Food and Agriculture
Organization of the
United Nations

Map Accuracy Assessment and Area Estimation

A Practical Guide

National forest monitoring assessment working paper No.46/E

National forest monitoring assessment working paper No.46/E

Map Accuracy Assessment and Area Estimation: A Practical Guide

Food and Agriculture Organization of the United Nations
Rome, 2016

Disclaimer

The designations employed and the presentation of material in this information product do not imply the expression of any opinion whatsoever on the part of the Food and Agriculture Organization of the United Nations (FAO) concerning the legal or development status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. The mention of specific companies or products of manufacturers, whether or not these have been patented, does not imply that these have been endorsed or recommended by FAO in preference to others of a similar nature that are not mentioned.

The views expressed in this information product are those of the author(s) and do not necessarily reflect the views or policies of FAO.

© FAO, 2016

FAO encourages the use, reproduction and dissemination of material in this information product. Except where otherwise indicated, material may be copied, downloaded and printed for private study, research and teaching purposes, or for use in non-commercial products or services, provided that appropriate acknowledgement of FAO as the source and copyright holder is given and that FAO's endorsement of users' views, products or services is not implied in any way.

All requests for translation and adaptation rights, and for resale and other commercial use rights should be made via www.fao.org/contact-us/licence-request or addressed to copyright@fao.org.

FAO information products are available on the FAO website (www.fao.org/publications) and can be purchased through publications-sales@fao.org

Cover photo: JAXA K&C imagery processed by Andreas Vollrath

Contents

Acknowledgements

This publication was written by the following authors: Yelena Finegold, Antonia Ortmann, Erik Lindquist, Remi d’Annunzio and Marieke Sandker. We would like to thank, Catherine Bodart, Carlos Riano Paramo, and the UN-REDD and FMA team members for their comments and insights on the document.

Acronyms

AD	activity data
aoi	area of interest
API	application programming interface
CI	confidence interval
CRS	coordinate reference system
EF	emission factor
F	forest
FAO	Food and Agricultural Organization of the United Nations
GE	Google Earth
GEE	Google Earth Engine
GFC	Global Forest Change dataset
GFOI	Global Forests Observation Initiative
GIS	geographic information system
IPCC	Intergovernmental Panel on Climate Change
Landsat	Land Satellite (US satellite series)
MMU	minimum mapping unit
NF	non-forest
OSS	open source software
PA	producer's accuracy
QGIS	Quantum GIS open source software
REDD+	reducing emissions from deforestation and forest degradation in developing countries; and the role of conservation, sustainable management of forests and enhancement of forest carbon stocks in developing countries
SRTM DEM	Shuttle Radar Topography Mission Digital Elevation Model
TOA	top of the atmosphere
TOC	table of contents
UA	user's accuracy
UN	United Nations
UN-REDD	The United Nations Collaborative Programme on Reducing Emissions from Deforestation and Forest Degradation
UNFCCC	United Nations Framework Convention on Climate Change

Introduction

Accurate and consistent information on forest area and forest area change is important given the reporting requirements for countries to access results based payments for REDD+[1] . Forest area change estimates usually provide data on the extent of human activity resulting in emissions (e.g. from deforestation) or removals (e.g. from afforestation), also called activity data (AD). A basic methodological approach to estimate greenhouse gas emissions and removals (IPCC, 2003), is to multiply AD with a coefficient that quantifies emissions per unit 'activity' (e.g. tCO_2e per ha), also called an emission factor (EF).

Activity data as part of emission/removal estimates should follow the IPCC good practice principle of neither over- nor underestimating emissions/removals and reducing uncertainties as far as is practicable. Uncertainty (lack of knowledge of the true value) is related to two issues: accuracy and precision (see Figure 1). Accuracy is a relative measure of the exactness of an estimate and accounts for systematic errors also referred to as bias. Therefore, an accurate estimate does not systematically over- or underestimate the true value. Map accuracy can be quantified by creating an error matrix (also commonly called a confusion matrix), which compares the map classification with a reference classification. Precision is related to the random error (see Figure 1), which can be quantified by a confidence interval. A confidence interval gives a range that encloses the true value of an unknown fixed quantity with a specified probability. A precise estimate would thus have a small confidence interval. Basic guidance on reducing uncertainty can be found in GFOI (2013).

---

[1] UNFCCC, 2013. Decision 14/CP.19, available at http://unfccc.int/resource/docs/2013/cop19/eng/10a01.pdf#page%3D39

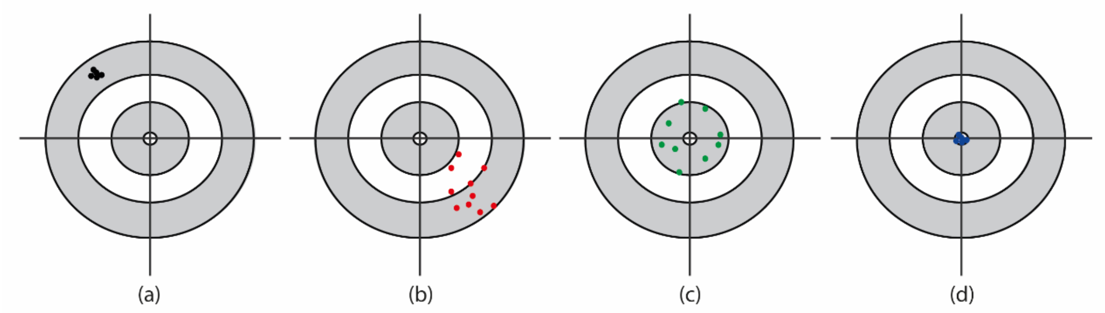


Figure 1: Illustration of accuracy and precision: (a) inaccurate but precise; (b) inaccurate and imprecise; (c) accurate but imprecise; and (d) precise and accurate (source: IPCC 2006)

The accuracy assessment of map data described in this document will provide:

1. a quantification of the map accuracy through the creation of error matrices;
2. area and area change estimates that are adjusted for the map bias, thus more accurate estimates;
3. and a quantification of the precision of the area and area change estimates through the calculation of confidence intervals.

The underlying principle of the accuracy assessment is that it compares the mapped land classification to higher quality reference data, collected through a sample based approach. The higher quality reference data can be obtained through ground collected data, but as this is expensive and labor intense it is more commonly obtained through satellite imagery or aerial photography with finer spatial resolution than the data that was used to create the map data. When relying on imagery for reference data and there is no high resolution imagery available higher quality data can be collected using a process considered more accurate, such as human interpretation of the reference data.

This document uses the publication *Good practices for estimating area and assessing accuracy of land change,* (Olofsson et al. 2014) as a framework to provide recommendations for designing and implementing an accuracy assessment for land cover and land cover change maps, and for estimating area based on the results from the accuracy assessment. This guide covers the theoretical background and recommendations for the setup of the sampling design, reference data collection and the analysis of the results.

The scope of this document

This document provides the methodology and practical implementation for the procedure for estimating area and assessing accuracy of a land cover map from a single period in time or for change between two time periods. The document guides the user through the aspects of an accuracy assessment which can be used to quantify and reduce uncertainty of map data for transparent reporting. The step-by-step guidance on how to implement such an assessment complements the theoretical background provided by Olofsson et al. (2014). Following the guidance of this document and the supplementary materials the user can produce tables of accuracy estimates, confidence intervals for area estimation, and comparison of area estimation derived from map data, reference data, and adjusted area estimates using both map data and reference data.

This document is split into two sections, the first covering the theoretical background for constructing an accuracy assessment and the second guiding the user through a practical implementation of an accuracy assessment of land cover and change maps. Step-by-step instructions rely on several open source software applications including *QGIS*, *R* and *Open Foris Collect Earth*. *R* is a free software programming language and software environment for statistical computing and graphics. *QGIS* is an open source desktop GIS application that provides data viewing, editing, and analysis capabilities. *Collect Earth* is a tool that facilitates sample-based assessments and collection of reference data from medium, high and very high spatial resolution satellite imagery in conjunction with Google Earth, Bing Maps and Google Earth Engine. Previous experience with these programs is recommended but not necessary. Experience with GIS, working with raster and point data, as well as in basic statistical analysis is required to complete the assessment.

Box 1: Land cover, land use and change Land cover, land use, and change are all different datasets. *Land cover* is the observed (bio)physical cover on the earth’s surface. *Land use* concerns the human- environment interaction and is characterized by the activities of people to maintain a certain land cover type. *Land use change or land cover change* is the quantitative change in land cover or land use types. This document describes the accuracy assessment of data created from satellite observations, therefore only refers to land cover and land cover change maps, although these methods can be applied to land use and land use change maps.

Part I

Theoretical background

1. | Overview of the steps of accuracy assessment

In an accuracy assessment of map data, the map is compared with higher quality data. The higher quality data, called reference data, is collected through a sample based approach, allowing for a more careful interpretation of specific areas of the map. The reference data is collected in a consistent manner and is harmonized with the map data, in order to compare the two classifications. The comparison results in accuracy measures and adjusted area estimates for each map category. This process is broken down into four major components: (i) a map, (ii) the sampling design (iii) the response design and (iv) the analysis (see Figure 2). The sampling design specifies how to select a subset of the map for which reference data will be collected. This is necessary because it is usually impractical to collect reference data for the entire study region (map area). The response design protocol provides guidelines for collecting the reference data. The analysis protocol includes all steps that lead to a decision whether the map and reference data are in agreement for the subset of the data that was sampled. How to derive accuracy and area estimates from the map and reference data is determined in the analysis protocol.

An accuracy assessment must provide full documentation of the map and reference data. To meet this prerequisite, this document first gives background information about the example map data used in the practical implementation, the Global Forest Change (GFC) dataset (Hansen et al., 2013) and about how to use other land cover data sets. The document explains how to define strata and calculate the area occupied by each stratum for the map data. This information is necessary for the calculation of the sample size that forms part of the sampling design. The response design chapter provides documentation of how the reference data can be obtained using sample-based approaches and the *Open Foris Collect Earth* application and alternatively, how to use other sources for reference data.

The steps of the accuracy assessment are outlined in Figure 2, including the software applications that can be used to complete the steps.

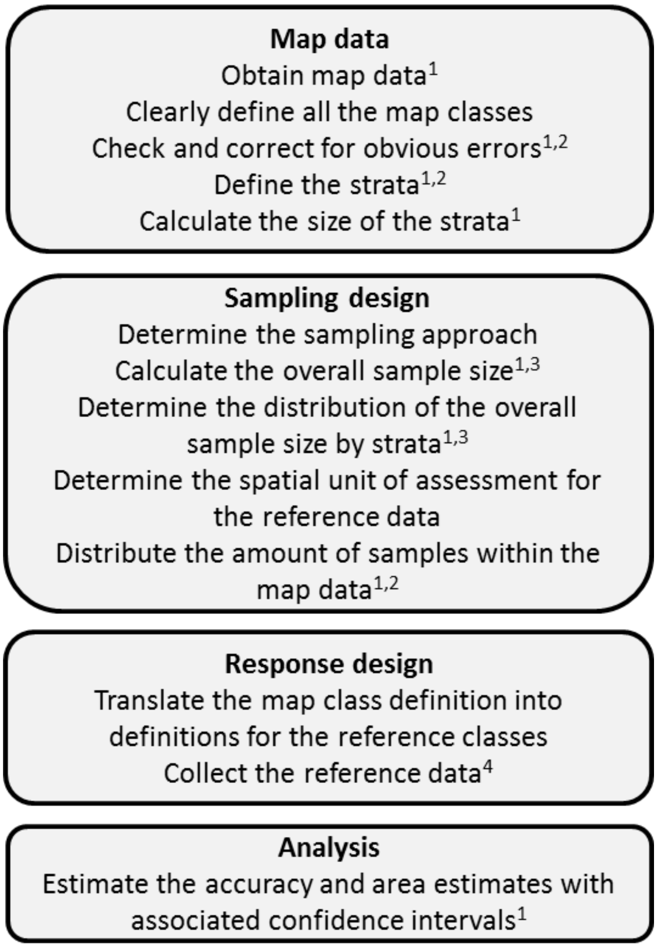


Figure 2: The main four steps of an accuracy assessment: Obtaining and finalizing the map data, sampling design, response design and analysis. The symbols show the software that can accomplish that step: ^{1} is *R*, ^{2} is *QGIS*, ^{3} is Excel, and ^{4}& is *Collect Earth*.

2. | Map data

Users may want to implement accuracy assessment for different types of thematic map data. Map data can be made from satellite images or be freely available map data for land cover or land use for a single time period or change between multiple time periods. The characteristics of the map determine some part of the methodology used for the accuracy assessment, e.g. the minimum mapping unit and the classes to be assessed.

The first step is a general quality control check of the map data. The user can do this by visually assessing the map for obvious errors. A preliminary analysis of comparing other similar data can help reveal map errors. If a map of change is being assessed, a method for checking the quality of the map is checking for impossible transitions, i.e., water to forest. Obvious errors should be accounted for and corrected before continuing with the accuracy assessment. The quality control check does not need to be an independent assessment; rather it can be beneficial if the map producer carries out the quality control check.

In the general preparation of the map data the user can consider aggregation. Aggregation of the map classes can reduce the burden during the collection of reference data and can increase the accuracy measures. Aggregation of the spatial unit from pixels to pixel blocks can also be considered. Aggregation can be considered on case by case basis and the justification for aggregating map classes and the spatial unit varies.

