PlotToMap Technical Documentation

Arnan Araza, Sytze de Bruin, Martin Herold

April 2020

0.1 Introduction

Here we automate aboveground biomass plot to map comparison and dubbed it as "PlotToMap" processing chain. *PlotToMap* is a workflow of steps starting from the filtering of deforested plots given that older (than map) plots could still be used as reference data and they might be already deforested. After filtering them, plots biomass are temporally adjusted to align with the map year by using growth increment data. These are applicable at any support unit (e.g. point, polygons, coarse aggregates), and since the map biomass is seldom 100% forests at support units, forest fraction is corrected at the support unit using a tree cover dataset. The filtered, temporally adjusted, and forest fraction-corrected plots at support units are then compared with the map. These steps are highlighted in Figure 1.

This documentation details the data requirement, functionalities, and expected output of *PlotToMap* to guide potential users and help with their own map validation and map usage requirement.
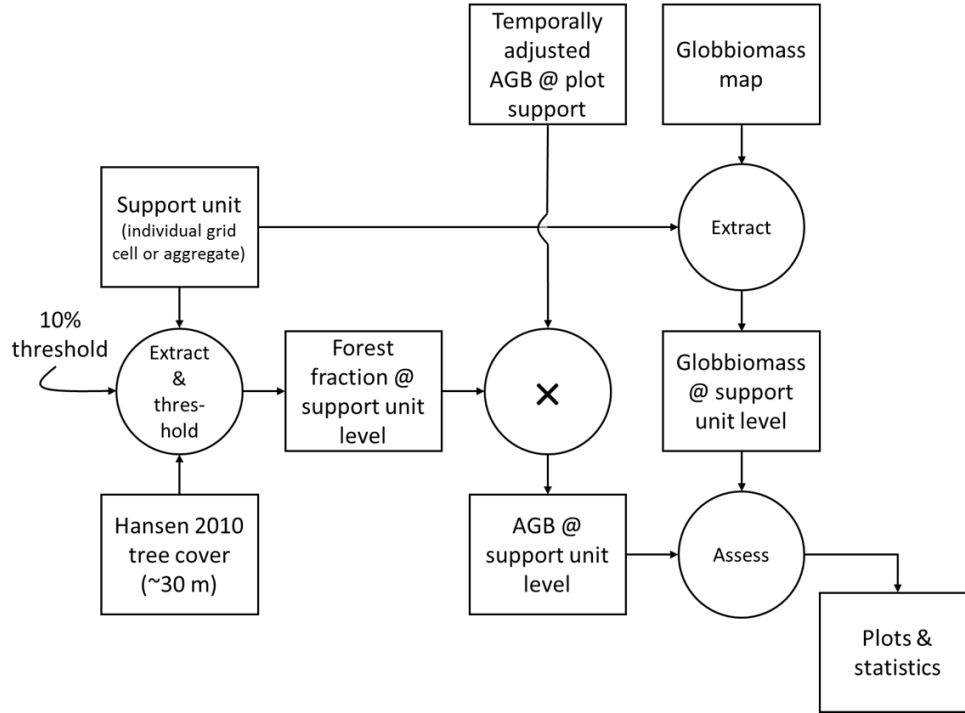

Figure 1: Schema of *PlotToMap* processing chain.

0.2   Preliminaries

Main folder and structure

The main folder structure is shown in Figure 2 containing three sub folders (data, results, and scripts) and the main script "main.R". The data folder has the two tabular data (plot data and growth data) and 2 raster dataset (biomass map and tree cover). There can be chances that latter will not be inside "data" folder because of file size. It's fine.
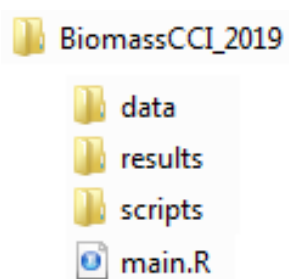

Figure 2: Main folder structure.

Installation

Clone or extract github repository at any desired machine directory:
https://github.com/arnanaraza/BiomassCCI

Input data

Default files

Included in the data folder are pre-processed dataset to assign ecological zones, continents, and to implement the temporal adjustment. Below are the list of *shapefiles* and tables included:

- *eco_zone.shp*. Shapefiles of the global ecological zones used by FAO Simons 2005. The coastal perimeter of this dataset is buffered to capture non-overlapping plots outside coastlines due to coarse resolution.

