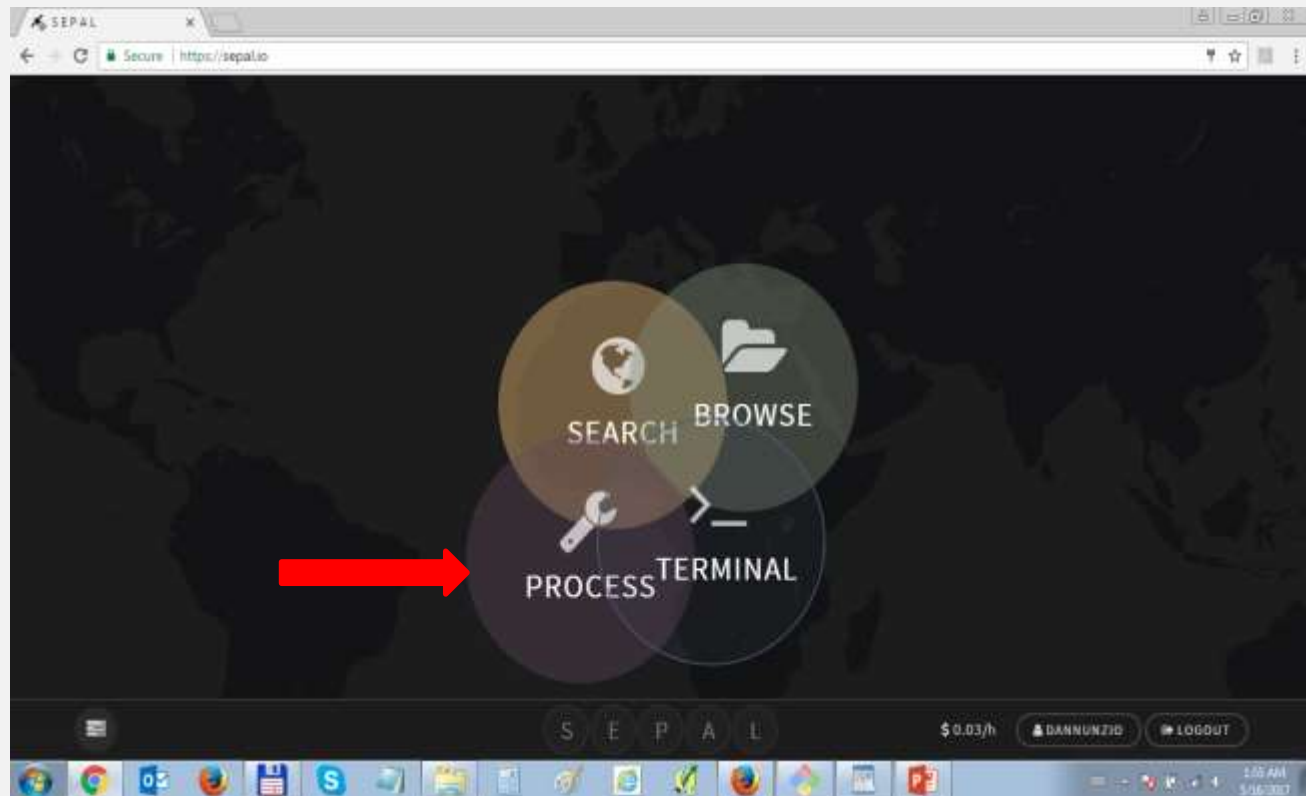
Navigation elements include a top-left arrow icon (highlighted with a red box) and a bottom-right 'DREMI' button (also highlighted with a red box).

Go back to main page (top left arrow)