The accuracy of custom land cover maps can be assessed using the steps described in this document. The appropriate methodology depends on the type of map data that is to be used. A map can be either raster or vector data; in the former the spatial assessment unit is usually a pixel or pixel block and in the latter the spatial assessment unit is a polygon.

Raster data can be either pixel or object-based. The example dataset used in this document is the Global Forest Change (GFC) data set. If the data being assessed is pixel-based, the same steps described in the practical implementation can be similarly implemented because the GFC data is a pixel based analysis. If it is object-based (or segment-based), estimating accuracy and area is more complex and is not covered by this document.

Vector data can be produced by visual interpretation of pixel-based satellite imagery, or also from segmented satellite data. If the visual interpretation is based on pixel-based satellite imagery, the resulting land cover map can be converted to raster data using the spatial resolution of the original satellite data (i.e. 30 m for Landsat data). Then the same methodology as for the GFC data can be applied. The accuracy assessment of object-based vector data is not covered by this document.

Once the custom land cover or land cover change map is available as raster data, the strata need to be defined and the size of each stratum needs to be calculated. The strata must be mutually exclusive, meaning that each pixel must be assigned to one strata class. The sum of all pixels in the strata defines the total study area. The calculation of the strata size can be done using GIS software or an *R* script. The process is the same for multiple change classes, i.e. between forest, woodland and cultivated land, as it is for changes between forest and non-forest. However, a high accuracy of many different change classes is increasingly difficult to achieve, and the assessment of many classes requires an increase in the amount of sample points.

3. | Sampling design

The sampling design defines how to select the subset of the map, which forms the basis for the accuracy assessment. Selecting a subset of the map is necessary because (i) a sampling approach allows more careful interpretation of the parameters of interest at each sample site thus satisfying the requirement of using 'higher quality' data than that used to create the map, even if the data used to make the map is also used in the accuracy assessment and (ii) it is usually not feasible to collect reference data for the whole study area. 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).

It is critical to use a probability sampling design that incorporates randomization in the sample selection protocol. Probability sampling is defined in terms of inclusion probabilities that quantify the likelihood of a given unit being included in the sampling design. The inclusion probability must be known for each unit selected in the sample and it must be greater than zero for all units in the area of interest. Non-response is the situation where the inclusion probability is unknown or zero, i.e. inaccessible plots or unavailable data due to cloud coverage. The circumstances must be clearly stated for non-response samples, for example by reporting the proportion of the selected sample units for which cloud cover or lack of reference imagery prevented assessment of the unit. If ground visits are used to collect reference data, a sampling design which considers non-response is advisable such as the protocol described by Stevens and Olsen (2004).

Commonly used probability sampling designs include simple random, stratified random, and systematic. For land cover maps it is recommended to use a **stratified sampling approach**, so only this sampling design is being addressed in this document. The sampling design can influence the results and information about it is necessary to properly interpret the error matrix.

1.1 Stratification of the map

Stratification is the division of the area of interest into smaller areas (strata), in which each assessment unit is assigned to a single stratum. Stratification could be for example by map class (i.e. forest and non-forest) or by sub-region (i.e. administrative units). The strata need to be mutually exclusive and inclusive of the entire study area, with no area that is in multiple strata classes or is omitted from the strata. The end use of the map also needs to be considered when creating the sampling design, i.e., a national forest change map that is being used to derive area estimation of forest change in different sub-national areas. In this case, the map is stratified by map class and administrative boundaries. The user should ensure the sampling design captures all of the strata.

There are two main purposes for stratification. Firstly strata can be of interest for reporting results, i.e. accuracy per land cover class or sub-region. The second purpose of stratification is to ensure a sufficient representation of rare classes (e.g. that only represent a small proportion of the area of interest). Land change often occupies a small fraction of the landscape, so a change stratum (i.e. forest loss) can be identified and the sample size allocated to that stratum can be large enough to produce a small standard error for the user's accuracy estimate. The stratification by map classes improves the precision of the accuracy and area estimates by increasing the sampling density in the change classes. For this reason, stratification in this study is based on land cover class and an independent sample is drawn for each land cover class.

When defining the strata, a feasible number of classes need to be chosen. For single date land cover maps, it is usually feasible to define a stratum for each map class (Wulder et al., 2007), but it is more challenging for a change map where the number of different types of changes might be too high. To reduce the number of strata, types of change that are very unlikely to occur could be eliminated. Strata could also be defined on the basis of generalized change categories, such as change from forest to non-forest instead of forest to cultivated land, forest to water etc. The feasibility of distinguishing these change classes in the reference data should also be taken into account. Strahler et al. (2006) provides additional examples for aggregating change classes. Even if a change type is not defined as stratum in the sampling design, accuracy and area estimates can still be derived for that change type, but the sample size might not be high enough to derive estimates at the desired precision.

1.2 Random versus systematic sampling

The two most common protocols for selecting the assessment units are simple random and systematic sampling. Systematic sampling is defined as selecting a starting point at random with equal probability and then sampling with a fixed distance between sampling locations. It is often implemented for field sampling activity, such as national forest inventories. In general, the simple random selection protocol is the recommended option, but systematic selection is also nearly always acceptable. If using simple systematic sampling it can be difficult to capture small classes, particularly change categories. Therefore when conducting an accuracy assessment for land cover change (such as for activity data) which includes the collection of reference data, it is recommended to use a stratified approach.

1.3 Determine sample size

The sample size should be representative of the population, the number of spatial units (i.e., pixels), making it large enough to get reliable estimates, but as small as possible in order to save costs. Determining this sample size is an inexact science because it depends on accuracy and area information that is not known prior to the assessment. A "best guess" about the accuracy and area information can be used for the sample size calculation. Although there are formulas which can calculate the overall sample size and distribution of the sample it is up to the user to decide how to best determine the sample size. Area information is usually based on the number of spatial units of the map, and accuracy is generally higher for larger classes, i.e. the estimated accuracy is higher for no-change classes than for change. Equation 1 (Cochran, 1977), 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), $S(\hat{O})$ 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*, and S_i is the standard deviation of stratum *i*.

$$n = \frac{(\Sigma \mathrm{W_i\,S_i})^2}{[\mathrm{S}(\widehat{\mathrm{O}})]^2 + (1/\mathrm{N})\Sigma \mathrm{W_i\,S_i}^2} \approx \left(\frac{\Sigma \mathrm{W_i\,S_i}}{\mathrm{S}(\widehat{\mathrm{O}})}\right)^2 \qquad (1)$$

The overall sample size resulting from this calculation can be allocated among the stratum in multiple ways. The samples need to be distributed between the strata balancing between equal sample size per stratum and proportional allocation. In proportional allocation, the overall sample size is allocated to the strata proportional to the area of the strata, so rare strata receive a small proportion of the overall sample size. In equal allocation, the overall sample size is distributed equally between the strata. Stratification is used for rare classes, such as the assessment of change, it is necessary to ensure there are a sufficient number of samples in the rare classes. Minimum sample size should be at the least 20 to 100 samples per strata (Congalton and Green, 2008).

Different allocations favor different estimation objectives, i.e. equal sample size favors estimation of user's accuracy, while proportional allocation usually results in smaller standard errors for producer's and overall accuracy. As a compromise, it is suggested to use a sample allocation somewhere in between same and proportional allocation, taking into account a minimum sample size per stratum.

2 | Response design

The response design defines how to determine whether the map and the reference data are in agreement. The response design establishes reference data sources to be compared with map data, assuming that the reference classification is sufficiently more accurate than the map classification being evaluated. The four major features in the response design are: the spatial unit, the sources of information used to determine the reference classifications, the labeling protocol for the reference classification, and the definition of agreement.

2.1 Spatial assessment unit

The spatial assessment unit is the unit at which the map was sampled. This is the unit at which the reference data is collected. The spatial unit is the basis for the location-specific comparison of the reference classification and map classification. It can be a pixel, polygon (segment), or pixel block. Usually, the pixel is chosen as the spatial unit, but any of the types can be used. If a pixel block is used, the map is coarsened to the size of the pixel block. The spatially explicit character should be retained in the accuracy assessment; therefore, the user should aim to have reference data with the same or finer level of detail. The choice of the spatial assessment unit has implications on the sampling design and analysis.

2.2 Sources of reference data

Various sources of reference data exist, ranging from ground visits to the use of satellite imagery. The reference classification needs to be of higher quality compared to the map classification which can be ensured in two ways: i) The reference source has to be of higher quality (e.g. higher spatial or radiometric resolution) than the source for the map classification, or ii) the process to create the reference classification has to be more accurate if using the same source material. For example, if Landsat imagery is the only available source for both map classification and reference data, the process for obtaining the reference classification has to be more accurate than the map classification, i.e. by including visual assessment from expert users. Olofsson et al. (2014) list various sources of reference data, and elaborate on their advantages and disadvantages. Additionally reference data should be temporally coincident with the map being assessed, e.g. if a land cover map of the year 2000 is being assessed, then the reference data should be from the year 2000. If reference data is collected from a year different than the year of the map, then adjusted areas will represent areas as of the time of the reference data.

A cost-effective tool for collecting reference data from very high, high and medium resolution satellite imagery is Collect Earth[2]. This Google Earth plugin allows the practitioner to visually assess the land cover/use of sample locations with the freely available data from Google Earth, Google Earth Engine, Here maps, and Bing maps. Chapter 10 addresses the setup of Collect Earth incorporating the major features of the response design.

2.3 Reference labeling protocol

The labeling protocol defines how to convert the information provided by the reference data into labels of the reference classification. When reference data contains a mixture of classes, it is especially to have a consistent approach for reference class labeling.

The specification of a minimum mapping unit (MMU) for the reference classification should be defined by the labeling protocol because it has important implications for the accuracy assessment and area estimations. The MMU is the smallest area that can receive a classification label in the reference data. A possible MMU is the spatial unit of the sample; however it is not necessary that the MMU for the reference data matches the spatial unit of the map. For example a smaller MMU for the reference data increases the specificity of the classification granting the ability to distinguish smaller patches of change for the reference classification.

Olofsson et al. (2014) provide more detailed suggestions for dealing with mixed reference units, and the practical section on reference data collection, section 10, addresses these suggestions with Collect Earth.

2.4 Defining agreement

After the map and reference classification for a given spatial unit have been obtained, rules for defining agreement need to be set up. In the simplest case, map and reference classification have the same classification scheme and if the labels agree, the map class is correct;

---

[2] http://www.openforis.org/tools/collect-earth.html

otherwise it is a misclassification. Defining agreement is more complicated for heterogeneous assessment units or different classification schemes. A heterogeneous assessment unit is a spatial unit covers that more than one class, such as a pixel block that is 60% non-forest and 40% forest. As for all steps, the rules for defining agreement need to be clearly stated.

3 | Analysis

The analysis protocol specifies how to translate the information contained in the comparison of map and reference data into accuracy and area estimates, and how to quantify the uncertainty associated with them. Most of the calculations are based on the error matrix (also commonly called a confusion matrix), which contrasts the map and reference classification.

This chapter gives an introduction about the error matrix, the measures used to summarize the accuracy assessment, and about estimating area.

3.1 The error matrix

The error matrix is a cross-tabulation of the class labels allocated by map and reference data. It is derived as a q x q matrix with q being the number of classes assessed. The elements show the number of data points which represent a map class i and reference class j (n_{ij}). Usually the map classes are represented in rows and the reference classes in columns. The diagonal of the matrix contains the correctly classified data points, whereas the cells off the diagonal show commission and omission errors. Commission error is the complimentary measure to user's accuracy, calculated by subtracting 100% from the user's accuracy for each class. Commission error, calculated for each of the map classes, is the probability that the spatial unit classified into a given category on the map represents that category in the reference data. Omission error is the complimentary measure to producer's accuracy, calculated by subtracting 100% from the producer's accuracy for each class. Omission error, calculated for each of the map classes, is the probability that the spatial unit classified into a given category in the reference data represents that category in the map data.

The sample based absolute counts, n_{ij}, can be converted into estimated area proportions $\hat{p}_{ij}$ (see table 1) with equation 2 when the strata correspond to the map classes if simple random, simple systematic or stratified random sampling are used

$$\hat{p}_{ij} = W_i \frac{n_{ij}}{n_{i\cdot}} \qquad (2)$$

W_i is the proportion of area classified as class i and can be calculated by dividing the number of pixels per stratum as derived in section 4.3 by the total number of pixels. Table 1 shows an example of an error matrix with four classes.

Table 1: Population error matrix of four classes. Cell entries (p_{ij}) represent proportion of area.

		Reference				
		Class 1	Class 2	Class 3	Class 4	Total
Map	Class 1	p_{11}	p_{12}	p_{13}	p_{14}	$p_{1.}$
	Class 2	p_{21}	p_{22}	p_{23}	p_{24}	$p_{2.}$
	Class 3	p_{31}	p_{32}	p_{33}	p_{34}	$p_{3.}$
	Class 4	p_{41}	p_{42}	p_{43}	p_{44}	$p_{4.}$
	Total	$p_{1.}$	$p_{.2}$	$p_{.3}$	$p_{.4}$	1

3.2 Estimating accuracy

The accuracy measures are derived from the error matrix and reported with their respective confidence intervals. They include overall accuracy, user's accuracy and producer's accuracy.

