CLIMATE CHANGE DEPARTMENT (FORESTRY COMMISSION) NATIONAL REDD+ SECRETARIAT

REPORT ON ADDITIONAL SAMPLE DATA COLLECTION IN RECALCULATION OF FRL

2nd-5th March, 2020

List of Acronyms

Acronym	Meaning
AOI	Area of Interest
CFP	Carbon Fund Participants
ER	Emission Reductions
ERPD	Emission Reductions Payment Document
FAO	Food and Agricultural Organization
FCPF	Forest Carbon Partnership Fund
FRL	Forest Reference Level
GEE	Google Earth Engine
IPCC	Intergovernmental Panel on Climate Change
LU	Land Use
LULUC	Land Use Land Use Change
MRV	Monitoring Reporting and Verification
QA/QC	Quality Assurance/Quality Control
REDD+	Reducing Emissions from Deforestation and Forest Degradation plus Conservation of carbon stocks, Sustainable management of forest and Enhancement of forest carbon stocks.

Table of Contents

1.0 Introduction

FAO has been contracted by WB to support Ghana with the improvement of its activity data, responding to the conditions raised by the CFP and enabling Ghana to assess REDD+ results in 2020. FAO will support the government in developing an improved estimate of activity data for deforestation, forest degradation and enhancement. To achieve this aim, two data collection exercises have taken place, but the first data collected resulted in a low deforestation rate, thus the need for another data collection that was done from 3rd of February to 7th of February 2020 at the World Bank office in Accra, followed with a second in-country mission from 11th to 14th of February.

As part of next steps derived from the last in-country mission and MRV meeting (14/02/2020), there was the need to collect extra data. This was done from the 2nd to 5th of March 2020 at the World Bank Office in Accra.

2. Approach

The team was tasked to interpret **1,204** plots on Collect Earth. The plots were grouped into 5 sample files, each file containing its requirements. 3 of the files were QAQC and 2 Recheck plots. Participants were taken through the specified requirements that came with each file in order to enhance understanding and create a consensus as to what is expected.

2.1 Plot Definition

The plots were shared amongst **4** team members with each collecting **301** points. Classification of the plots was done in Collect Earth mainly with the support of Google earth, Bing maps, planet data and Google earth engine (GEE). Each plot had **49** points and classified under the main land use systems in Ghana which is also in line with the IPCC Classification hierarchy.

The image below shows the response design and labelling for data collection on Collect Earth.
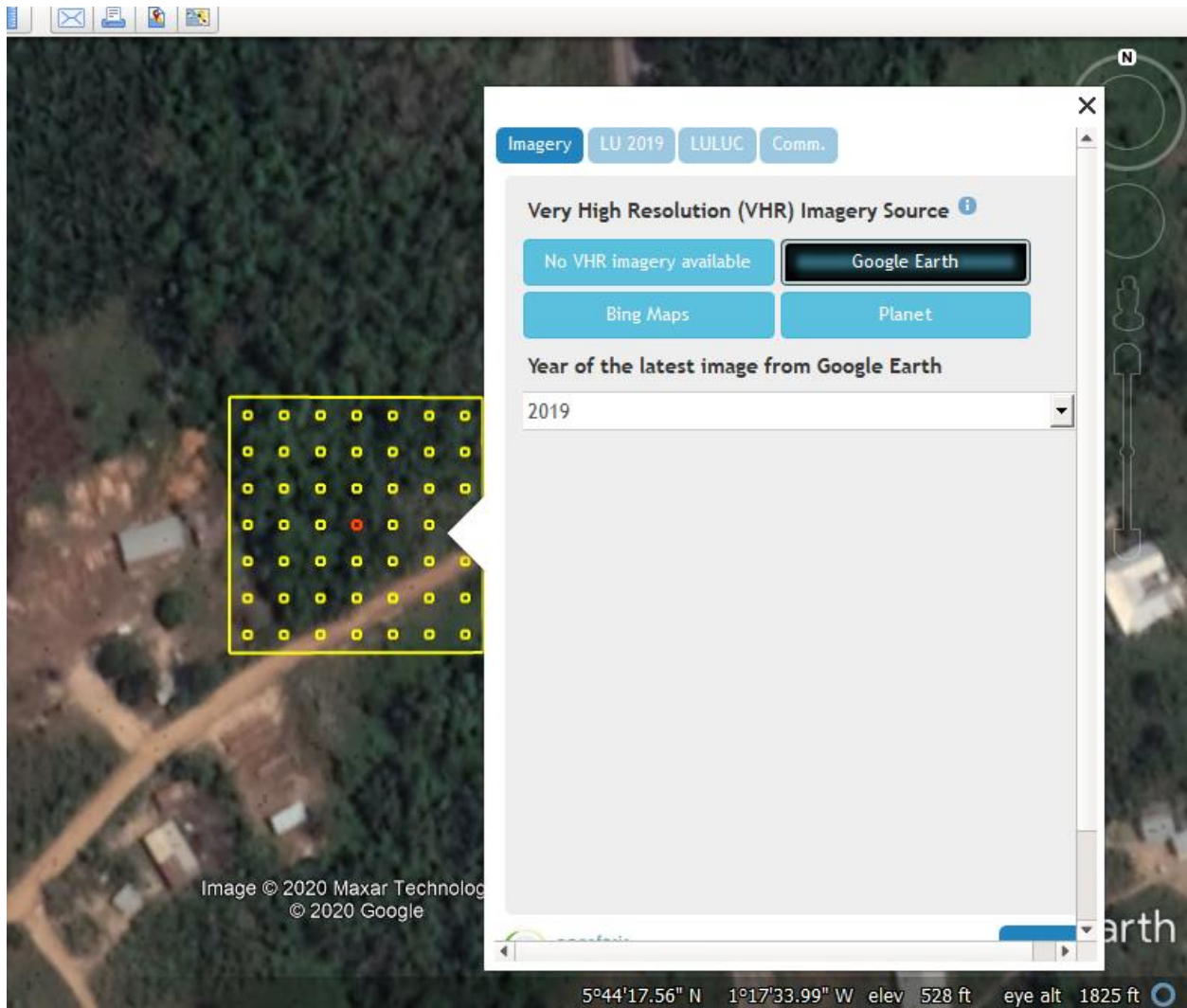
3.0 Quality Assurance/ Quality Control