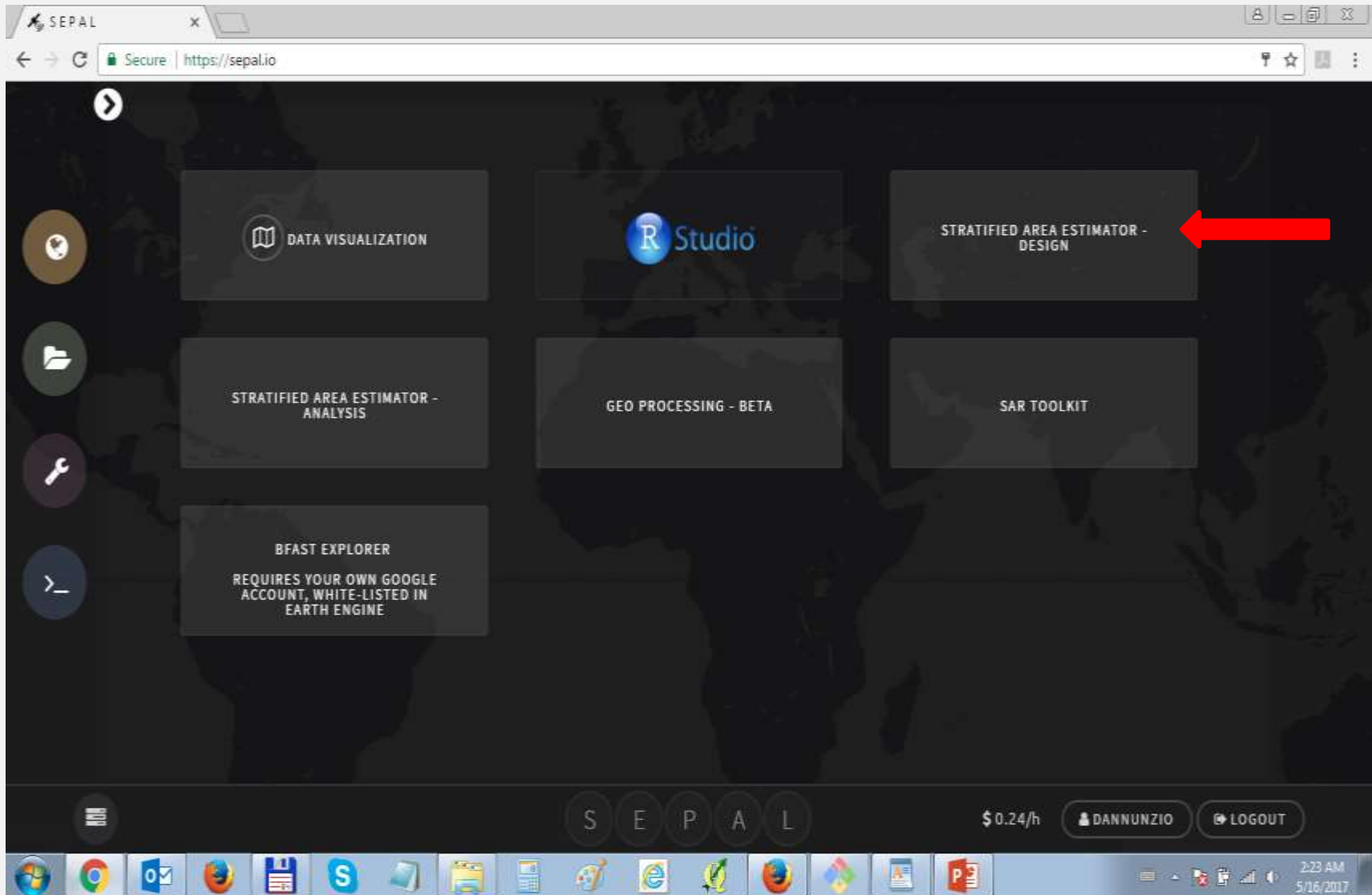
Start the PROCESS tab

There are four fields in SEPAL

- SEARCH for imagery and creating mosaics
- BROWSE through your personal folders and visualize your data
- TERMINAL to access all the command lines possibilities of the LINUX server
- PROCESS access pre-loaded tools and chains of processing



Select SAE – DESIGN tool



Introduction TAB: Select your language

Check the description tabs and the tool structure.
Steps need to be performed one after the other.

The screenshot displays the SEPAL Stratified estimator web application. On the left is a dark sidebar with navigation links: Introduction, Map input, Strata areas, Strata selection, Sampling size, Sample allocation, Source code, and Bug reports. The main content area has a green header with the title 'SEPAL Stratified estimator'. Below the header, there are three main sections: 'Language', 'Description', and 'Background'. The 'Language' section features a dropdown menu with 'English' selected and highlighted by a red box. The 'Description' section contains text about the tool's purpose and a link to the Open Foris support forum. The 'Background' section explains the tool's design and its use in area estimation. At the bottom, there is a 'How to use the tool?' section with a list of five steps. The footer includes logos for SEPAL, UN-REDD, and Open Foris, along with a disclaimer.

SEPAL Stratified estimator

Language

English

English

Français

Español

Description

This interactive tool creates stratified designs to estimate areas. The objective of this tool is to provide a simple user interface for generating a probability dataset with stratified random sampling. For support ask [Open Foris support forum](#)

Background

The aim of this stratified sampling design tool is to produce a sampling design that can be used for area estimates. The idea is to combine a map (used as a stratification of the landscape of interest) with a visual map interpretation of samples to produce an area estimation.

The concept is derived from map accuracy assessment principles; characterized frequency of errors (omission and commission) for each map class may be used to compute area estimates and also to estimate the uncertainties (confidence intervals) for the areas for each class.

How to use the tool ?

You have to go through all the steps in the left panel, in this order:

1. Select the map data which will be assessed. The required input is either vector (.shp supported) or raster (.tif supported)
2. Compute the areas of each strata
3. Select the expected accuracies of the strata
4. Compute the sampling size
5. Draw the sampling points and export as a Collect Earth file

Disclaimer **Reference and Documents**

FAO declines all responsibility for errors or deficiencies in the database or software or in the documentation accompanying it for program maintenance and upgrading as well as for any damage that may arise from them. FAO also declines any responsibility for updating the data and assumes no responsibility for errors and omissions in the data provided. Users are, however, kindly asked to report any errors or deficiencies in this product to FAO.

SEPAL
UN-REDD
Open Foris

Introduction TAB: reference documents

Access reference and background documents, the link will bring you directly there

SEPAL Stratified estimator

Language

English

How to use the tool ?

You have to go through all the steps in the left panel, in this order:

1. Select the map data which will be assessed. The required input is either vector (.shp supported) or raster (.tif supported)
2. Compute the areas of each strata
3. Select the expected accuracies of the strata
4. Compute the sampling size
5. Draw the sampling points and export as a Collect Earth file

Description

This interactive tool creates stratified designs to estimate areas. The objective of this tool is to provide a simple user interface for generating a probability dataset with stratified random sampling. For support ask [Open Foris support forum](#)

Background

The aim of this stratified sampling design tool is to produce a sampling design that can be used for area estimates. The idea is to combine a map (used as a stratification of the landscape of interest) with a visual map interpretation of samples to produce an area estimation.

The concept is derived from map accuracy assessment principles: characterized frequency of errors (omission and commission) for each map class may be used to compute area estimates and also to estimate the uncertainties (confidence intervals) for the areas for each class.

Disclaimer **Reference and Documents**

[REDQ Compass](#)

[Olofsson et al. \(2014\): Good practices for estimating area and assessing accuracy of land change](#)

[FAO NPMA paper N46: Map accuracy assessment and area estimation](#)

STEP 1 Map input

1. Download the test dataset (note it is downloaded to your SEPAL workspace)
2. Select the map you just downloaded (INPUT/aa_data_test/aa_test_congo.tif)

The screenshot shows the SEPAL Stratifed estimator web application. The interface is divided into a left sidebar, a main content area, and a right sidebar.