The overall accuracy is the proportion of area classified correctly, and thus refers to the probability that a randomly selected location on the map is classified correctly (see equation 3). User's accuracy is the proportion of the area classified as class i that is also class i in the reference data (see equation 4). It provides users with the probability that a particular area of the map of class i is also that class on the ground. Producer's accuracy is the proportion of area that is reference class j and is also class j in the map (see equation 5). It is the probability that class j on the ground is mapped as the same class.

$$A = \sum_{j=1}^{q} p_{jj} \qquad (3)$$

$$U_i = p_{ii}/p_{i\cdot}. \qquad (4)$$

$$P_j = p_{jj}/p_{\cdot j} \qquad (5)$$

For all three accuracy measures, the confidence intervals need to be derived as well. The formula for the variance are presented in equations 5, 6 and 7 in Olofsson et al. (2014), and the 95 % confidence interval can be calculated by multiplying the square root of the variance by 1.96.

The kappa coefficient is also often reported as a measure of map accuracy. However, its use has been questioned by many articles and is therefore not recommended (Pontius Jr and Millones, 2011).

3.3 Estimating area

The error matrix provides information on the accuracy of the map. It is also recommended using the information for estimating the area of classes, such as the area of deforestation, and their standard errors. The reference data can be used to adjust the area estimate as obtained from the map. It is recommended to base that estimation on $p_{.\,k}$, the proportion of area derived from the reference classification, because in contrast to $p_{k\,.}$, the proportion mapped as class k, it should have smaller bias. $p_{.\,k}$ is the column total of reference class k in the error matrix (see equation 6).

$$\hat{p}_{\cdot k} = \sum_{j=1}^{q} \hat{p}_{ik\cdot} \qquad (6)$$

The standard error for the stratified estimator of proportion of area can be calculated using equations 10 and 11 in (Olofsson et al., 2014), and the 95 % confidence interval is obtained by multiplying the standard error by 1.96.

4 | Interpretation of the results

The main purpose of the accuracy assessment is to quantify the accuracy of the map and to generate new area estimates to correct for bias in the map. For both accuracy and area estimates, the accuracy assessment provides confidence intervals. This section has a closer look at what these estimates mean, and what needs to be taken into account when reporting the results of the accuracy assessment.

4.1 Interpretation of the accuracy estimates

There is no general rule as to which level of accuracy is good and which is not. Judgment on the data validity depends on the purpose of the map and thus needs to be dealt with on a case by case basis. The steps of the accuracy assessment, described in this document, need to be considered when assessing the accuracy value. Accordingly, UNFCCC does not provide any thresholds for the accuracy of data provided for the construction of forest reference levels.

For land cover change, it is necessary to look at the accuracy of change and not at the accuracy of two single land cover maps. Even if both land cover maps have high accuracy measures for a single point in time, it does not provide any information about the accuracy for change classes. A new change analysis using remote sensing images is necessary rather than comparing maps from different times. This is because change usually occupies a small portion of an area and is frequently smaller than the cumulative error of the individual map productions (GFOI, 2013). Forest change is often less than 1% of the total area; in two land cover maps with an overall accuracy of 99%, the change can be attributed to error between the two land cover maps.

The overall map accuracy is not always representative of the accuracy of individual classes (GFOI, 2013). High overall map accuracy does not guarantee high accuracy for forest loss. Therefore, both producer's and user's accuracy for all single classes need to be considered. A high user's accuracy and low producer's accuracy for forest loss, for example, indicate that most of the forest loss in the map was also forest loss in the reference data, but that the map missed catching a fair amount of forest loss. Additionally total sample size, the number of strata and the allocation of the total sample size to the strata can favor one accuracy measure over the other.

Accuracy is usually higher for stable classes than for change classes. Furthermore, accuracy is variable in different landscapes. Global products, like the GFC data, need to be assessed for each study region instead of relying on global accuracy estimates. For example Potapov et al., (2014a) opted not to use the global forest change classification model because it had a conservative estimate of forest loss. For a study area in Eastern Europe, Potapov et al. (2014a), reports for forest loss between 2000 and 2012, the GFC data has a user's accuracy of 65 % and a producers accuracy of 68 % while their customized classification model products had higher accuracy measures of 94 % user's accuracy and 88 % producer's accuracy.

4.2 Interpretation of the area estimates

The accuracy assessment serves to derive the uncertainty of the map area estimates. Whereas the map provides a single area estimate for each land cover class without confidence interval, the accuracy estimates adjusts this estimate and also provides confidence intervals as estimates of uncertainty (Figure 3). The adjusted area estimates can be considerably higher or lower than the map estimates.

Such area estimates with confidence intervals could also be derived from the reference data alone, but the combination of map and reference data increases the precision of the final estimate (Figure 3). A higher precision means that the confidence intervals are smaller. Therefore, it is highly recommended to use this combination of map and reference data for area estimates. Even from a map with low accuracy measures, meaningful and precise area estimates can be derived if the reference data is collected in a thorough way. The adjustment of the map area, whether the adjusted area is greater or smaller than the map area, can be comprehended by comparing the error matrix which shows the sample data. The map area for a particular stratum is adjusted to a larger area when the number of reference units sampled is greater than spatial map units for that stratum. For example if 1000 units are sampled, and within forest strata for the selected sample, the map data has 100 spatial units labeled as forest and the reference data has 150 spatial units labeled as forest, the resulting adjusted area is greater than the area only derived from the map data.

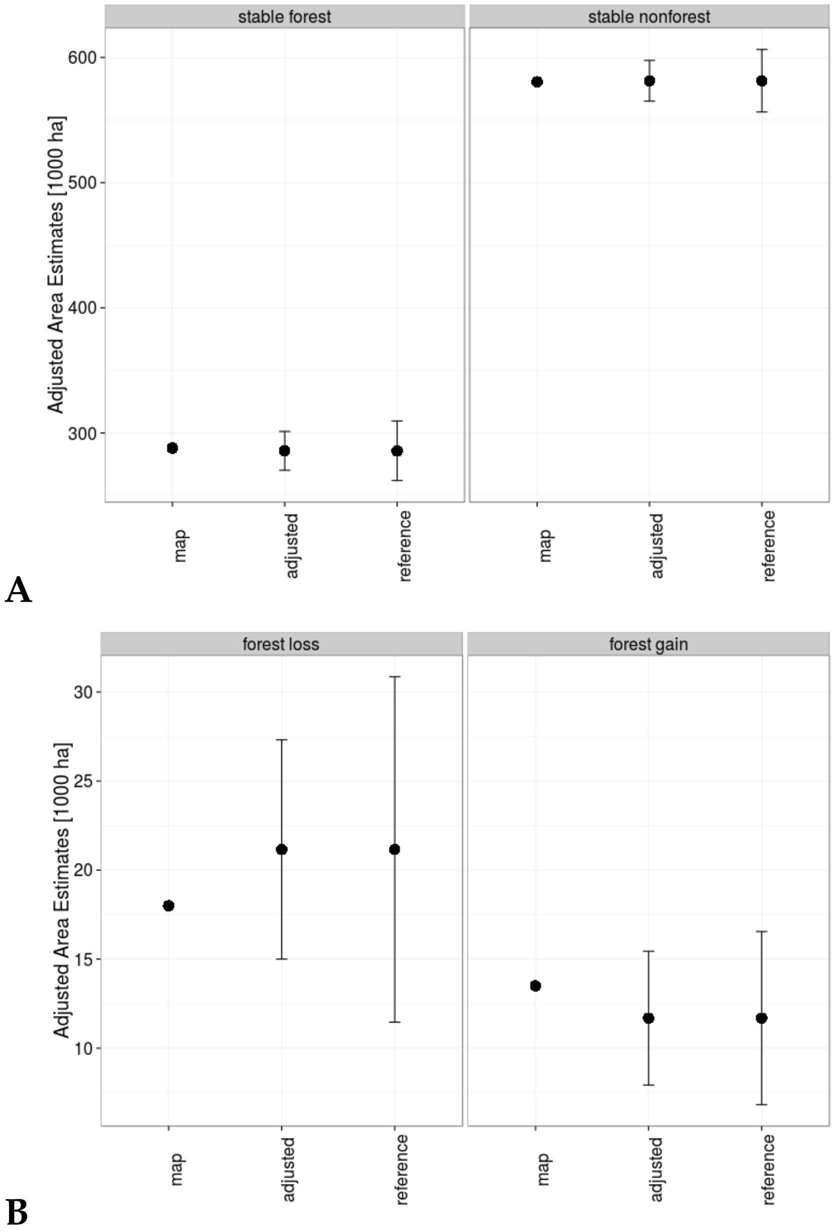


Figure 3: Graphs A and B show the area estimates from the map data alone (map), the combination of map and reference data (adjusted), and reference data alone (reference) for each of the four strata. Graph A shows the area estimates for the stable classes, forest and non-forest and graph B shows the area estimates for the non-stable classes, forest loss and forest gain. Each of the points includes confidence intervals. The confidence intervals are larger for the non-stable classes because they cover a smaller area. The map data does not have confidence intervals because it is represents the entire population that is being sampled (all of the pixels in the map). The *R* script will output this graph in addition to the values of the areas and confidence intervals.

4.3 Reporting results

When reporting the results of accuracy assessment, the report should not only include the estimates of accuracy assessment, adjusted area and their respective confidence intervals but also the assumptions implied by several elements of the accuracy assessment. The assumptions can influence the level of accuracy, and include, but are not limited to:

1. the minimum mapping unit and the spatial assessment unit
2. the sampling design
3. the forest definition
4. the source of reference data
5. the confidence level used for calculating the confidence intervals (typically 95 %).

The estimates always need to be reported with their respective confidence intervals. Additionally, it is recommended to present the error matrix in terms of estimated area proportions instead of absolute sample. The estimated area proportions normalize the absolute sample counts by the map area and are used to calculate the users and producer's accuracy. Because the producer's accuracy is based on the map data, it can be calculated using either the estimated area proportions or the sample counts; however, the user's accuracy is based on the reference data and yields different results whether the calculation uses sample counts or estimated area proportions. It is recommended to calculate the accuracy measures based on the estimated area proportions and not the absolute sample counts, therefore showing only the sample counts does not explicitly demonstrate how the user's accuracy is calculated.

Part II

Practical implementation

Introduction to the practical implementation

The practical implementation gives instructions on how the different steps of the accuracy assessment can be achieved. These are only suggestions and the same can also be achieved using other tools and methodologies. This implementation includes instructions for downloading a global data set for any area of interest, sampling the data set, collecting reference data, and producing accuracy and adjusted area estimates.

5 | Obtaining map data

5.1 The Global Forest Change Data set

The Global Forest[3] Change (GFC) data is used to exemplify a dataset that provides land cover and land cover change data. The GFC product provides estimates of global tree cover and tree cover changes on an annual basis from 2000 through 2013 based on Landsat satellite imagery (Hansen et al., 2013). The study area includes all global land except for Antarctica and a number of Arctic islands. Trees are defined as all vegetation taller than 5 m in height. Tree cover loss is defined as stand replacement disturbance, and gain was defined as inverse of loss.

The global dataset resulting from this study is available through Global Forest Change 2000 – 2013 data download[4]. The data is divided into 10x10 degree tiles. Each tile consists of seven files per tile:

- Tree canopy cover for year 2000 (treecover2000):
  Tree cover in the year 2000, defined as canopy closure for all vegetation taller than 5 m in height. Encoded as a percentage per output grid cell, in the range 0 – 100.
- Global tree cover loss 2000 – 2013 (loss):
  Tree cover loss during the period 2000 – 2013, defined as a stand-replacement disturbance, or a change from a tree cover to non-tree cover state. Encoded as either 1 (loss) or 0 (no loss).
- Global tree cover gain 2000 – 2013 (gain):
  Tree cover gain during the period 2000 – 2013, defined as the inverse of loss, or a non-tree cover to tree cover change entirely within the study period. Encoded as either 1 (gain) or 0 (no gain).

---

[3] ”Forest” refers to tree cover, not land use

[4] http://www.earthenginepartners.appspot.com/science-2013-global-forest/download.html

- Year of gross tree cover loss event (lossyear):
  A disaggregation of total tree cover loss to annual time scales. Encoded as either 0 (no loss) or else a value in the range 1 – 12, representing loss detected primarily in the year 2001 – 2013, respectively.
- Data mask (datamask):
  Three values representing areas of no data (0), mapped land surface (1), and permanent water bodies (2).
- Circa year 2000 Landsat 7 cloud-free image composite (first):
  Reference multispectral imagery from the first available year, typically 2000. If no cloud-free observations were available for year 2000, imagery was taken from the closest year with cloud-free data, within the range 1999 – 2012.
- Circa year 2013 Landsat cloud-free image composite (last): Reference multispectral imagery from the last available year, typically 2013. If no cloud-free observations were available for year 2013, imagery was taken from the closest year with cloud-free data, within the range 2010 – 2012.

All files contain unsigned 8-bit values and have a spatial resolution of one arc-second per pixel, or approximately 30 m per pixel at the equator. For most purposes, only the first five layers are needed. The Landsat cloud-free image composites are furthermore very large, so it is not recommended to download them unless specifically needed.

To convert tree cover data from this product into approximate forest cover data in line with the country's forest definition, the product will most likely need to be modified. For example, pixels may be grouped to approximate the minimum area in the forest definition and a canopy cover threshold may be applied. In some locations the direct application of the tree cover threshold in the forest definition may give different forest area estimation than in the forest definition. Differing tree cover thresholds can greatly impact the forest extent estimates (Potapov et al., 2014b).

