Pontus Olofsson *(olofsson@bu.edu)*
Earth & Environment, Boston University

4. Response Design

Training module: Use of global tree cover and change datasets in REDD+ Measuring, Reporting and Verifying. Boston, May 2015.

4.1 Introduction

At this stage a sample has been selected which needs to be interpreted using a suitable source of reference data. This step is referred to as the response design and includes providing reference labels for each unit in the sample. With each unit having a map label and a reference label we can construct an **error matrix** which contains all the information needed to perform the analysis (Section 5). The use of TimeSync (http://timesync.forestry.oregonstate.edu/) is recommended, especially if the stratification covers a large area as Landsat data can be retrieved without having to download it – see point 7 below.

4.2 Interpret sample

1. Display the reference data in QGIS, i.e., display the data you will use to interpret the sample you just created. This is likely a combination of different data sources, such as Landsat, RapidEye and Google Earth, acquired around the same times as the data used to create the map (in this case 2000 and 2012), and preferably also in-between. **Do not display the map!**
2. Display the shapefile containing the sample, i.e. the file you created in Section 3.
3. Right-click shapefile in *Layer* pane; *Open Attribute Table*; then  and then  ; delete the STRATUM column.
4. Click the *New column* button  to add a column; name it "reference"; leave options as default except *Width* which should be set to 3.
5. Now provide a label for each of the units in the sample by manually examining the reference data. Add label
   "1" for stable non-forest,
   "2" for forest,
   "3" for water,
   "4" for forest loss,
   "5" for forest gain (note that this stratum does not exist in this example),
   "6" for forest loss/gain (note that this stratum does not exist in this example).
   Since your final area estimates are based on the interpretation of this sample it is important that the labels are correct – if you can't provide a correct label then delete the unit rather than guessing. You can click  to jump to the highlighted unit. Make

sure you save the shapefile regularly.

6. If you want to open the sample in Google Earth TM, right click the shapefile with the sample > *Save As...* > in the *Save As* dialog, set *Format* to *Keyhole Markup Language [KML]*, specify an output file and set *NameField* to *ID*; leave other options as default > click *OK* . You can also use the GDAL program "ogr2ogr" (www.gdal.org/ogr2ogr.html) to create the KML file: either paste the following into the terminal: `ogr2ogr -f "KML" test_ge.kml test.shp -dsco NameField=ID`
7. In addition to (or instead of) the workflow described above, the use of the TimeSync software is highly recommended (http://timesync.forestry.oregonstate.edu/). TimeSync is by the time of writing not fully operational but is expected to be soon. The main benefit of TimeSync is the retrieval of annual Landsat data for each unit in the sample without having to download the data.

4.3 Construct the error matrix

1. With each unit having a map label and a reference label we can construct an **error matrix**. This can be done in various ways but we recommend using a home-made script that executes in the terminal; if not present, download the script from https://raw.githubusercontent.com/ceholden/accuracy_sampler/master/script/crosstab.py and place it in the directory where the sample shapefile is located.
2. Open a MATE terminal and navigate to the directory where the sample shapefile and "crosstab.py" are located.
3. Type `python crosstab.py -v -a [column] [map].tif [shapefile].shp errormatrix.txt` where "[column]" is the column in the shapefile that contains the reference labels, "[map].tif" is the map that is being assessed (the stratification created in Section 3 in this case) and "[shapefile].shp" is the sample shapefile. This will create textfile that contains the error matrix called "errormatrix.txt".