- Left Sidebar:** Contains navigation links: Introduction, **Map input** (highlighted with a red box), Strata areas, Strata selection, Sampling size, and Sample allocation.
- Main Content Area:**
 - Data type:** A section with instructions: "First choose the type of data used for the stratification - the map. The map can be in raster or vector format. The map area will be calculated in the next tab. The input map can represent a single time or multiple times change made from satellite images. It can also be any acquired from available map data of land cover or land use." Below this, the 'Input' field is highlighted with a red box and labeled with a red '2'. It shows a 'Missing' status.
 - Browse:** A file selection dialog is open, showing a folder named 'aa_data_test' and a file named 'aa_test_congo.tif' (highlighted with a red box).
- Right Sidebar:**
 - Download test data:** A section with a button labeled 'Download test dataset' (highlighted with a red box and labeled with a red '1'). Below it, a text field shows the file path: "INPUT/aa_data_test/aa_test_congo.tif".
 - Output folder:** A section with text: "All products of the random stratified sampling design will be stored here: areas of the map, sampling sizes, point file".

STEP 2 Strata areas

Display map by checking box

Generate the legend by clicking on the button: you can further EDIT the legend

SEPAL - Mozilla Firefox

Request access to sep - SEPAL

https://sepal.io

SEPAL Stratified estimator

Introduction
Map input
Strata areas
Strata selection
Sampling size
Sample allocation
Source code
Bug reports

Area calculation

Map areas are calculated by counting the frequency of the pixels for each map class or by summing the areas of all the polygons. If using raster data the map area can be calculated using R or Open Foris Geospatial Toolkit (OFT). R is compatible with all systems and OFT is only compatible with Linux. Area calculations of large raster files using R will take some time.

☐ OFT
☒ R

Area calculation and legend generation

☒ Do you want to display the map?

Legend and Areas

The areas for each of the map categories need to be calculated in order to calculate the overall and stratified sample size. Make sure to click on the submit legend button to load the map area table.
[Click on area calculation and legend generation](#)
[Click on submit legend before continuing](#)

Legend labeling

The legend classes need to be specified and submitted. Please wait for the map values to appear. Then type the names of the classes and submit the legend.
After submitting the legend the table with the map classes and area will appear. The legend names can be modified at any time in this tab.

[Submit legend](#)

Edit class name for map value: 2
Non Forest

Edit class name for map value: 4
Water

Edit class name for map value: 11
Forest

Edit class name for map value: 12
Forest Secondary

Edit class name for map value: 13
Forest Swamp

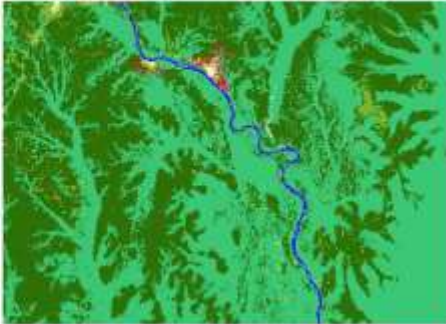
Edit class name for map value: 31
Loss primary forest

Edit class name for map value: 32
Loss secondary forest

Edit class name for map value: 33
Loss swamp forest

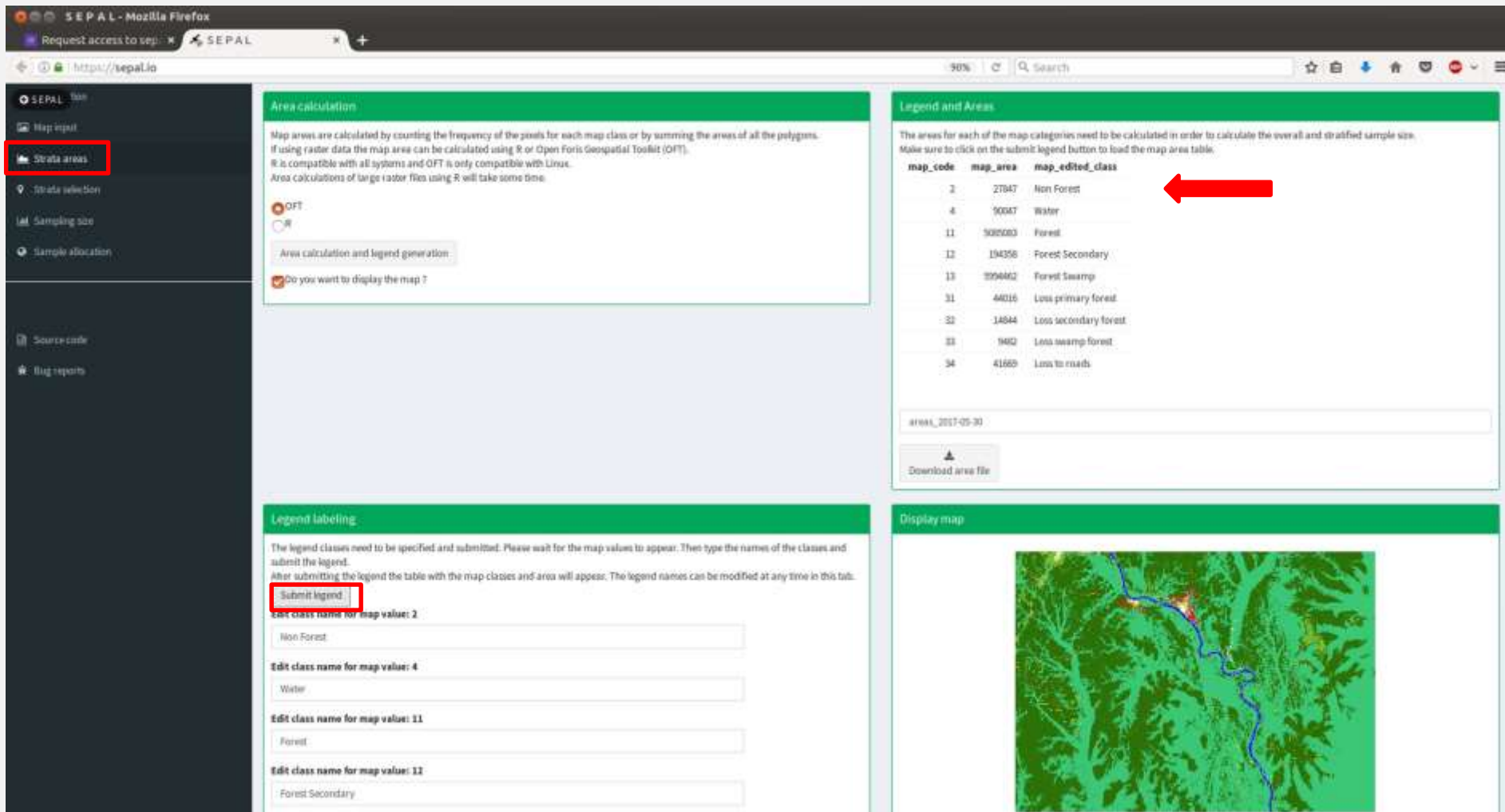
Edit class name for map value: 34
Lost to roads

