

SERVIR GLOBAL

Tools for Carbon Emissions Estimation and Mapping in East Africa Guide

The following information was adapted from Ned Horning's *Workshop on Tools for Carbon Emissions Estimation and Mapping in East Africa*, developed as a part of a NASA-SERVIR funded project "Forest carbon assessment for REDD+ in the East Africa SERVIR region" – a collaboration of the Woods Hole Research Center, the American Museum of Natural History, and the Regional Centre for Mapping of Resources for Development.

Objectives of the training:

The main purpose of the training is to learn how to use open source software to estimate carbon emissions from deforestation using reference geospatial datasets, such as multi year land cover land use maps, and above ground biomass data.

Main analyses to be learned using R:

- Perform geospatial analyses, including vector and image processing
- Create land cover map classifications using Random Forest method
- Perform land cover land use change analysis and identify deforestation
- Calculate accuracy assessments and validation of land cover maps
- Calculate carbon emissions from deforestation based on temporally comparable biomass and deforestation maps

This guide explains how to navigate and use the training materials, including R scripts, presentations and sample datasets.

I. Preparing the Software

These materials utilize the open source software R, RStudio and QGIS. You can download these programs directly from their websites. They are available for Mac OS X, Windows, and Linux:

- a. QGIS: <http://www.qgis.org/en/site/>
- b. R: <https://www.r-project.org/>
- c. RStudio: <https://www.rstudio.com/products/rstudio/download/>

Udemy hosts a free overview of R Basics that is very helpful for new R users.
<https://www.udemy.com/r-basics/>

There are several packages you will need to install in RStudio in order to run the analyses. Not all of these packages are necessary for every step but they will be used in order to run through all the exercises, They can be installed directly through the RStudio install feature by searching for the package name (Figure 1). These packages are:

- | | | |
|---------------|------------------|-------------------|
| 1. car | 16. labeling | 31. raster |
| 2. caret | 17. lme4 | 32. RColorBrewer |
| 3. chron | 18. magrittr | 33. Rcpp |
| 4. colorspace | 19. maptools | 34. RcppArmadillo |
| 5. data.table | 20. MatrixModels | 35. RcpEigen |
| 6. dichromat | 21. minqa | 36. reshape2 |
| 7. digest | 22. munsell | 37. rgdal |
| 8. doParallel | 23. nloptr | 38. rgeos |
| 9. e1071 | 24. pbkrtest | 39. RStoolbox |
| 10. foreach | 25. plyr | 40. scales |
| 11. gdalUtils | 26. quantreg | 41. sp |
| 12. geosphere | 27. R.methodsS3 | 42. SparseM |
| 13. ggplot2 | 28. R.oo | 43. stringi |
| 14. gtable | 29. R.utils | 44. stringr |
| 15. iterators | 30. randomForest | 45. XML |