5.2 Downloading the global forest change data

The global forest change (GFC) data is freely available for almost all forested regions in the world5. The specifications of the GFC dataset are described in section 3.1. The data can be obtained in multiple ways including: i) using the user interface for downloading, clipping and mosaicking, ii) through command line or iii) manually through the internet. The user interface[5] runs using the R package, 'gfcanalysis' and 'shiny.' First R must be downloaded (for further information about R and how to install it, see Appendix A.1.1).

Downloading the GFC data and accompanying instructions can be found here[5]. To complete the download via command line in R, the sample script analyze_GFC.R[6] provides an example of how to use the gfcanalysis package. A customized version of this script for the accuracy assessment can be found here[7].

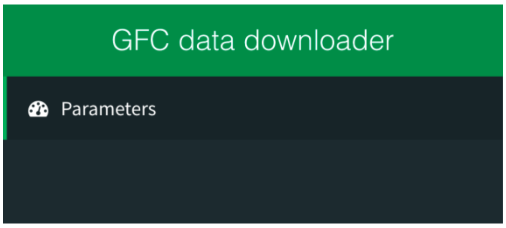


Figure 4: User interface[5] for downloading GFC data

Using the user interface[5] to download the GFC data for a specific study area is highly recommended because it automates the merging, stacking, and clipping of the GFC data to the specified area of interest. However, it takes a considerable amount of computing power and time, especially for large study areas, such as country boundaries.

The manual download can be done using the GFC download[8]. For the manual download the user must download each tile and data layer individually, then merge and clip the layers to the study area. This can be accomplished using programs with GIS capability such as QGIS, however the process is time consuming and tedious.

---

[5] https://github.com/openforis/accuracy-assessment

[6] https://raw.githubusercontent.com/azvoleff/gfcanalysis/master/inst/examples/analyze_GFC.R

[7] https://github.com/openforis/accuracy-assessment/blob/master/gfcdownload.R

[8] http://www.earthenginepartners.appspot.com/science-2013-global-forest/download.html

5.3 Define the strata

The GFC data is divided into four strata: stable forest, stable non forest, forest loss, and forest gain. The layers necessary to determine these strata include treecover2000, loss, gain, datamask, and aoi.

- **Stable forest** is determined by applying a user defined threshold to the treecover2000 layer. Stable forest excludes areas outside the data mask and water bodies, forest loss and forest gain between 2000 to 2013.

- **Stable non forest** is similarly determined by applying a threshold to the treecover2000 layer. The areas below the forest threshold are considered non forest. Stable non forest excludes areas outside the data mask and water bodies, forest loss and forest gain between 2000 and 2013.

- **Forest loss** is determined by all pixels within the data mask (excluding water bodies) that show loss and no gain and are forest in 2000.

- **Forest gain** is determined by all pixels within the data mask (excluding water bodies) that show gain and no loss and are nonforest in 2000.

5.4 Calculate the size of each stratum

The size and proportion of each stratum is calculated to determine the sample size. The size of each stratum is calculated using the pixel as the unit. The user interface[5] can be used to find the area of each stratum from the GFC data.

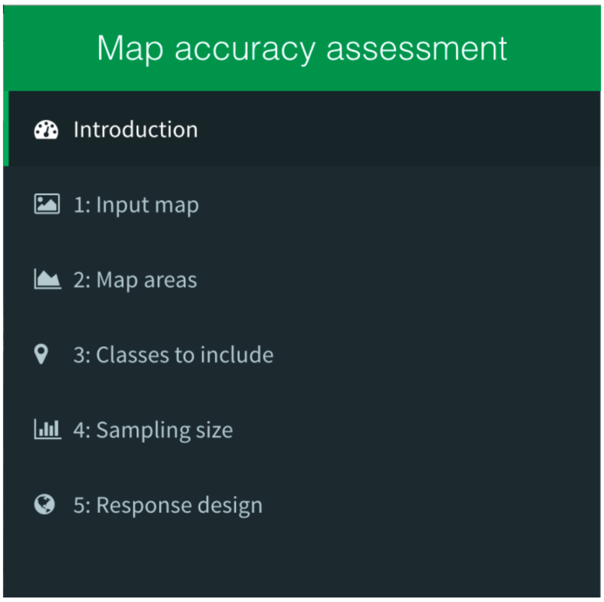


Figure 5: User interface[5] to create a probability sample from a thematic map.

6 | Sampling design

The sampling design requires a probability sampling design. First, the user determines the stratification of the map data, described in Section 8.1. Next, the sampling protocol is determined; in this document, the sampling protocol is random stratified sampling. The sampling design chapter takes the user through calculating the sample size, deciding the sampling spatial unit, and allocating the sample points on the map data. The sampling design can be completed in the user interface[5], which guides the user through each of the steps to create the samples.

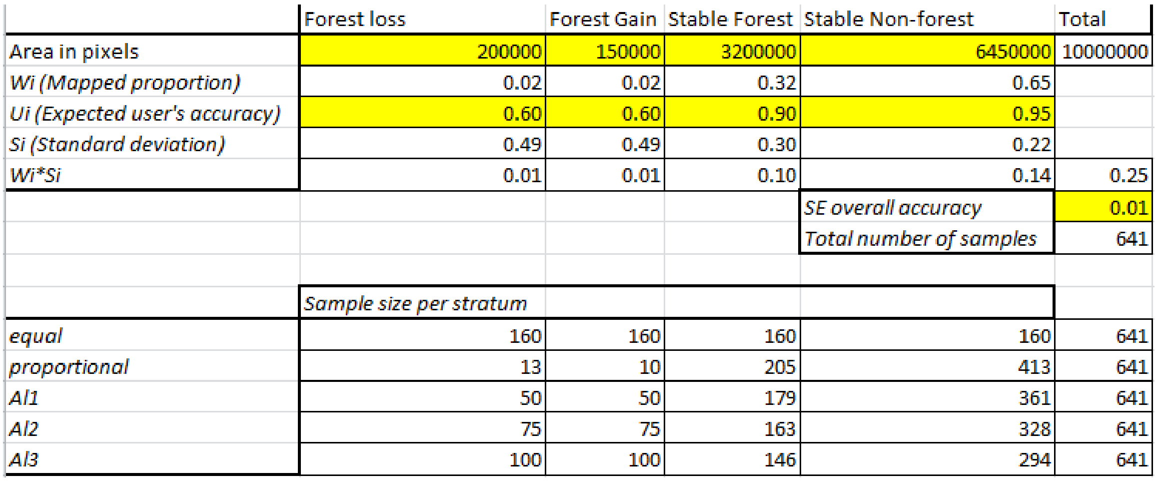
6.1 Calculating the sample size

The sample size is calculated using the formula 1 implemented the user interface[5]. The calculation can also be explored hands on in the spreadsheet[9] shown in Figure 6, which can be accessed with the spreadsheet application of choice (e.g. Excel or LibreOffice Calc). The only necessary user input in this spreadsheet is the area of each stratum in the map. The values for target standard error, suggested by Olofsson et al. (2014), are used for the overall accuracy and user's accuracy. The target standard error for overall accuracy is 0.01 while the user's accuracy for the change classes, the target standard error is lower because stable classes are known to be more accurate.

An example sample size calculation is shown in Figure 6. The area of each stratum for the area of interest is entered into the yellow fields in addition to the overall desired accuracy. The total number of samples based on the total area is 641. The total number of samples is then distributed by stratum equally, proportionally, and three options balancing between these allocations. The change classes, forest loss and forest gain, occupy a very small proportion of the total map area, making it necessary to have an allocation with a minimum sample size to adequately sample these strata. The three 'Al' allocations all have minimum sample sizes, for example the second allocation option (Al2) has a minimum sample size of 75 per stratum

---

[9] https://www.dropbox.com/s/wsihmlldebjc024/sample_size_stratified_simple_random.xlsx?dl=0



	Forest loss	Forest Gain	Stable Forest	Stable Non-forest	Total
Area in pixels	200000	150000	3200000	6450000	10000000
Wi (Mapped proportion)	0.02	0.02	0.32	0.65	
Ui (Expected user's accuracy)	0.60	0.60	0.90	0.95	
Si (Standard deviation)	0.49	0.49	0.30	0.22	
*Wi*Si*	0.01	0.01	0.10	0.14	0.25
				SE overall accuracy	0.01
				Total number of samples	641
	Sample size per stratum				
equal	160	160	160	160	641
proportional	13	10	205	413	641
AI1	50	50	179	361	641
AI2	75	75	163	328	641
AI3	100	100	146	294	641

Figure 6: Spreadsheet to determine sample size. The required inputs are the map areas and the expected user's accuracy for the sample, which are entered in the highlighted cells. The remainder is calculated using the formulas from Olofsson *et al.* (2014)

6.2 Spatial unit

In this tutorial, considering the map data, the spatial unit for the sampling data can be either a pixel or a pixel block. The spatial unit is the how the map data and reference data are compared. In the sampling procedure, the spatial unit can either match the resolution of the map or can involve the aggregation of pixels to a pixel block. At this step it is necessary to think ahead to the size of the sampling unit that will be chosen for the response design. The spatial unit in the sampling design can match the spatial assessment unit in the response design to aid the user in defining agreement, however this is not necessary. It is helpful and can reduce labeling errors if the spatial assessment unit from which the sample points are derived matches the minimum mapping unit of the reference data. If the user decides to have the sample spatial unit as the pixel, no modification of the map data needs to take place. This example uses the pixel as the spatial unit; the size of the spatial unit is equal to the resolution for the GFC data, 30 m x 30 m.

Otherwise the user can decide to use the pixel block as the spatial unit. The pixel block could be favorable if the user wants to exclude small fragmented patches or if it is difficult to visually assess 30 m x 30 m pixels for the response design. The user can use a larger spatial assessment unit for the sampling design by coarsening the spatial resolution of the map data (i.e. a 3 x 3 pixel block, equivalent to 90 m x 90 m for the GFC data). In *R*, aggregation begins in the upper left corner of each of the map data layers, then calculates the mean for each pixel block, and creates a new

RasterLayer for each data layer with a lower spatial resolution (larger cells), as shown in Figure 7.

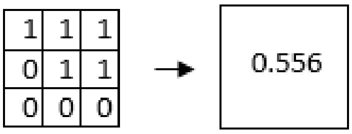


Figure 7: An example of aggregating 9 pixels to a pixel block. The original Boolean data is averaged to show the proportion of pixels in the pixel block.

If the map data is aggregated to pixel blocks the result is mixed pixels. A mixed pixel is a cell where the corresponding land area contains more than one class. For example the loss layer that was aggregated into a 3 x 3 pixel block, such as the illustration in Figure 7, originally contained 0 (absence of loss) and 1 (presence of loss), now has values of 0, 0.11, 0.22, 0.33, 0.44, 0.56, 0.67, 0.78, 0.89, 1 depending on the number of presence pixels in the pixel block. The user must decide on the threshold that determines which pixel blocks will be sampled. The threshold can be defined to include mixed pixel blocks, by defining the threshold as a value between 1 – 8 pixels within each block. Otherwise the threshold can be defined to only include pure pixel blocks, by defining the threshold as the maximum number of presence pixels (9 pixels) within the pixel block. For example pure loss pixel blocks could be used to define forest loss in the sampling for the reference data, therefore all pixel blocks with any non-loss pixels are excluded from the sample. Using pixel blocks could enhance the chance for the interpreter to detect the changes through visual assessment; however using pixel block is more methodologically complicated compared to using the pixel as the spatial unit.

6.3 Create random points

After the number of samples per strata has been established and, if desired, the aggregate layers have been created, the strata need to be randomly sampled. The map accuracy assessment designer automates the random selection of the specified amount of sample points.

The user interface[5] first assists in calculating the sample size, then randomly samples each of the map categories. The output of the user interface is a CSV file, which is formatted to be directly input into *Collect Earth*. The CSV file contains the specified amount of sample points for each stratum and for each randomly sampled pixel contains: the coordinates, the name of the administrative region and the country.

Optionally the random points can be created using GIS software such as *QGIS*. Appendix A.2 details how to create random points using *QGIS*.

In order to use these points in Collect Earth, they need to be in a CSV file containing the following columns separated by comma:

ID, YCOORD, XCOORD, ELEVATION, SLOPE, ASPECT, ADM1_NAME, COUNTRY, GFC_TREE_COVER_2000, GFC_FOREST_GAIN, GFC_FOREST_LOSS, GFC_FOREST_LOSS_YEAR, GFC_DATA_MASK.6

The ID needs to be a unique identifier and can be a string or a number. Coordinates need to be given in latitude-longitude-format (EPSG:4326). The first three columns, ID YCOORD XCOORD, need to be filled for Collect Earth to be able to use the data. All other columns are optional and can be filled with the value of 0 if there is no information, but can help the interpreter in assessing the land cover and land use of the points, or in analysing the data. Elevation, slope and aspect refer to the topography of the point; Adm1_name is the name of the administrative unit, i.e. province, and country the country name. The last five columns refer to the GFC data and it is recommended to fill them also when using custom land cover and land use maps because this will provide additional information in the analysis.