Display map



STEP 2 Strata areas

Submit the legend : you can edit as much as you want but remember to SUBMIT
The classes name will be showing in later stage for data collection



The screenshot shows the SEPAL web application interface. The left sidebar contains a menu with 'Strata areas' highlighted. The main content area is divided into three sections: 'Area calculation', 'Legend and Areas', and 'Legend labeling'.

Area calculation: This section explains that map areas are calculated by counting the frequency of the pixels for each map class or by summing the areas of all the polygons. It mentions that if using raster data, the map area can be calculated using R or Open Foris Geospatial ToolKit (OFT). It is compatible with all systems and OFT is only compatible with Linux. Area calculations of large raster files using R will take some time. There are radio buttons for 'OFT' and 'R', with 'R' selected. Below this is a button 'Area calculation and legend generation' and a checkbox 'Do you want to display the map?' which is checked.

Legend and Areas: This section states that the areas for each of the map categories need to be calculated in order to calculate the overall and stratified sample size. It instructs to click on the submit legend button to load the map area table. Below this is a table with columns 'map_code', 'map_area', and 'map_edited_class'. A red arrow points to the 'map_edited_class' column.

map_code	map_area	map_edited_class
2	27847	Non Forest
4	90047	Water
11	5025003	Forest
12	194356	Forest Secondary
13	3994462	Forest Swamp
31	44016	Loss primary forest
32	14844	Loss secondary forest
33	1902	Loss swamp forest
34	41669	Loss to roads

Below the table is a text input field with the value 'areas_2017-05-30' and a 'Download area file' button.

Legend labeling: This section states that the legend classes need to be specified and submitted. It instructs to wait for the map values to appear, then type the names of the classes and submit the legend. After submitting the legend, the table with the map classes and area will appear. The legend names can be modified at any time in this tab. There is a 'Submit legend' button. Below this are four text input fields for editing class names for map values 2, 4, 11, and 12. The values entered are 'Non Forest', 'Water', 'Forest', and 'Forest Secondary' respectively.

Display map: This section shows a map of a forested area with a blue line representing a river or road.

STEP 3 Select strata of interest and associated EUA

SEPAL Stratified estimator

Introduction

Map input

Strata areas

Strata selection

Sampling size

Sample allocation

Source code

Bug reports

What are the expected accuracies?

Some classes are identified easier than other classes. Usually common classes, which occupy the majority of the map, are the easiest to identify. Rare classes, such as land change classes, which occupy a small portion of the map area, can be very difficult to identify. This measure will influence the overall sample size.

More classes with lower confidence will increase the overall sample size

- Stable classes are expected to have high user accuracies and should be assigned a higher confidence. Here the value chosen is 0.9
- Rare classes are expected to have the lower user accuracies and should be assigned a low confidence. Here the value chosen is 0.7

Choose classes expected user's accuracies

high confidence (Expected UA = 0.9)

Non Forest Water Forest Forest Secondary Forest Swamp

low confidence (Expected UA = 0.7)

Loss primary forest Loss secondary forest Loss swamp forest

Loss to roads

Expected User's Accuracy (EUA) values for specific classes

High expected user accuracy

0.5 0.75 0.9 1

0.9

Low expected user accuracy

0.5 0.75 0.9 1

0.7

STEP 4 Sampling distribution

Tweak the standard error of overall accuracy & minimum sample size

Samples are allocated by minimum size first and the rest is distributed proportionally to strata size

SEPAL

Stratified estimator

Introduction

Map input

Strata areas

Strata selection

Sampling size

Sample allocation

Source code

Bug reports

Sampling size

In the sampling design, the sample size for each map category is chosen to ensure that the sample size is large enough to produce sufficiently precise estimates of the area of the class (GFOI, 2013)

Standard error of expected overall accuracy

0.01

Minimum sample size per strata

50

☐ Do you want to modify the sampling size?

Formula to calculate the overall sample size

The equation below calculates an adequate overall sample size for stratified random sampling that can then be distributed among the different strata.