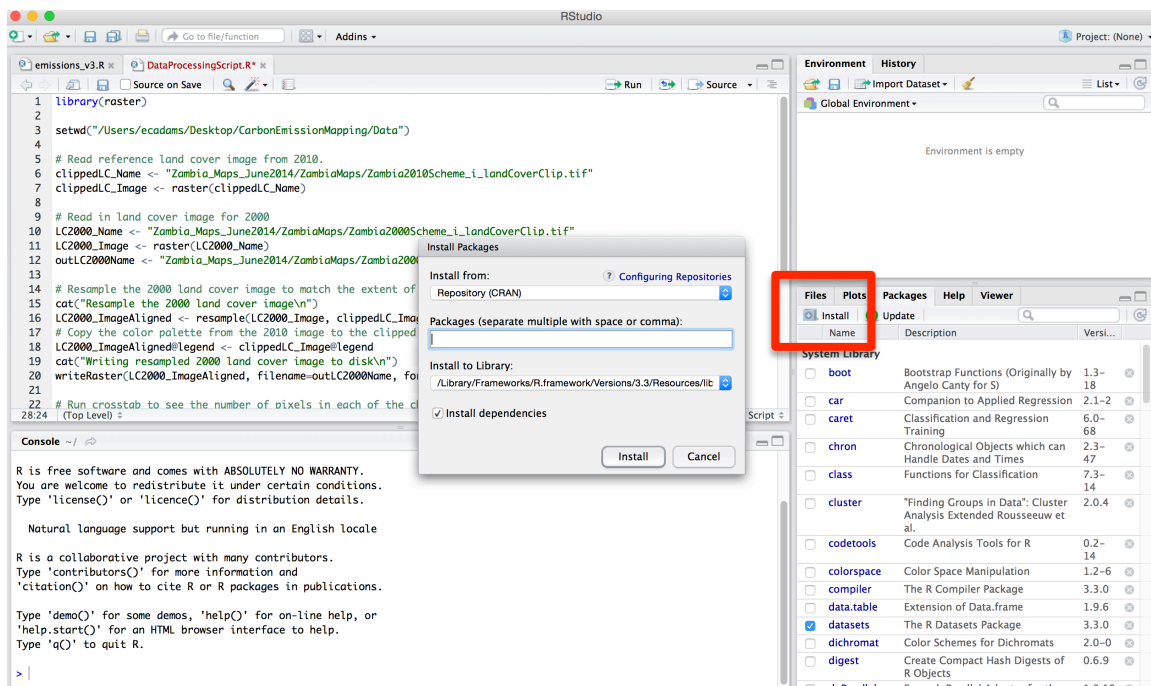


Figure 1. R studio platform screenshot - showing how to install packages.

II. Review of R

This folder contains 4 folders, 1. Vector Processing In R Demo, 2. Image Processing In R, 3. Recent Geospatial Updates In R, and Data.

Each of these folders contain information relevant to utilizing R for vector and raster processing in general. If you are a novice user of R for geospatial analysis the material in these folders will serve as a good review with examples of some of the core capabilities of R. It is recommended to go through the materials in each of these folders in order to gain an understanding of some of the processing behind estimating carbon emissions. Each of the folders have presentations that contain reference information as well as R code. The Data folder contains all of the referenced files for each of the R scripts in each folder except for the Landsat data. In order to download the Landsat data, go to earthexplorer.usgs.gov. Enter search criteria as: Path 174 and row 69 for Landsat 8, January 2, 2016. Save this scene in the Landsat8 folder. Navigate to Landsat 5 and download the scenes for November 20, 2010 and October 20, 2000 for the same path and row. Save them in the LT5 folder. You will need a USGS account in order to download the scenes.

Opening the scripts within RStudio allows for easy manipulation, visualization, and execution of the codes. You will need to set your working directory to the ReviewOfR/Data folder. The codes are commented to indicate what is occurring at each step. Below are helpful suggestions for each of the scripts in the folders

1. Vector Processing in R
 - a. VectorOperationsInR.R – Run the code line by line to see examples of various vector data functions using the example data
 - i. *Line 4* - Change the working directory
 - ii. *Line 105 & 106* – This function is to get coordinates from the plot, which are user defined. When you run this line, it will activate the plots window for you to click the location you would like to identify the coordinates for. Press finish in the plot window after you have clicked to complete the analysis. You will repeat this process again for the following lines.
2. Image Processing in R
 - a. ImageOperationsInR.R- Run the code line by line to see examples of various raster processing capabilities
 - i. *Line 2* – Change the working directory
 - ii. *Line 48 & 49* – This function is similar to lines 105 and 106 in the Vector Operations In R script. You can click and interact with the plot window. Clicking finish completes the analysis
3. Recent Geospatial Updates in R
 - a. testRStoolbox_demo.R – Run the code line by line to illustrate some of the capabilities of the RStoolbox, including supervised (random forest) and unsupervised classifications. Land cover classifications will be needed to for the carbon emissions calculations later in this demonstration. Make

sure your working directory is pointed to the Data folder in the ReviewOfR folder.

- b. Accuracy Assessment
 - i. accuracyAssessment.R – serves as a general example on how to perform accuracy assessments on user generated land cover classification maps, example data is provided in the Data folder and the script is currently written to use this example data. Should the user wish to apply this method to their own data, they will need to update more than the working directory as indicated in the comments of the script
 - ii. validateMapExample.R – uses external data to validate an example map, similar to the accuracy assessment example, you will have to update the fields with information relevant to your particular study area to work outside the example given.