6.4 Assign values to the sample points

The user interface[5] can be used to assign the necessary values to the CSV file containing the sample point coordinates. But if the points are created using *QGIS* or any other software, it is still useful to add this information. Even if the map data is not the GFC data, it is still useful to add the information of the GFC data (percent tree cover, loss gain, loss year, mask) to the points for further analysis.

An alternative to using the *R* script to populate the CSV file with the required information of the GFC data, topography from SRTM digital elevation data (DEM) and country boundaries, is Google Earth Engine (GEE) application program interface (API) run through Google Earth Engine Playground[10]. GEE is a cloud-based platform for environmental data analysis, and the API allows the user to access the functionality of GEE in both JavaScript and Python. The GEE API can be used if the sample points are input into a fusion table and the fusion table link is specified in the script. The output from the GEE

---

[10] https://ee-api.appspot.com/236fc0f6ec08e4e366def230d8de38f3

script is formatted to be used as the plot data in Collect Earth. A Google account and trusted tester sign up[11] is required to access the Google Earth Engine Playground. Note: The GEE script assigns the value directly from the Hansen et al. (2013) data, using the value of the pixel. If using the method in this script and if the sample is derived from the aggregated pixel block, then the sample has the potential to be labeled incorrectly and then the analysis can possibly be incorrect; showing error in classification when it is actually a labeling error. To avoid this error the aggregated data should be used to assign the label to the sample.

---

[11] https://signup.earthengine.google.com/

7 | Using *Collect Earth* for reference data collection

Collect Earth is a cost effective tool for collecting sample-based reference data. This section explains how to utilize *Collect Earth* to implement the reference design methodology. While this section only covers using *Collect Earth* to collect reference data, however other sources of data can be used as reference data as long as the sampling design and reference design protocols are followed. Other sources of data can include satellite imagery, aerial photography, field work, georeferenced photographs, national forest inventory data, etc.

7.1 *Collect Earth* set up

Collect Earth is a Google Earth plugin for visual land assessment through freely available satellite imagery that was developed by the Food and Agriculture Organization of the United Nations (FAO) under the Open Foris Initiative[12]. *Collect Earth* can be conveniently used for the collection of reference data. The *Collect Earth* User Manual[13] provides a thorough step by step guide to installation, data import/export, and the website[14] provides additional resources and a support forum[15].

First, *Collect Earth* can be installed following the documentation of installation and set up in section 2 of the *Collect Earth* User Manual. Afterwards, the specifications given in this section provide guidance for using *Collect Earth* for reference data collection that can be used in addition to the *Collect Earth* User Manual. This section covers the specification of the spatial assessment unit in the plot layout, the labeling protocol and reference classification uncertainty.

---

[12] http://www.openforis.org/home.html

[13] http://openforis.org/fileadmin/user_upload/Collect_Earth_Tutorials/Collect_Earth_User_Manual_20150618_highres_full.pdf

[14] http://www.openforis.org/tools/collect-earth.html

[15] http://www.openforis.org/support

7.1.1 Project

Collect Earth comes with several custom projects that have different dialogue boxes and survey definitions. The project Global Forest Change Accuracy Assessment[16] has been tailored towards the requirements of an accuracy assessment of a land cover/use change map with four map classes: stable forest, stable nonforest, forest loss and forest gain (see Figure 6). Other classification schemes can be implemented by building a new custom project using the Collect Survey designer. By downloading the latest version of Collect[17], a new survey can be created through the Collect Designer using the Collect Earth template which includes the necessary basic fields that every Collect Earth Project.

ID: 286 - Elevation: 0m, Aspect: 0°, Slope: 0°

Year of change N/A

Canopy Cover (in %) 81-90 %

Figure 8: The *Collect Earth* dialogue box for the GFC project enables the user to define the classification of the reference data.

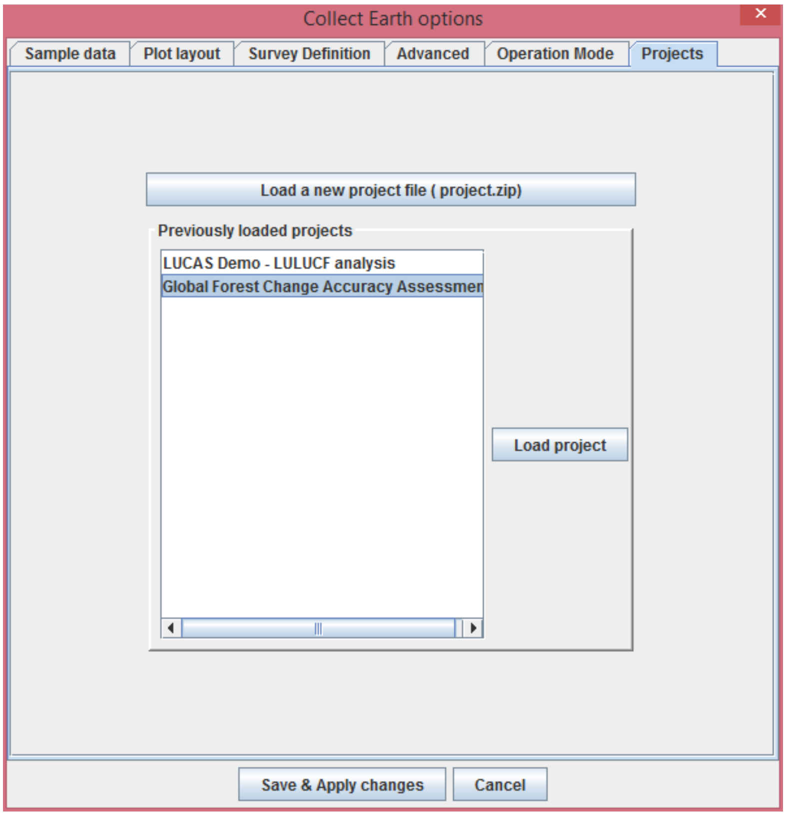
---

[16] http://www.openforis.org/fileadmin/installer/GlobalForestChangeAccuracy.cep

[17] http://www.openforis.org/tools/collect.html

To choose the project, follow these steps:

1. After the installation and set-up of Collect Earth as explained in the user manual, click on the Tools tab and select Properties.
2. Under the Project tab, click Load a new project file (project.zip). Locate the GlobalForestChangeAccuracy.cep file and select the path as the project file (see Figure 9).
3. Click Load project to save the changes. The custom GFC interface for the dialogue box (see Figure 8) will pop up in Google Earth when the area inside a sample plot is clicked.



LUCAS Demo - LULUCF analysis
Global Forest Change Accuracy Assessmen

Figure 9: Specify the GFC project to load the proper interface for the dialogue box in *Collect Earth*.

7.1.2 Loading the sample point locations

The sample point locations are imported as a CSV file that contains the following columns separated by comma (see section 9.3): ID, YCOORD, XCOORD, ELEVATION, SLOPE, ASPECT, ADM1_NAME, COUNTRY, GFC_TREE_COVER_2000, GFC_FOREST_GAIN, GFC_FOREST_LOSS, GFC_FOREST_LOSS_YEAR, GFC_DATA_MASK. It is important the CSV file contains all of these columns with some value for each sample point. It is necessary for the ID, YCOORD and XCOORD to be filled in with the appropriate values. For the remainder of the columns it is optional to fill these in with the accurate information however there must be some value in these columns, i.e., 0, for CE to read the CSV file. Additionally the ID points must be a unique number that has not been used in another CE project. The server for CE stores the information according to the ID number, and if a number is repeated for different datasets, then the classification from the previous dataset will appear for the matching ID in the new dataset.

Follow instructions in section 3.1.1 *Adding Collect Earth data files* in the Collect Earth User Manual. Select the CSV file with the details of the sample plots (the name of this file was specified in *R* in section 9.3) as the path to ced/csv file with plot data.

7.1.3 Modifying the plot layout

When the sample points are imported into *Collect Earth*, they are visualized as plots with a systematic grid of points inside a square. The dimensions of the plot are specified by the user. The plot layout defines the size and shape of the plot as well as the number of points shown in the plot. The size and shape should be in accordance with the spatial assessment unit that had been defined in the sampling design. The plot layout can be adjusted by following the instructions in section 3.1.3 *Modifying the plot layout* in the *Collect Earth* User Manual. It is recommended that the control points[18] in the plot are at least a 4 x 4 plot, in order to have an adequate systematic grid of points to assign the classification to the reference data. The concept of assigning the classification to reference data is further explained in 10.2.1.

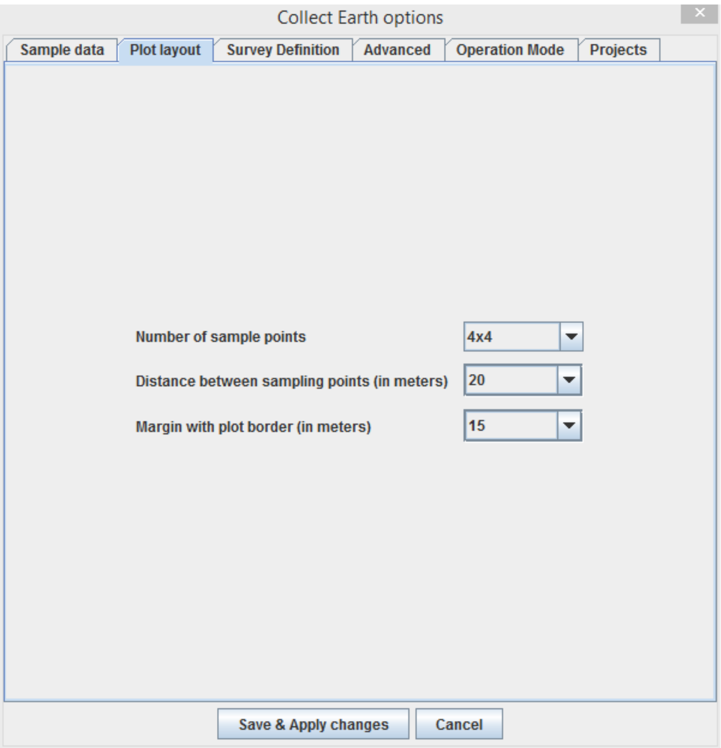
---

[18] In *Collect Earth* the systematic grid of points inside the sample plot are referred to as 'sample points.' In the context of this document, we refer to the points in this grid as control points because we refer to sample points as the collection of the sampling unit for the area of interest.

7.2 Labeling protocol

The labeling protocol is the process of setting up clear rules for defining the classification of the reference data. Generally the labels for the reference data should match the classification of the map data to keep the labeling protocol clear and the definition of agreement between the map and reference data simple.

In *Collect Earth* the labeling protocol is defined by the classification options. For the "Global Forest Change Accuracy Assessment" project, these are the land use dynamics (e.g. F to F, NF to NF, F to NF, and NF to F), confidence (e.g. Y or N), year of change (e.g. 1981 – 2014), canopy cover (e.g. 0 – 100 %), and comments (see Figure 6). If other classification options are required, a custom project needs to be created (see section 10.1.1). Recommendations for the use of these classification options are given below.



Number of sample points: 4x4
Distance between sampling points (in meters): 20
Margin with plot border (in meters): 15

Figure 10: An example of the plot layout for 4 x 4 sample plots for a 90 m x 90 m sample unit

7.2.1 Land use dynamics

Land use dynamics in *Collect Earth* match the map classes: stable forest (F to F), stable nonforest (NF to NF), forest loss (F to NF) and forest gain (NF to F). They are observed by the interpreter using aerial photography and satellite imagery as the reference sources through *Collect Earth*.

Very high resolution imagery in Google Earth or Bing maps should be used first to identify the most recent land cover/use. The plot layout with the dots inside helps in distinguishing between forest and nonforest. The interpreter can count the number of dots that intersect with trees, and then calculate the percentage of dots on tree cover. For example, if 4 out of 16 dots intersect with trees, tree cover would be 25%. Depending on the forest definition, this can then be classified as forest or non-forest.

If at all available, very high resolution imagery often only exists for one point in time. In this case high or medium resolution imagery can be used to identify the dynamics of the land use. Landsat 7 Greenest Pixel TOA reflectance composite, accessed in Google Earth Engine, provides a consistent time series needed to identify the land use dynamics of the sample plot. The visualization of bands 4, 5, 3 effectively shows forest cover in reddish brown (see Section 3.4 *Imagery with Google Earth Engine* in the *Collect Earth* User Manual).

The time series of Landsat 7 also facilitates the identification of the year of change. The year of change only needs to be specified when a change in land cover/use occurs. It is independent from the time span of the map data. The time span of the GFC data is from 2000 to 2013, but if for example forest loss is observed after this period, e.g. in 2014, this could be recorded in the reference data. When comparing the map and reference data later, the year of change will be considered when labeling the reference data. When specifying the change years as 2000 to 2013, this plot would be labeled as stable forest because it was stable forest from the period between 2000 and 2013 and the change occurred after 2013.