- N is number of units in the area of interest (number of overall pixels if the spatial unit is a pixel, number of polygons if the spatial unit is a polygon)
- S(O) is the standard error of the estimated overall accuracy that we would like to achieve
- W_i is the mapped proportion of area of class i
- S_i is the standard deviation of stratum i.

$$n = \frac{(\sum W_i S_i)^2}{[S(\bar{O})]^2 + (1/N) \sum W_i S_i^2} \approx \left(\frac{\sum W_i S_i}{S(\bar{O})} \right)^2$$

Distribution of samples

The computed overall size is : 904

Map Class	Proportional	Adjusted	Final
Non Forest	2	50	50
Water	7	50	50
Forest	401	277	277
Forest Secondary	15	50	50
Forest Swamp	473	327	327
Loss primary forest	3	50	50
Loss secondary forest	1	50	50
Loss swamp forest	0	50	50

Basename of csv to export

sampling_2017-05-30

Download csv with sample design

STEP 5 Sample allocation

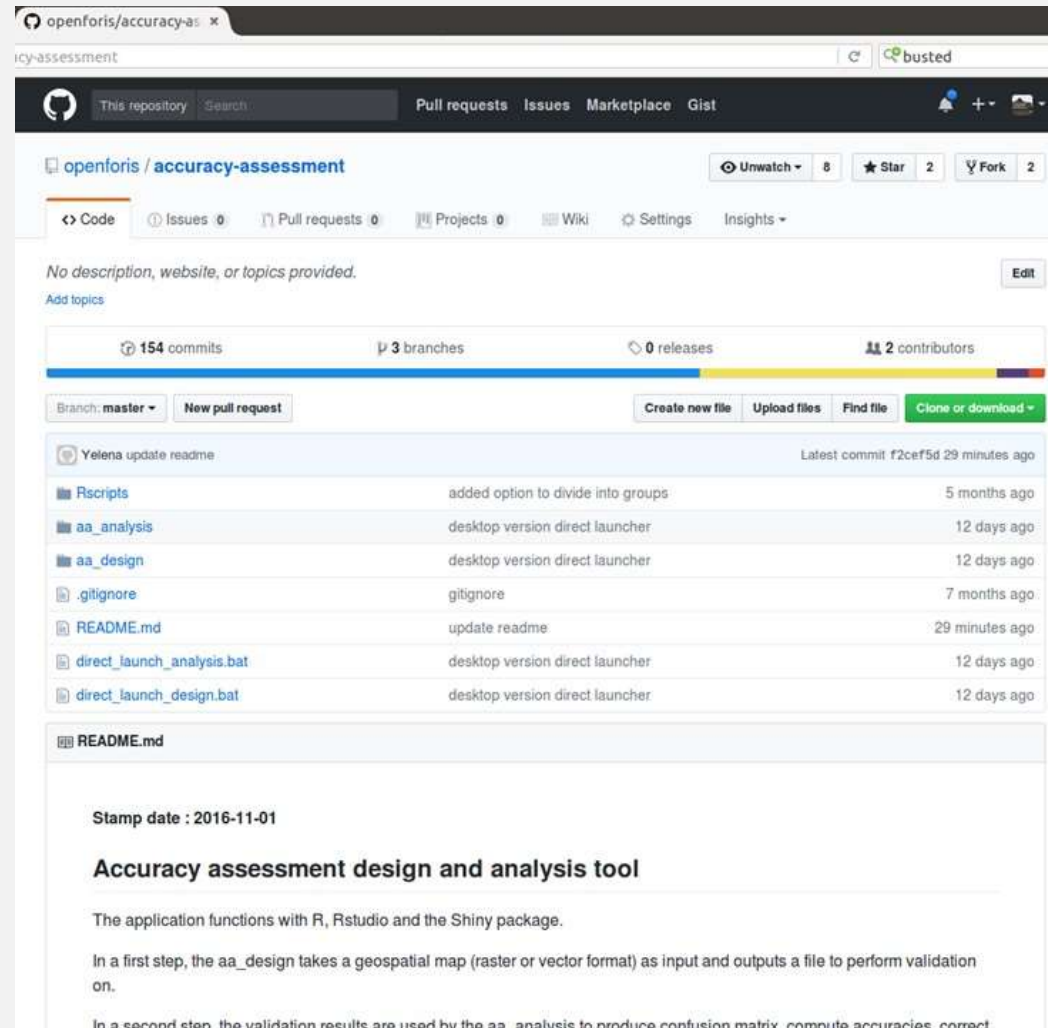
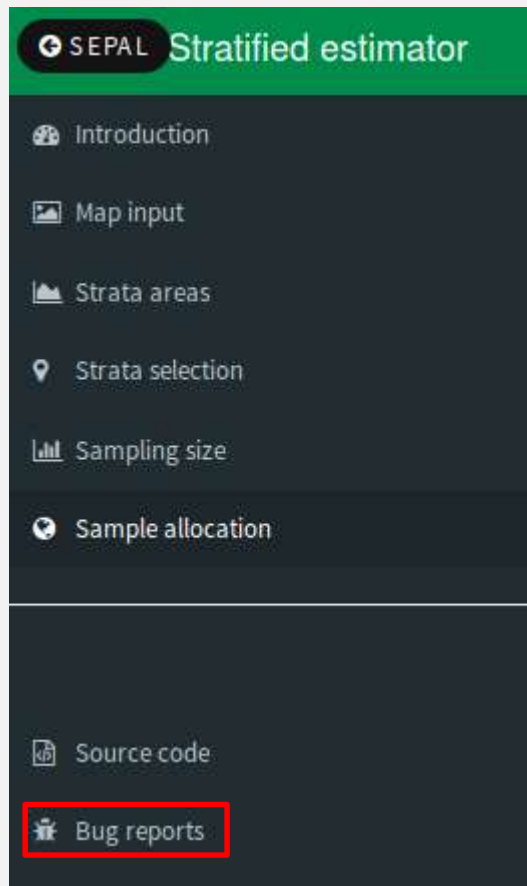
Click on “Generate sampling points” → wait (see task bottom right)

Download CEP project AND tabular data → saves to your computer (and backup in SEPAL workspace)

The screenshot displays the SEPAL Stratified estimator web application. The left sidebar contains navigation links: Introduction, Map input, Strata areas, Strata selection, Sampling size, **Sample allocation** (highlighted with a red box), Source code, and Bug reports. The main content area is titled 'Create a stratified random sample on the map' and includes a map of the Republic of Congo with sampling points. A red box highlights the 'Generate sampling points' button. The right-hand panel, titled 'Create a Collect Earth Project file (.cep) to start validation work', contains fields for 'Choose country name if you want additional national data for the samples' (Republic of Congo), 'Number of operators' (5), and 'Size of the interpretation box (in m)' (30). Below these fields are buttons for 'Download as Collect Earth project (.cep)', 'Download as tabular data (.csv)', and 'Download as vector data (.shp)', with red arrows pointing to them. A file download dialog is open in the foreground, showing the file 'CE_2017-05-30.cep' and options for how to handle it. At the bottom right, a progress bar indicates 'Generating random points'.