III. Carbon Emissions

The Carbon Emissions folder contains a presentation, which serves as an overview to some of the concepts regarding carbon emissions mapping, 2 R scripts, emissions_v3.R and DataProcessingScript.R, and a Data folder. The DataProcessingScript.R script is useful for preprocessing the necessary data to calculate carbon emissions using the “emissions_v3.R” script. Example data has been provided in the Data folder. The details below will outline the data requirements and steps to calculate carbon emissions.

Data Preparation: Collect the following data to calculate carbon emissions using the emissions_v3.R script:

- A biomass map for your region in .tiff format. Woods Hole Research Center has produced a national dataset for the pantropical region of aboveground live woody biomass derived from field/LiDAR(GLAS) and MODIS at 500 m spatial resolution in tons per hectare. The dataset and more information can be found at: <http://whrc.org/publications-data/datasets/pantropical-national-level-carbon-stock/>
 - For demonstration purposes the ZambiaBiomass500m.tif file will be used
- A land use land cover change map highlighting the areas of forest loss also in .tiff format. This can be created using 2 land cover maps from different years and the script DataProcessing.R included with these materials.
 - The land use land cover maps used in this example are available at apps.rcmrd.org/landcoverviewer
 - Alternatively, you can use the capabilities highlighted in the testRStoolbox_demo.R, the accuracyAssessment.R, and the validateMapExample.R scripts to create land cover classification maps for your region.
 - For demonstration purposes the Zambia2000scheme_i_landCoverClip.tif file and the Zambia2010Scheme_i_landCoverClip.tif file in the ZambiaMaps

folder can be used. The accuracy assessment information is also included for reference. These maps used the 6 IPCC, the International Panel on Climate Change, top-level land categories for greenhouse gas (GHG) inventory reporting.

- You will also need to determine the land cover class values that represent deforestation. If you used the DataProcessingScript.R those values were combined in the legend of the land use change image. For example: the legend from the RCMRD land cover viewer is: 1. Forestland, 2. Grassland, 3. Cropland, 4. Wetland, 5. Settlement, 6. Otherland. A value of 12 in the land cover change image would indicate the previous image was forestland but is now grassland; a value of 43 would indicate that area was wetland and is now cropland. Thus the values of 12, 13, 14, 15, and 16 would represent deforestation using these data.
- The emissions_v3.R script performs a number of processing steps. After the biomass and land cover change image are input they are compared and then both are cropped to the common area covered by both images. Next the biomass image is resampled so it is the same resolution as the land cover change image. The biomass per pixel is then calculated since the original biomass units are biomass per hectare. The biomass map should reflect pre-deforestation conditions. The last analysis step is to use zonal statistics to calculate the sum of the biomass per pixel values for each of the deforestation classes and then convert that number to emitted CO₂ by multiplying by 0.5 (convert weight of dry biomass to carbon)* and multiply the result by 3.66 (convert weight of carbon to CO₂). The emitted CO₂ values are output in tons in the terminal window along with the total CO₂ for all classes. If an output image was defined in the variables section then the last processing step is to output an image with biomass per pixels values. The R script has comments to help you understand each processing step.

*The 2006 IPCC Guidelines for National Greenhouse Gas Inventories, Chapter 4, page 48 provides conversion factors to calculate carbon from biomass depending on your region and land cover class.

Guide for using the example data:

1. Create the land use land cover change map following the DataProcessingScript.R and using the Zambia Maps for 2000 and 2010. The only step involved here is to make sure you set your working directory. The following outlines the steps occurring in the script
 - a. The script outputs the LC_Change2000Times10Plus2010.tif file in the working directory, which is the land cover land use image file for the carbon emissions calculations
2. In the emissions_v3.R script you set the variables as follows
 - i. Setwd(your own working directory)
 - ii. The biomassImageName is ZambiaBiomass500m.tif
 - iii. Nodata value is 0

- iv. The name and path of the land cover change image is the file we generated in the previous step LC_Change2000Times10Plus2010.tif
- v. The pixel values for deforestation are 12, 13, 14, 15, 16
- vi. The land cover change image nodata value is 0
- vii. targetImage is lcc
- viii. You can name the out image whatever you would like, as well as change the destination path. For this example it is set to write the BiomassPerPixel.tif into the data folder.
- ix. Run the entire script at once, not line by line
- x. The script will print the total CO₂ emissions for each of the deforestation classes followed by the total emissions and the output image, ZambiaBiomass500m.tif, in tons of biomass per pixel instead of tons of biomass per hectare