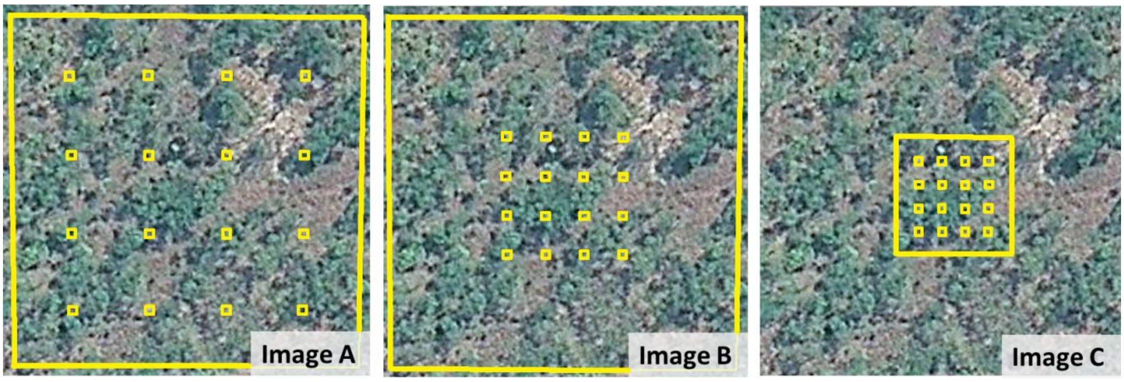


Figure 11: Examples illustrating different configurations of the *Collect Earth* plot layout for different spatial units. Image A: The spatial unit is a pixel block of 3 x 3. The plot layout is a bounding box of 90 m x 90 m with a margin of 15 meters between the plot border and the points and a 20 meter distance between the sampling points. Image B: The spatial unit is a pixel. The plot layout is a bounding box of 90 m x 90 m with a margin of 30 meters between the plot border and the points and a 10 meter distance between the sampling points. Image C: The spatial unit is a pixel. The plot layout is a bounding box of 30 m x 30 m with a margin of 6meters between the plot border and the points and a 6 meter distance between the sampling points. The user is encouraged to experiment with the plot layout to determine which plot layout can include the spatial unit of the reference data and compliment the project specifications. In this example images B and C have the same spatial unit, however the user might find the visual assessment is easier in image B, where the bounding box is larger.

7.2.2 Confidence

The confidence option allows the interpreter of the land cover for the reference data to indicate their level of certainty with their classification. This refers to the interpreter's confidence of the land cover dynamics classification of the sample plot. An example of a sample with no confidence is a plot with consistent cloud cover in all available satellite imagery. If a large portion of the sample cannot be determined the rigor and validity of the sample to accurately represent the population must be evaluated.

Plots labeled with no confidence should be validated by at least one additional interpreter. If another interpreter can confidently classify the plot, the confidence rating can be changed to "yes".

The confidence rating is used in the analysis to subset the reference data by confidence rating. Ultimately if no clear consensus can be determined for a sample plot, it should be excluded from the area calculation and error matrix. The excluded plots classified with no confidence should be reported as plots with an inclusion probability of zero.

7.2.3 Year of change

Year of change in the reference data corresponds with the lossyear layer in the GFC data. Only select a year of change if the land use dynamics are a change category, F to NF or NF to F. The year of change indicates the year the disturbance or tree growth occurred. Because it is more difficult to pin point a specific year for tree growth, year of change usually refers to the loss year. Landsat 7 imagery in Google Earth Engine is a useful resource to determine the year of change because yearly composites are available from 1999 to 2014. It is recommended to also record years of change that are before or after the period of interest as they can be excluded later in the analysis (see section 6).

7.2.4 Canopy cover

Canopy cover in the reference data provides a quantitative description of canopy cover and corresponds with the treecover2000 layer in the GFC data. The canopy cover from *Collect Earth* is not directly used in the accuracy assessment, but is useful information for further analysis.

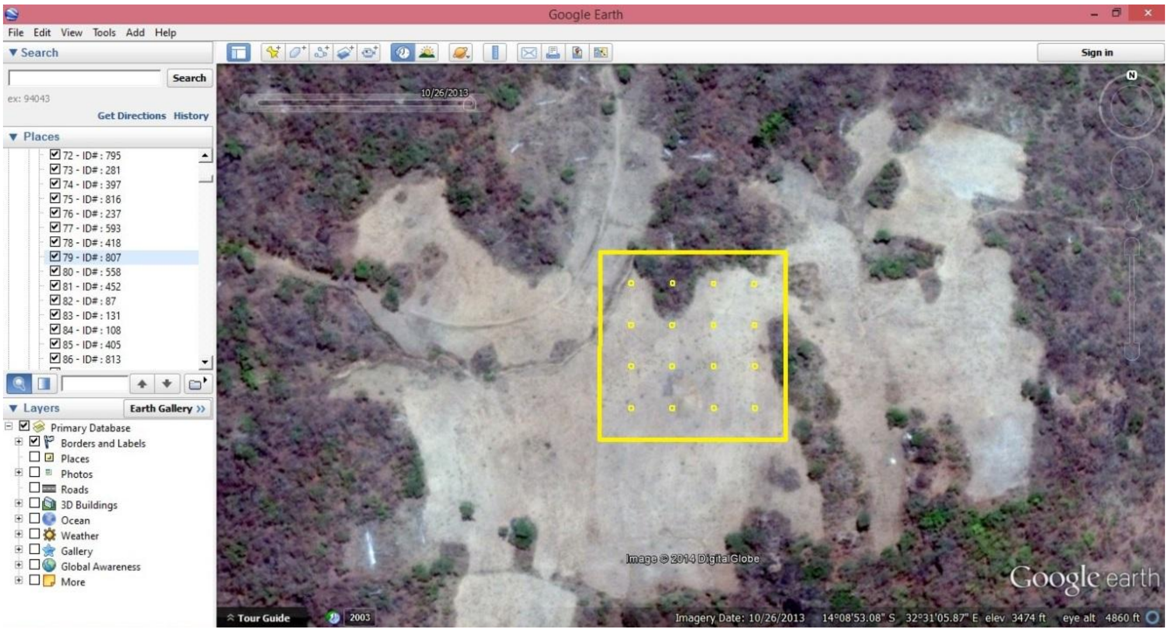
It is not likely that high resolution imagery will be available for 2000, therefore it is acceptable to use the high resolution imagery from the most recent year available. However, if reference data is collected from a year different from the year of the map, then the adjusted area will represent the area of the time of the reference data. Therefore if using more recent imagery for the reference data to access the land cover map 2000, then the adjusted area estimate will not be for the year 2000, but for the years used in the reference data collection.

The sample plot in *Collect Earth* contains a system grid of control points that facilitates the classification of the reference data. The dot grid area estimation method is used to approximate canopy cover information by counting the number of control points in the sample plot, which intersect a tree canopy and divide by the total number of control points in the plot (Barrett and Philbrook, 1970) (example in Figure 10). The ratio of points intersecting with a tree to total points can be used to approximate the percent canopy cover for the sample plot (some reference here to dot-grid method). If no high resolution imagery is available leave this option blank.

The canopy cover should especially be specified for the F to F and F to NF land use dynamics, but also for NF to NF if the treecover threshold is higher than 10 %. If non-forest is defined by a treecover below 10 %, the canopy cover does not necessarily need to be specified in Collect Earth.

7.2.5 Comments

Comments provide a space for the interpreter to voice additional concerns or observations about the sample plot. If a plot is labeled with no confidence the user should include comments explaining why there is no confidence. The comment section can serve as a means of communication between multiple interpreters. The comments are included in the data export and when imported by another interpreter, the previous interpreter's comments are loaded and can be responded to, forming a dialogue between interpreters.



▼ Places
☑ 72 - ID# : 795
☑ 73 - ID# : 281
☑ 74 - ID# : 397
☑ 75 - ID# : 816
☑ 76 - ID# : 237
☑ 77 - ID# : 593
☑ 78 - ID# : 418
☑ 79 - ID# : 807
☑ 80 - ID# : 558
☑ 81 - ID# : 452
☑ 82 - ID# : 87
☑ 83 - ID# : 131
☑ 84 - ID# : 108
☑ 85 - ID# : 405
☑ 86 - ID# : 813

▼ Layers Earth Gallery »
☑ Primary Database
  ☑ Borders and Labels
  ☐ Places
  ☐ Photos
  ☐ Roads
  ☐ 3D Buildings
  ☐ Ocean
  ☐ Weather
  ☐ Gallery
  ☐ Global Awareness
  ☐ More

Imagery Date: 10/26/2013 14°08'53.08" S 32°31'05.87" E elev 3474 ft eye alt 4860 ft

Figure 12: To calculate the tree cover of the sample plot, count the number of control points which directly intersect tree canopy cover. In this example from 2013 in Zambia there is one point that intersects tree canopy cover, equivalent to 6% canopy cover for the plot. This constitutes non-forest according to the national forest definition. The sample plot needs to be compared with other points in time to determine if this plot shows forest loss (forest to non-forest) or stable non-forest (non-forest to non-forest).

7.3 Defining agreement

After the map and reference classifications have been obtained, the rules for defining agreement need to be established. Defining agreement is relatively simple due to the simple classification scheme. A single label class in the reference, i.e. F to F, is compared to the map classification, stable forest. If these labels agree, the map class is correct for that unit. If the labels disagree, the type of misclassification is identified. Defining agreement leads to constructing an error matrix that presents the proportions of agreement and disagreement between the map and reference classifications (see section 6).

7.4 Reference classification uncertainty

The reference classification is not a perfect representation of the condition on the ground and is subject to uncertainty. An assessment of the reference classification uncertainty should be conducted because small errors in the reference data set can lead to large bias of the estimators of both classification accuracy and class area. Interpreter uncertainty can be feasibly assessed in two ways: 1) quantifying the no confidence plots; 2) assessing the variability between the reference classification assigned to the same plot by different interpreters.

Plots marked with no confidence by one interpreter should be assessed by other interpreters. The Saiku Server, which comes with *Collect Earth*, can be used to identify the no confidence plots. Follow the steps in Section 4 *Analyzing data with Saiku Server* of the *Collect Earth* User Manual.

Continuous communication among the multiple interpreters to discuss and document difficult cases is important to foster enhanced consistency and accuracy of the reference labeling process. Disagreements among interpreters evaluating the same plot can be resolved by consensus agreement on the reference class. To maximize agreement between interpreters, the validation should commence with all interpreters validating the same plots and engage in discussion about the labeling protocol. An interpretation key for a region can be developed as a useful tool for maintaining consistency. Such a key can show classes that could be easily mistaken for forest such as a coconut plantation (if it is defined as agriculture) or sparse shrubland.

8 | Deriving accuracy and area estimates in *R*

Accuracy levels and uncertainty of the area estimates are calculated by comparing the map and reference classifications in an error matrix. The accuracy and area estimates and their confidence intervals can be derived using *R*. An interactive application using R shiny has been developed which implements the formula for stratified random sampling with the map classes being the strata.

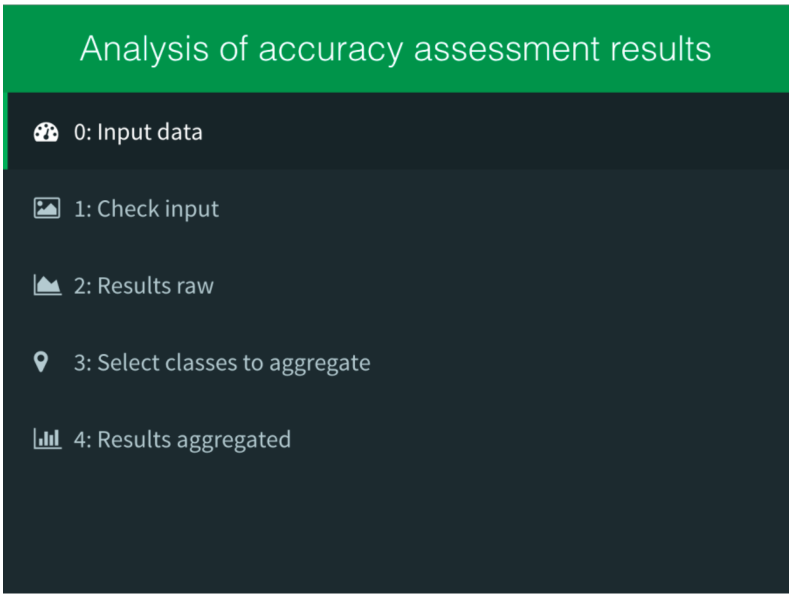


Figure 13: User interface[5] for calculating accuracy measures and adjusted areas

After running the application, the output file can exported in CSV format and opened in LibreOffice, Calc or Excel. It includes all relevant measures to report for a robust accuracy assessment. The error matrix, graphical display of the adjusted areas and the accuracy measures can be readily viewed and exported.

9 References

Barrett, J. P., & Philbrook, J. S. 1970. Dot grid area estimates: precision by repeated trials. *Journal of Forestry* 68, 149-51.

Bey, A., Sanchez-Paus Diaz, A., Pekkarinen, A., Patriarca, C., Maniatis, D., Weill., Mollicone, D., Marchi, G., Niskala, J., Rezende, M., Ricci, S. 2015. Collect Earth User Manual: A guide to monitoring land use change and deforestation with free and open-source software. Food and Agriculture Organization of the United Nations, Rome, Italy. Available online here: http://www.openforis.org/tools/collect-earth.html

Cochran, W. G., 1977. *Sampling techniques*. John Wiley & Sons.

Congalton, R. G. & Green, K. 2008. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices, Second Edition.* CRC Press

GFOI. 2013. Integrating remote-sensing and ground-based observations for estimation of emissions and removals of greenhouse gases in forests: Methods and guidance from the global forest observations initiative. Pub: Group on Earth Observations, Geneva, Switzerland, 2014.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina,A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L. C., Justice, O., Townshend, J. R. G. 2013. High-resolution global maps of 21st-century forest cover change. *Science*, 342(6160):850–853.IPCC, 2003. Good Practice Guidance for Land Use, Land Use Change and Forestry. Institute for Global Environmental Strategies, Japan.