- *world_region.shp*. Shapefiles of the world administrative regions. The the coastal perimeter of this dataset is buffered to capture non-overlapping plots outside coastlines due to coarse resolution.

- *GR_Uniques.csv*. Growth data unique per ecological zone and continent.

- *GR_SD.csv*. Standard deviation of the growth data.

Plot data and format

The plot data should include the coordinates of plots or corner/boundaries of polygons. The plot data should also have plot identification label, plot size, year of inventory, and the plot biomass. See Table 1 for sample.

Table 1: Sample and required columns and naming of plot data.

Entity/Column	Data type	Description	Example
PLOT_ID	Character	Unique ID	FOS1
POINT_X	Numeric	Longtiude	-3.007917056
POINT_Y	Numeric	Latitude	6.01019559
AGB_T_HA	Numeric	Biomass value	106.66
AVG_YEAR	Numeric	Average year	2007.5
SIZE_HA	Numeric	Plot size (ha)	0.1

For polygon plot data, the coordinates should be separate from the other plot-level data. See the default data folder for example.

Raster inputs

The default biomass map used for this workflow is the GlobBiomass 2010 map for demonstration purposes. Aside from that, a tree cover dataset is needed preferably one from Global Forest Change (GFC) in 2010. Tiles of GlobBiomass and GFC Tree Cover can be accessed from these respective links:

- http://globbiomass.org/wp-content/uploads/GB_Maps/Globbiomass_global_dataset.html
- https://glad.umd.edu/gladmaps/globalmap.php

0.3   main.R

The "main" script loads libraries, defines global variables, and calls other functions to perform the PlotToMap steps sequentially.

It uses 7 packages:

- 2 for spatial operations *(rgdal and raster)*
- 2 data frame operations *(plyr and dplyr)*
- 3 for parallel processing *(foreach, parallel, doParallel)*
- 1 for graphing *(plotrix)*
- 1 for downloading tiles *(gfcanalysis)*

There are 9 global variables as follows:

- *mainDir.* Main directory
- *scriptsDir.* Directory of scripts
- *outDir.* Directory of all results
- *dataDir.* Directory of data
- *plotsFile.* Filename of plot data
- *agbTilesDir.* Directory of AGB tiles
- *treeCoverDir.* Directory of Hansen tree cover tiles
- *forestTHs.* Forest threshold, can be a number or vector
- *mapYear.* Year of map biomass in nearest tens e.g. 2010 = 10
- *flDir.* Desired/existing directory of GFC forest loss tiles

0.4   Subscripts

Polygonize.R

Function:

- *Polygonize.* Creates a polygon from coordinates of (1) plot corners and (2) convex hull of within-plot coordinates.

Variables:

- *x*. Numeric. Longitude of plot data

- *y*. Numeric. Latitude of plot data

- *pltID*. Character. Plot identification label for each coordinate

Deforested.R

Function:

- *Deforested*. Removes deforested plots <= the map year by overlapping plot locations into at most four forest loss (FL) tiles from Global Forest Change by Hansen et al. 2013. For relatively bigger plots (e.g. 25ha), if partial deforestation is >50% then the plot is removed, otherwise plot is retained.

Variables:

- *plt*. Data frame. Default plot data

- *fdir*. File path of FL directory folder

BiomePair.R

Function:

- *BiomePair.R*. Overlays the plots to corresponding continent referred to as Zones and FAO eco-zones (FEZ) referred as biomes. From the FEZ classes, the general eco-zone (GEZ) class is obtained e.g. FEZ = tropical mountainous forest; GEZ= tropical forest.

Variables:

- *plt*. Data frame. the plots without deforested ones.

TempFix.R

Function:

- *TempFix*. Adjusts plot biomass values to align with the map year by adding or subtracting annual growth increment to older or newer plots. The growth data comes from a model-based estimate of growth-age relationships using permanent and re-measured plots Suarez et al. 2019. This dataset is adopted and reported by the new Inter-Panel in Climate Change Guidelines 2019 Report. The growth-adjusted biomass also incorporates uncertainties from the growth data used.

Variables:

- *df*. Data frame. Default plot data.
- *domain*. Character. Continent of interest.
- *year*. Numeric. Map AGB year i.e. 2010 for GlobBiomass.

TempEffect.R

Functions:

- *HistoShift*. Creates histogram of pre and post temporal adjustment as well as its overlap. Returns histogram and saves it at *outDir*.
- *ChangeTable*. Creates a table summarizing the temporal adjustment effect per biomass range. Returns and saves the table at *outDir*.