Report any bugs, it is work in progress

<https://github.com/openforis/accuracy-assessment>

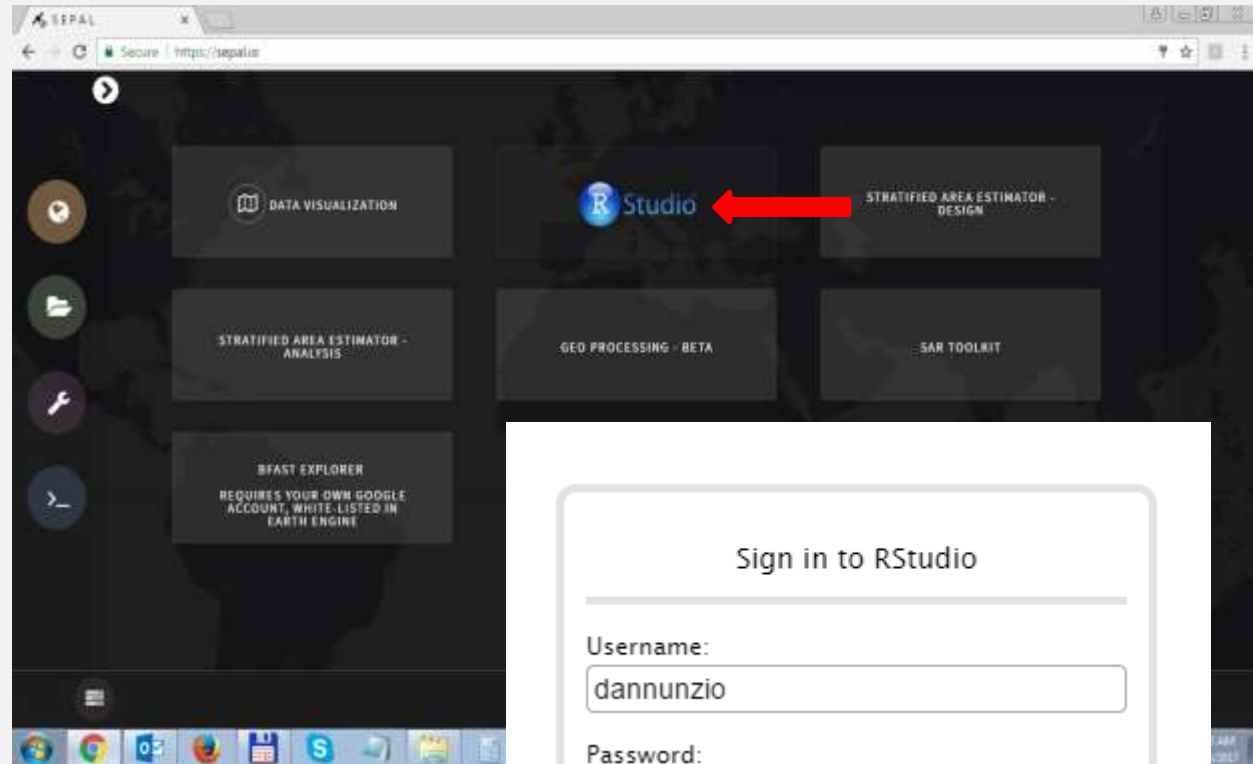
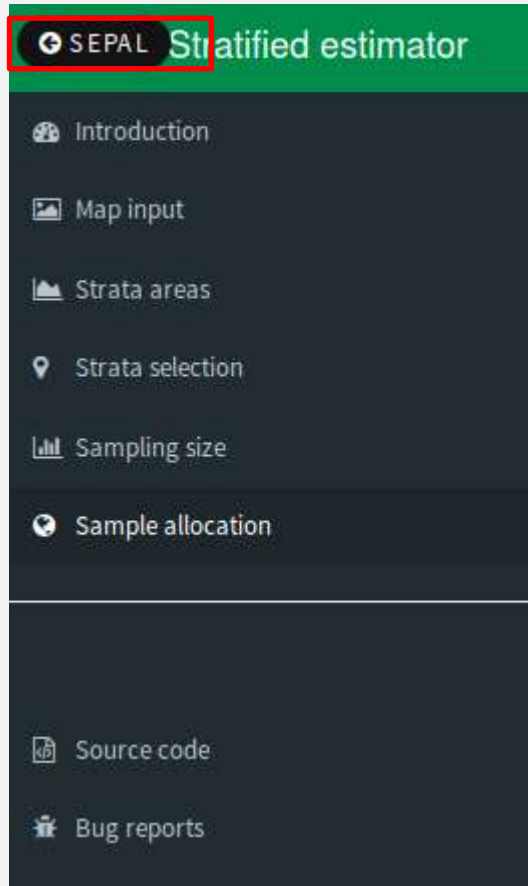




How to upload your own map in SEPAL ?