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., Wulder, M. A. 2014. Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, 148:42–57.

Pontius Jr, R. G. & Millones, M. 2011. Death to kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing*, 32(15):4407–4429.

Potapov, P. V., Turubanova, S. A., Tyukavina, A., Krylov, A. M., McCarty, J. L., Radeloff, V. C., Hansen. M. C. 2014b. Eastern Europe's forest cover dynamics from 1985 to 2012 quantified from the full Landsat archive. *Remote Sensing of Environment,* (0) doi: http://dx.doi.org/10.1016/j.rse.2014.11.027.

Potapov, P. V., Dempewolf, J., Talero, Y., Hansen, M. C., Stehman, S. V., Vargas, C., Rojas, E. J., Castillo, D., Mendoza, E., Calderón, A., Guidice, R., Malago, N., Zutta, B. R. 2014a. National satellite-based humid tropical forest change assessment in Peru in support of REDD+ implementation. *Environmental Research Letters,* 9.

Stevens, D. L. and Olsen, A. R. 2004. Spatially balanced sampling of natural resources. Journal of the American Statistical Association, 99: 262-278.

Strahler, A. H., Boschetti, L., Foody, G. M., Friedl, M. A., Hansen, M. C., Herold, M., Mayaux, P., Morisette, J. T., Stehman, S. V., Woodcock, C. E. 2006. Global land cover validation: Recommendations for evaluation and accuracy assessment of global land cover maps. *European Communities, Luxembourg,* 51.

UNFCCC, 2013. Decision 14, CP 19. Warsaw, Poland

Wulder, M. A., White, J. C., Magnussen, S., McDonald, S. 2007. Validation of a large area land cover product using purpose-acquired airborne video. *Remote Sensing of Environment,* 106(4):480–491.

A | Appendix

A.1 Software instructions

This document relies solely on open source software (OSS), which is freely available. This section provides instructions on the installation and getting started as well as with additional resources for finding help. These OSS have big user communities which document exchanges on technical issues and difficulties online, making it fairly easy to find help on particular problems. Previous experience with these programs is recommended but not necessary.

Note: Large aoi takes significant processing power and computer memory. Be aware of memory usage and large files which have the potential to crash the system. You might consider splitting up the aoi into smaller parts, using another software or being very patient with your computer.

A.1.1 R

R is a software programming language and software environment for statistical computing and graphics. It is widely used for statistical and data analysis, including spatial data. Its popularity has increased drastically in recent years, and *R* is among the most used software for statistical analysis. It runs on Windows, UNIX platforms and MacOS.

Installation

R itself is only a programming language that can be used directly from a command line, or from an editor that provides a user interface.

To download *R*, refer to the website of the *R* project[19] and choose a CRAN mirror close to your country (this determines where the software will be downloaded from) and the version according to your operating system. After successfully installing *R*, you can install *R*Studio as an editor. Follow the instructions on the *R*Studio website[20].

When installing *R*, basic packages are installed, and by installing additional packages there are many more functions available for specific purposes. Additional packages need to be loaded into *R* before making use of their functionality. The *R* scripts provided in this document automatically take of additional packages.

Getting started

To get started with *R* it is recommended to make use of one of the many well-made tutorial materials online. Free online tutorials such as Datacamp introduction to *R*[21] provide interactive explanations of *R* basics. To be able to use the code provided by this document, you should know how variables are defined, how to specify file paths, why you sometimes need to install additional packages, and how data is loaded into *R*. Additional instructions are provided in the scripts and the remainder of the commands are automated.

A variable is a basic concept in programming. It allows you to store a value (like a number) or an object (such as a function description or a table) in *R*. Through the name of the variable you can later easily access the value or object stored in it.

```
a_number <- 4 # a_number is the name of the variable, <- tells R
    to assign something to it, and 4 is the value that will be
    stored in the variable
a_number # prints what is stored in the variable
my_name <- "Sara" # If you want to store a character string in a variable,
    make sure to quote it. Unquoted sequences will be treated as
    variables.
```

Before starting the script, it is recommended to clear the script to avoid confusion with variable names that were in use before:

```
rm(list = ls())
```

Of course you do not always have to enter the data manually in *R*, you can also read in existing files. For that purpose, you should first

---

[19] http://www.r-project.org/

[20] http://www.rstudio.com/products/rstudio/download/

[21] https://www.datacamp.com/

define your working directory where your data is stored now and where the data you produce by the *R* analysis will be saved. You can set the working directory with the command setwd():

```
setwd ( " path  to  your  working  directory" )
```

Always make sure to quote the path to the working directory because it is a character string. Remember to use the proper *R* syntax by using either forward slashes “/” or double backward slashes “\\” for folder or file paths.

In Windows, folders need to be separated by “//”, for example C: //Users//Myname“. Then files saved in the directory can be assessed, for example read in a csv file:

```
df <- read.csv ( " name_ of _my_ csv _ file.csv" ) head
( df )
```

df is the variable that will be created, read.csv is the command used to read in csv files, and as all commands it is not quoted. The file name needs to be quoted because otherwise *R* looks for a defined variable and will throw out an error message. head(df) prints out the first six lines of your file. *R* does not read in Excel files, but in Excel you can save your data as csv-file and then import it into *R* as shown above. In addition to reading in text or csv files, *R* can also read in many other types of data, such as raster images (with the function raster()), shapefiles (with the function readOGR()) and many more.

If you are not sure how to use a specific command, instructions for that command will appear by typing a question mark followed by the name of the command.

```
? raster
```

The example script above provides information about the raster function.

To use the scripts provided with this document, open them in *R*Studio. First, the user specifications in the beginning of the script need to be changed, such as defining the working directory, etc.

Resources for learning *R* can be found online, such as on the RStudio website[22], which gives a good overview of online material. Additional links include:

- An introduction to R by Computerworld[23]
- Try R[24]
- The datacamp introduction to R[25]

These links explain the structure of RStudio, how to load data, set the working directory and other tips for starting to use *R*.

A helpful resource for using spatial data in *R* is "spatial.ly[26]".

Finding help

Additional resources for technical issues include:

- The stackoverflow[27] site: If you have a question, it has most likely already been asked and answered here
- additional resources on R[28] on the official R website
- additional documentation of R raster package[29]
- R raster package manual[30] for working with raster data in R
- R GFCanalysis package manual[31]
- the website of the R GFCanalysis package[32]

---

[22] http://www.rstudio.com/resources/training/online-learning/#R

[23] http://www.computerworld.com/article/2497143/business-intelligence/beginner-s-guide-to-r-introduction.html

[24] http://tryr.codeschool.com/

[25] https://www.datacamp.com/

[26] http://spatial.ly/r/

[27] http://stackoverflow.com/questions/tagged/r

[28] http://cran.r-project.org/manuals.html

[29] http://www.inside-r.org/packages/cran/raster/docs

[30] http://cran.r-project.org/web/packages/raster/raster.pdf

[31] http://cran.r-project.org/web/packages/gfcanalysis/gfcanalysis.pdf

[32] http://azvoleff.com/articles/analyzing-forest-change-with-gfcanalysis/

A.1.2 QGIS

QGIS is an open source desktop geographic information systems (GIS) application that provides data viewing, editing, and analysis capabilities. It runs on Windows, Linux, MacOS, BSD and Android and has an international support community of enthusiastic users, developers and supporters. The QGIS[33] website provides comprehensive information about the installation, getting started and where to get help, and much more. Here we provide you with links to the most important information.

Installation

To install QGIS Desktop, choose your operating system on the QGIS download website[34] and follow the instructions. Additional functionalities are provided in plugins that you can add later. Just use the plugin manager that you can find under Plugins -> Manage and Install Plugins. The available plugins are listed on this website[35].

Getting started

For GIS beginners a good starting points is introduction to GIS[36] provided by QGIS. QGIS Desktop provides similar functionalities as other Desktop GIS applications, such as ArcGIS. It can deal with a variety of data types and formats, such as vector and raster data. The user interface consists of five main parts (see Figure A.12). In the data sources, you can browse through the files available on your computer. The table of contents shows the layers that you have already loaded into QGIS. The map area displays those layers. With the tools, you can zoom, select certain features and many more. The system menu also provides additional functionalities.

QGIS also provides an example data set[37] that you can use to test the functionalities of the software.

---

[33] http://qgis.org/

[34] http://qgis.org/en/site/forusers/download.html

[35] http://plugins.qgis.org/plugins/

[36] http://docs.qgis.org/2.2/en/docs/gentle_gis_introduction/index.html

[37] http://docs.qgis.org/2.2/en/docs/user_manual/introduction/getting_started.html#sample-data

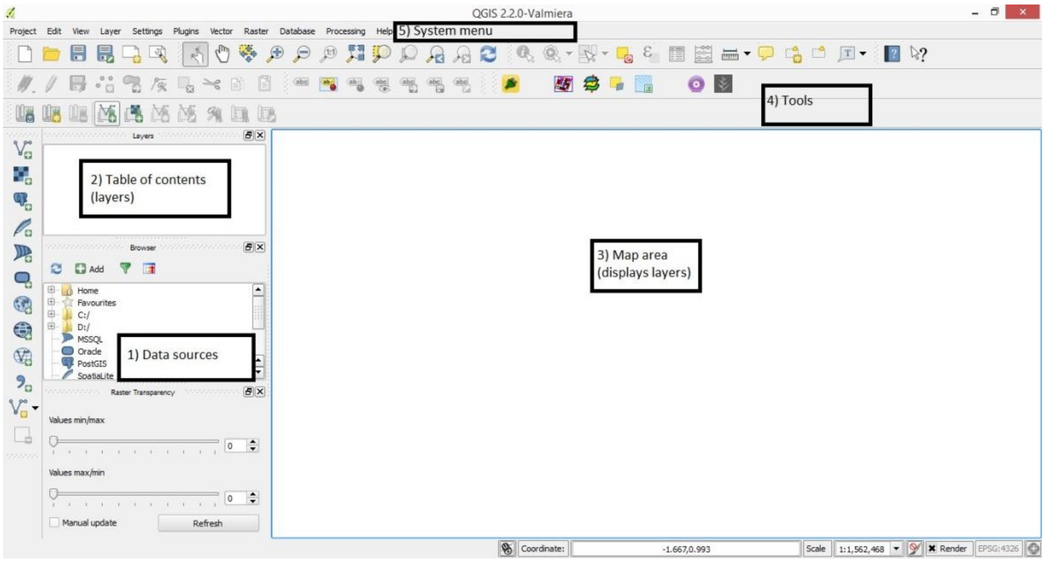


Figure 14: QGIS Desktop interface.

Finding help

QGIS is driven by a big international community of users, developers and supporters, so there are plenty of resources available. As for all other parts, the QGIS website also provides a good overview about where to find help[38].

---

[38] http://qgis.org/en/site/forusers/support.html

A.2 Create random points in QGIS

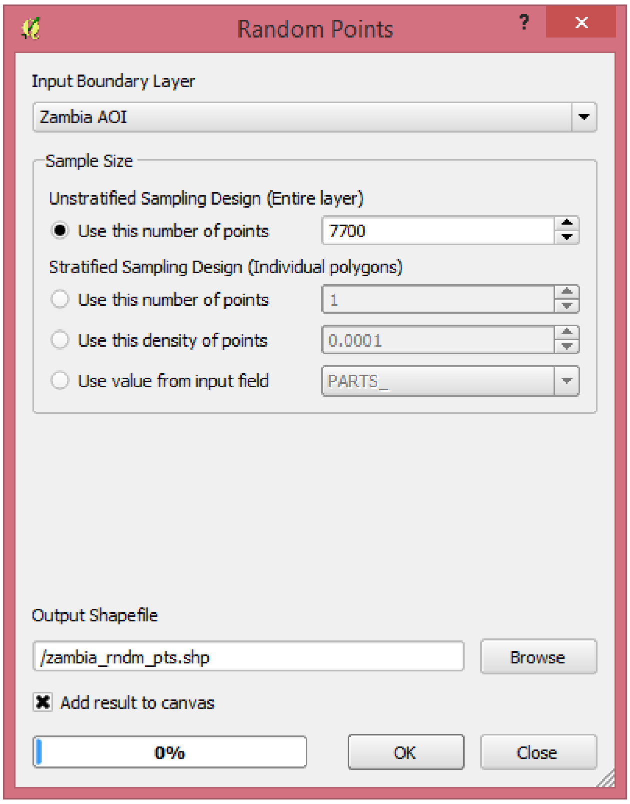
In the document instructions for stratified random sampling are provided using an *R* script. This Appendix section offers an alternative software instructions in *QGIS* for creating stratified random points. Figure 2 shows a few of the steps in the accuracy assessment can be completed using *QGIS*. *QGIS* is provided as an alternative to the *R* script because GIS users are often accustomed to using an interface where it is convenient to visualize intermediate steps and results. Step by step instructions are provided in order to create random points within the stratified map data. The coordinates of the resulting points can then be exported as an CSV file and imported into a Google fusion table to assign the necessary attributes to the points using a Google Earth Engine API script.