Variables:

- *old*. Data frame. Default plot data
- *new*. Data frame. Temporally adjusted plot data

MakeBlockPolygon.R

Function:

- *MakeBlockPolygon*. Makes a polygon out of original plot location of point reference data to be aligned with AGB map pixel or larger cell over which AGB is aggregated. Returns a Spatial Polygon.

Variables:

- *x*. Numeric. Longitude of plot data
- *y*. Numeric. Latitude of plot data
- *size*. Numeric. Target aggregation scale in degrees

TileNames.R

Function:

- *TCtileNames*. Takes up to four tree cover (TC) tile names covered by polygon. Returns 4 TC file names.

- *AGBtileNames.* Takes up to four AGB tile names covered by polygon. Returns 4 AGB file names.

Variables:

- *pol.* Spatial Polygon. The created polygon from desired aggregated scale.

BlockMeans.R

Function:

- *sampleTreeCover.* Generates block mean mapped forest over polygon.

Variables:

- *pol.* Spatial Polygon. The created polygon from desired aggregated scale.
- *thresholds.* Numeric. Forest threshold, defaults at 10%.
- *wghts.* Boolean. Enables block mean of forest cover of polygon.

Function:

- *sampleAGBmap.* Generates block mean mapped AGB over polygon.

Variables:

- *pol.* Spatial Polygon. The created polygon from desired aggregated scale.
- *wghts.* Boolean. Enables block mean of map AGB of polygon.

InvDasymetry.R

Function:

- *InvDasymetry.* Applies forest scaling and spatial aggregation. Returns a data frame with mean AGB from plots satisfying selection criteria of 5 plots per block polygon. Maximizes computer cores via parallel processing.

Variables:

- *clmn.* Character. Choice between "biome" or "continents".
- *value.* Character. Continent name or biome name i.e. "boreal"
- *aggr.* Boolean. Whether to aggregate or not.
- *minPlots.* Numeric. Minimum number of plots to create a block polygon. Default is 5.
- *wghts.* Boolean. Enables block mean of forest cover of polygon.

0.5   Operation

Error handling

Errors from the following are handled by try-stop operations with error messages and suggestions

- Errors from not attaining 5 plots minimum = "There are no records satisfying the selection criterion".

- Few plots after aggregation = "too few plots selected, try decreasing min-Plots or run non-aggregated model"".

- Errors from limited computer memory mainly due to raster aggregation = see *?memory.limit()*

Run time

For 100,000 plot data and a machine that has 8 gig RAM with 8 cores:

- Deforested plot removal = 4 hours

- Biome and continent labelling = 2 hours

- Temporal adjustment = 2 minutes

- Forest scaling, with aggregation = 1 hour

- Forest scaling, no aggregation = 3 hours

Bibliography

M. C. Hansen, P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, A. Kommareddy, A. Egorov, L. Chini, C. O. Justice, and J. R. G. Townshend. High-resolution global maps of 21st-century forest cover change. *Science*, 342(6160): 850–853, 2013. doi: 10.1126/science.1244693.

H. Simons. Global ecological zoning for the fao global forest resources assessment 2000 entwicklung einer globalen ökologischen zonierung für die weltweite bewertung der waldressourcen im jahr 2000 seitens der fao. *Anwendung und Auswertung der Karte der natürlichen Vegetation Europas Application and Analysis of the Map of the Natural Vegetation of Europe*, page 55, 2005.

D. R. Suarez, D. M. A. Rozendaal, V. D. Sy, O. L. Phillips, E. Alvarez-Dávila, K. Anderson-Teixeira, A. Araujo-Murakami, L. Arroyo, T. R. Baker, F. Bongers, R. J. W. Brienen, S. Carter, S. C. Cook-Patton, T. R. Feldpausch, B. W. Griscom, N. Harris, B. Hérault, E. N. H. Coronado, S. M. Leavitt, S. L. Lewis, B. S. Marimon, A. M. Mendoza, J. K. Ndja, A. E. NGuessan, L. Poorter, L. Qie, E. Rutishauser, P. Sist, B. Sonké, M. J. P. Sullivan, E. Vilanova, M. M. H. Wang, C. Martius, and M. Herold. Estimating aboveground net biomass change for tropical and subtropical forests: Refinement of IPCC default rates using forest plot data. *Global Change Biology*, 25 (11):3609–3624, 2019. doi: 10.1111/gcb.14767.