Select RSTUDIO and sign in

Go back to tools



Sign in to RStudio

Username:
dannunzio

Password:

☐ Stay signed in

Sign In

Only lowercase
in the username

Upload data from your computer

If multiple files (e.g. shapefile), make a zip first

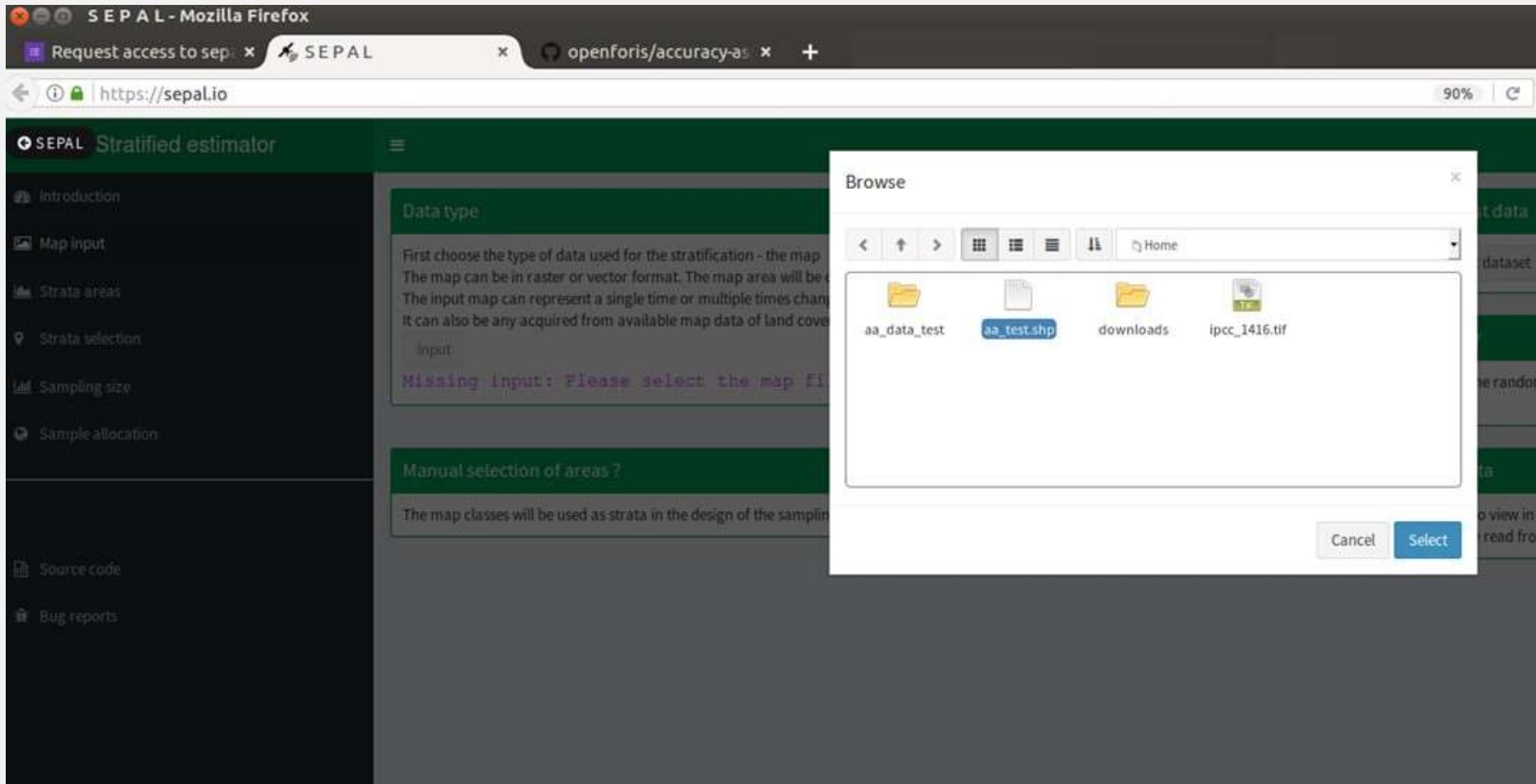
The screenshot displays the SEPAL web interface. On the left, a 'File Upload' dialog is open, showing a file explorer view of the user's local file system. The file 'test_shapefile.zip' is selected. The dialog includes a sidebar with navigation options like 'Recent', 'Home', 'Desktop', 'Documents', 'Downloads', 'Music', 'Pictures', 'Videos', and 'Other Locations'. The main area shows a list of files with columns for 'Name', 'Size', and 'Modified'. The 'test_shapefile.zip' file is highlighted in orange. At the bottom of the dialog are 'Cancel' and 'Open' buttons.

The main interface shows the 'Environment' tab, which is currently empty. Below the environment view, there is a toolbar with buttons for 'New Folder', 'Upload', 'Delete', 'Rename', and 'More'. The 'Upload' button is highlighted with a red box. Below the toolbar, a file explorer view shows the contents of the 'Home' directory, including folders named '.R', 'aa_data_test', and 'downloads'.

Name	Size	Modified
aa_design_output		19:34
aa_test.cpg	5 bytes	31 Mar 2016
aa_test.dbf	212.6 kB	31 Mar 2016
aa_test.prj	390 bytes	31 Mar 2016
aa_test.qpj	599 bytes	31 Mar 2016
aa_test.shp	5.0 MB	31 Mar 2016
aa_test.shx	8.5 kB	31 Mar 2016
aa_test_congo.tif	737.5 kB	23 Aug 2016
FC_2014_Kampongthom_province.dbf	1.5 MB	13 May
FC_2014_Kampongthom_province.prj	402 bytes	12 May
FC_2014_Kampongthom_province.sbn	221.0 kB	13 May
FC_2014_Kampongthom_province.sbx	15.2 kB	13 May
FC_2014_Kampongthom_province.shp	85.5 MB	13 May
FC_2014_Kampongthom_province.shx	174.5 kB	13 May
ipcc_1416.tif	6.2 MB	15 May
test_shapefile.zip	3.8 MB	17:36

Go back to SAE – Design application

Your own map is available for use inside the application



SEPAL - Mozilla Firefox

Request access to sep... x SEPAL x openforis/accuracy-as... +

https://sepal.io 90%

SEPAL Stratified estimator

- Introduction
- Map input
- Strata areas
- Strata selection
- Sampling size
- Sample allocation
- Source code
- Bug reports

Data type

First choose the type of data used for the stratification - the map. The map can be in raster or vector format. The map area will be... The input map can represent a single time or multiple times change... It can also be any acquired from available map data of land cover... input

Missing input: Please select the map file

Manual selection of areas?