In *QGIS*

1. Open *QGIS* and add area of interest (aoi) in *QGIS* using the add vector layer.

2. Vector->Research tools->Random points

3. Select the aoi as the input boundary layer. The sample needs to be larger than the total sample points, it is recommended to use at least ten times the amount of total sample points. Enter this amount under unstratified sampling design (entire layer). Select add result to canvas to view the random points for the entire aoi and click OK.



Input Boundary Layer: Zambia AOI

Unstratified Sampling Design (Entire layer)

● Use this number of points: 7700

○ Use this number of points: 1

○ Use this density of points: 0.0001

○ Use value from input field: PARTS_

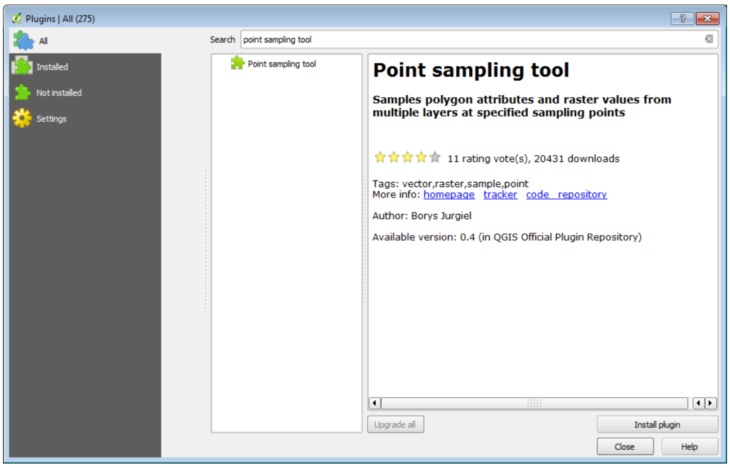
Output Shapefile: /zambia_rndm_pts.shp

☒ Add result to canvas

0%

4. Next in the menu bar go to Plugins -> Manage and Install Plugins…

5. Search 'point sampling tool.'



Search: point sampling tool

Point sampling tool

Samples polygon attributes and raster values from multiple layers at specified sampling points

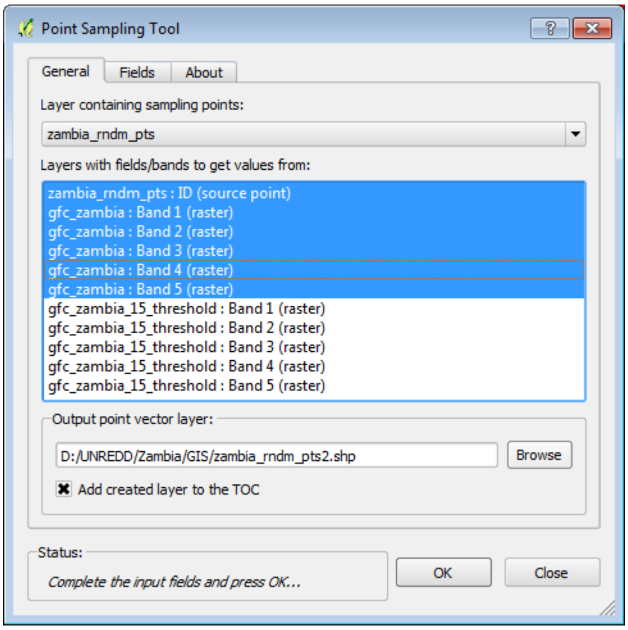
11 rating vote(s), 20431 downloads

Tags: vector,raster,sample,point
More info: homepage tracker code repository

Author: Borys Jurgiel

Available version: 0.4 (in QGIS Official Plugin Repository)

6. Select point sampling tool and select install plugin.
7. Add the layers with the strata. In this example the GFC layer has the information for the strata: forest loss, forest gain, stable forest, and stable non-forest.
8. Plugins -> Analyses -> Point sampling tool
9. Select the output shapefile from the random points as the layer containing sample points.
10. Select the layers containing the desired strata and random point layer as the layers with fields/bands to get values from. Using the GFC output from the *R* script the primary output contains information for the gains, losses, percent tree cover, and data mask.
11. Enter an output point vector layer name and check add created layer to the table of contents (TOC).



Layer containing sampling points: zambia_rndm_pts

Layers with fields/bands to get values from:

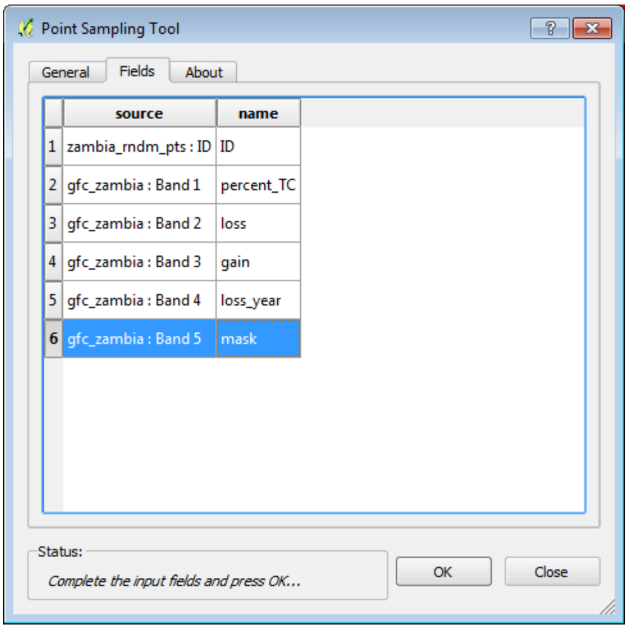
zambia_rndm_pts : ID (source point)
gfc_zambia : Band 1 (raster)
gfc_zambia : Band 2 (raster)
gfc_zambia : Band 3 (raster)
gfc_zambia : Band 4 (raster)
gfc_zambia : Band 5 (raster)
gfc_zambia_15_threshold : Band 1 (raster)
gfc_zambia_15_threshold : Band 2 (raster)
gfc_zambia_15_threshold : Band 3 (raster)
gfc_zambia_15_threshold : Band 4 (raster)
gfc_zambia_15_threshold : Band 5 (raster)

Output point vector layer: D:/UNREDD/Zambia/GIS/zambia_rndm_pts2.shp

☒ Add created layer to the TOC

Status: *Complete the input fields and press OK...*

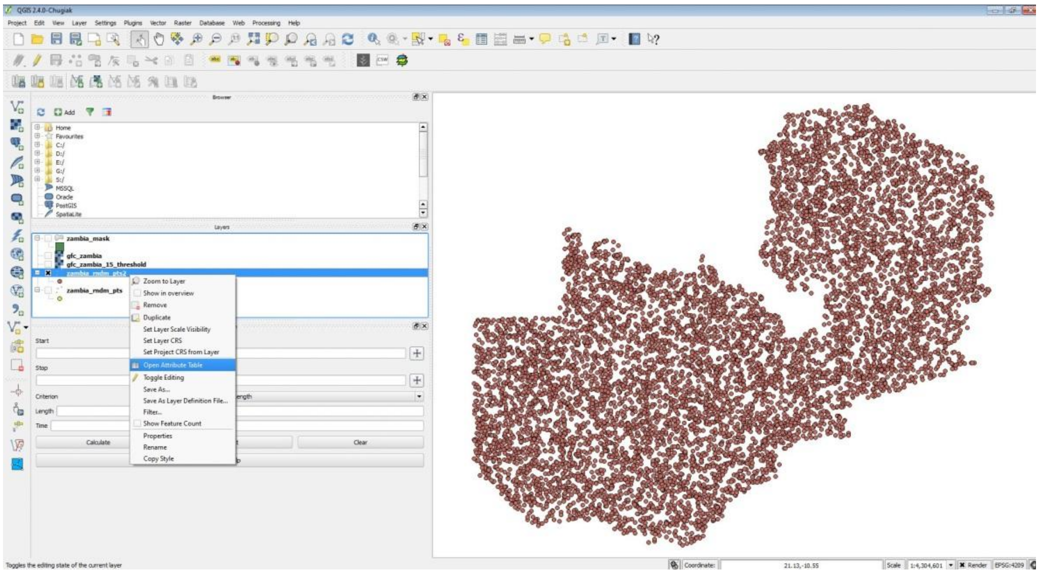
12. Go to the ’fields’ tab and enter the field name for each data layer, then press OK.



	source	name
1	zambia_rndm_pts : ID	ID
2	gfc_zambia : Band 1	percent_TC
3	gfc_zambia : Band 2	loss
4	gfc_zambia : Band 3	gain
5	gfc_zambia : Band 4	loss_year
6	gfc_zambia : Band 5	mask

Status:
Complete the input fields and press OK...

13. The new output point file will look exactly like the randomly sampled points however now the attribute table contains forest information for each point. Open the attribute table for the newly created point file by right clicking the layer.



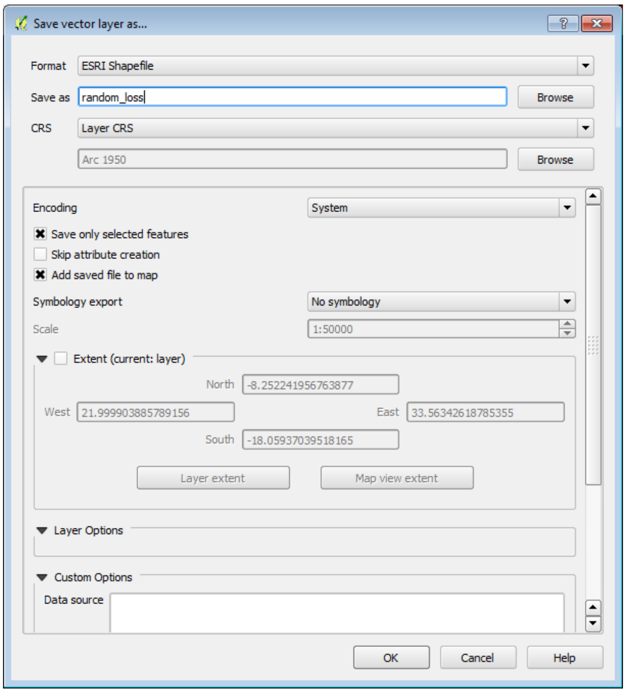
14. Click on the select by expression symbol in the attribute table.

15. Now the strata will be separated into individual point files. The select by expression is used to distinguish between the strata. An expression which shows desired categories is entered, for example to the expression below selects forest loss within the data mask which is not forest gain this expression. Forest loss sample points: "loss" = 1 AND "mask" = 1 AND "gain" = 0

16. Click select and ensure that the amount of points selected is at least equal to the desired sample size. If the selection is smaller go back to step 4 and create a sample that is larger or see step âA˘ e¸ for an alternative.

17. Next the layer is saved as ... by right clicking on the point sample layer and selecting 'Save as ...'

18. The format of the layer is ESRI Shapefile and create a unique name for save as output layer name. The coordinate reference system (CRS) can be the layer CRS and be sure to check 'save only selected features' and 'add saved file to map' under encoding. Then press OK.



Save vector layer as...

Format: ESRI Shapefile

Save as: random_loss

CRS: Layer CRS

Arc 1950

Encoding: System

☒ Save only selected features

☐ Skip attribute creation

☒ Add saved file to map

Symbology export: No symbology

Scale: 1:50000

☐ Extent (current: layer)

North: -8.252241956763877

West: 21.999903885789156

East: 33.56342618785355

South: -18.05937039518165

19. Repeat steps 14 – 19 to create unique sample point layers for each of the strata. In this example to achieve strata for forest loss, stable forest, stable non-forest the following expressions for step 16 used Forest loss sample points: "loss" = 1 AND "mask" = 1 AND "gain" = 0 AND "f_nf_15" = 1
    Forest gain sample points: "gain" = 1 AND "mask" = 1 AND "loss" = 0 AND "f_nf_15" = 0
    Stable forest sample points: "f_nf_15" = 1 AND "loss" = 0 AND "gain" = 0 AND "mask" = 1
    Stable non-forest sample points: "f_nf_15" = 0 AND "loss" = 0 AND "gain" = 0 AND "mask" = 1
    In the example the forest gain only returned 1 sample point, while the desired number of sample points is 100. This requires more intensive sampling, such as 100 times the amount of sample points. Creating a very large number of data points can cause *QGIS* to crash.

20. An alternate option to creating a random sample for a stratum that is small and is not captured within an unstratified sampling design is to randomly sample only this particular stratum. If the stratum is in vector format we can randomly sample it using the random points module demonstrated in step 4. If the stratum is in raster format each pixel must be converted to a point.

21. To convert the pixels to a point first make sure the coordinate system for the stratum layer is projected by double clicking on the dataset and in the Table Contents selecting the metadata tab and reviewing the metadata information. If it is not a projection CRS, go to Raster -> Projections -> Warp (reproject) and select an appropriate projection. Note the pixel size and unit information of the dataset.

22. Next with the projected stratum we will create a grid of regular points, so each pixel is a point. Select Vector -> Research Tools -> Regular Points. The Input Boundary Layer should be set to the aoi. Again note the pixel size and unit of the projected strata and in the Grid Spacing section set the point spacing value to the pixel size. In order to make the regular points fall within the center of each pixel we will add an offset of half our pixel size so divide the pixel size by two and assign this value to the Initial inset from corner (LH side) value.

23. To format the points created in QGIS to facilitate their use in Collect Earth, follow the instructions in section 9.3.

www.fao.org/forestry

I5601E/1/04.16