The map classes will be used as strata in the design of the sampling

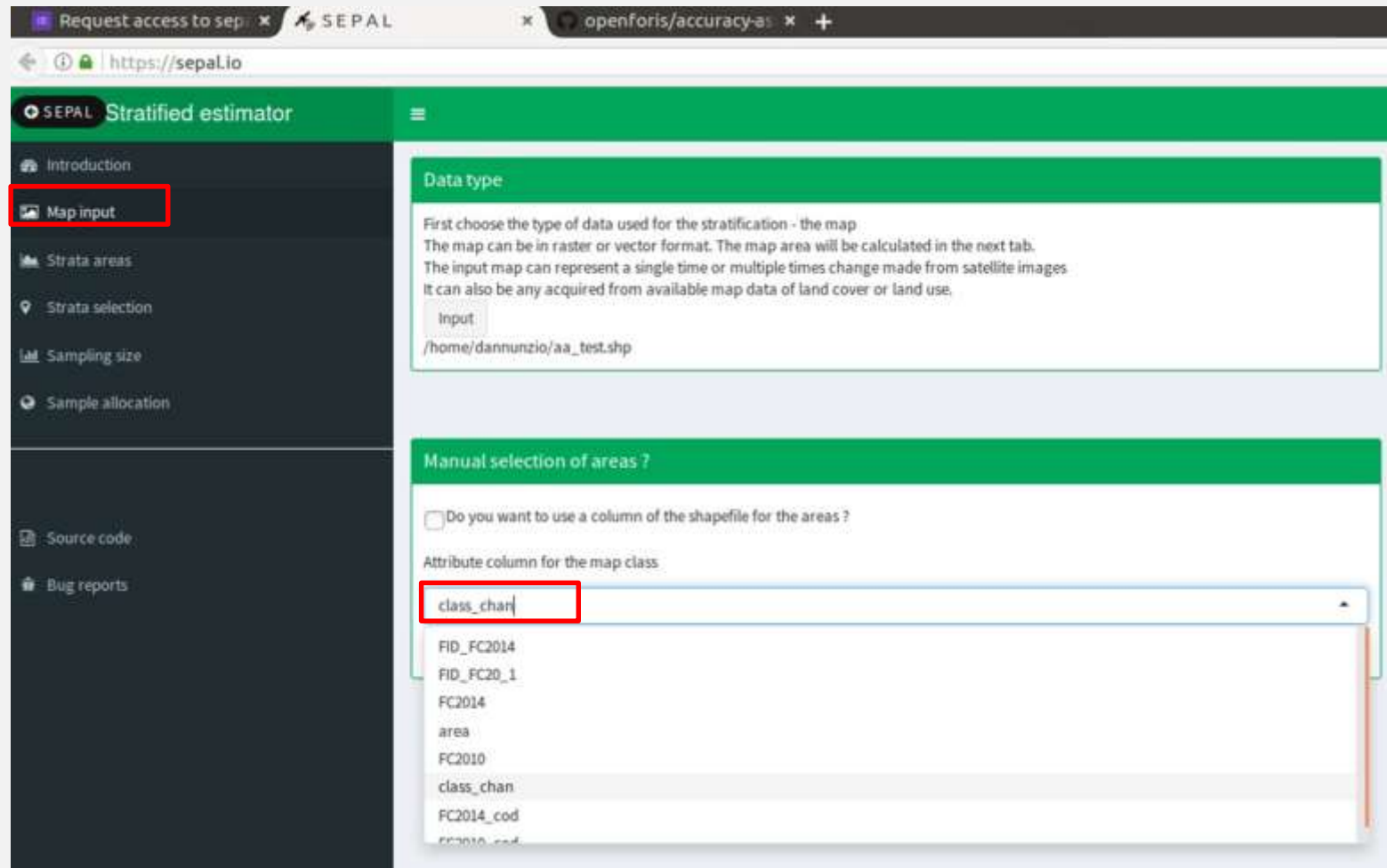
Browse

aa_data_test aa_test.shp downloads ipcc_1416.tif

Cancel Select

Shapefile as input map

Select the right column for class attribute
Rest of the process is identical



The screenshot shows the SEPAL web application interface. The left sidebar contains a menu with the following items: Introduction, Map input (highlighted with a red box), Strata areas, Strata selection, Sampling size, Sample allocation, Source code, and Bug reports. The main content area is titled 'Stratified estimator' and has a green header. Below the header, there is a 'Data type' section with a green header. The text in this section reads: 'First choose the type of data used for the stratification - the map. The map can be in raster or vector format. The map area will be calculated in the next tab. The input map can represent a single time or multiple times change made from satellite images. It can also be any acquired from available map data of land cover or land use.' Below this text is an 'Input:' label and a text box containing the path '/home/dannunzio/aa_test.shp'. Below the 'Data type' section is a 'Manual selection of areas ?' section with a green header. It contains a checkbox labeled 'Do you want to use a column of the shapefile for the areas ?' which is currently unchecked. Below the checkbox is a label 'Attribute column for the map class' and a dropdown menu. The dropdown menu is open, showing a list of columns: FID_FC2014, FID_FC20_1, FC2014, area, FC2010, class_chan (highlighted with a red box), FC2014_cod, and FC2010_cod.