It is ideal that every data collected is always of exceedingly high quality, devoid of error and biases. This is therefore achieved, through the process of Quality Assurance and Quality (QA/QC) control to trace any error that emerges throughout the data collection process.

However, most of the samples collected during this exercise were QAQC plots, therefore there was no need to do a separate QAQC this time.

4.0 Challenges encountered with Interpretation of plots

- Planet data images was cloudy and blurry when zoomed in thus making interpretation difficult.

- The supportive maps such as Bing, GEE and Planet data were not stable for some interpreters therefore delaying the work.

- One set of QAQC plot had error messages when classifying but this was rectified on the last day of work.
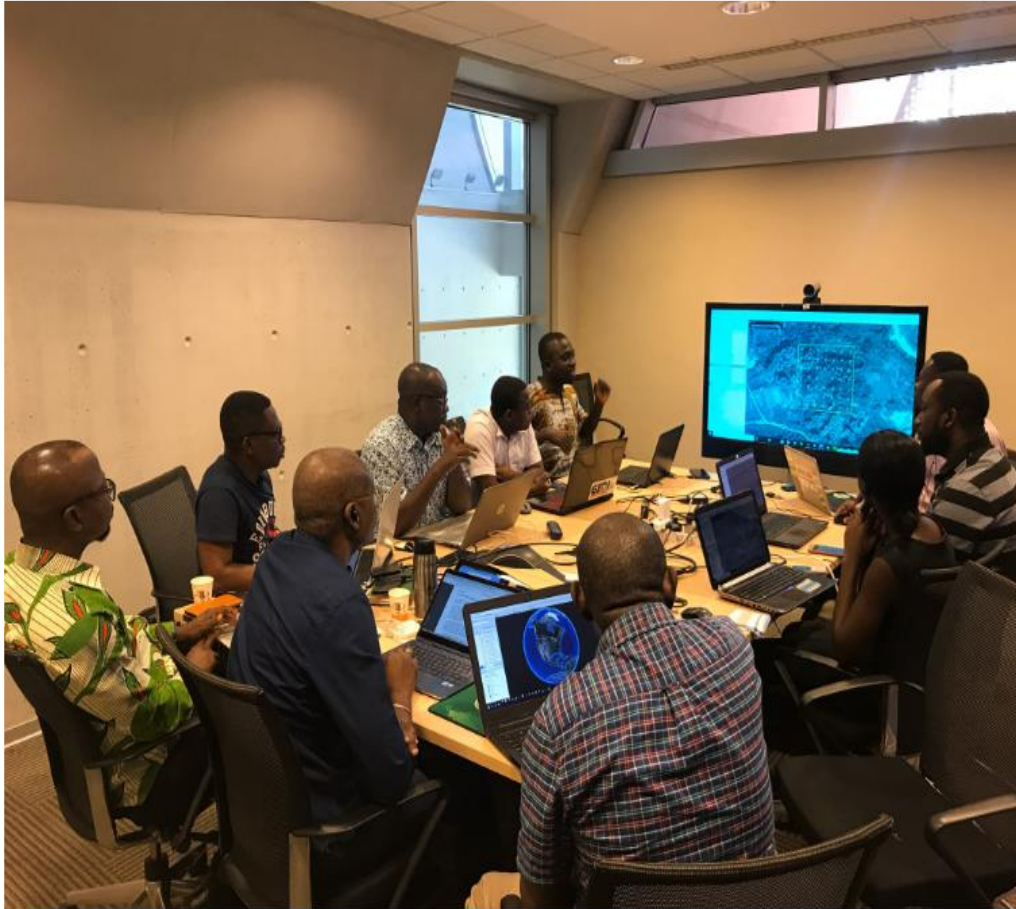
REPORT ON GHANA CAPACITY BUILDING ONLINE TRAINING SESSIONS FOR REDD+ FCPF REPORTING

28th September – 9th October, 2020

List of Acronyms

Acronym	Meaning
AD	Activity Data
CE	Collect Earth
CFP	Carbon Fund Participants
CfRN	Coalition for Rainforest Nations
EF(S)	Emission Factor(S)
ER	Emission Reductions
ERPD	Emission Reductions Payment Document
FAO	Food and Agricultural Organization
FCPF	Forest Carbon Partnership Fund
FRL	Forest Reference Level
GIS	Geographic Information System
HIA	Hotspot Intervention Area
IPCC	Intergovernmental Panel on Climate Change
MRV	Monitoring Reporting and Verification
QA/QC	Quality Assurance/Quality Control
QGIS	Quantum GIS
REDD+	Reducing Emissions from Deforestation and Forest Degradation plus Conservation of carbon stocks, Sustainable management of forest and Enhancement of forest carbon stocks.
RMSC	Resource Management Support Centre
SOPs	Standard Operating Procedures
WB	World Bank

Table of Contents

Introduction

Ghana's Emission Reductions (ER) Programme was included in the Forest Carbon Partnership Fund (FCPF) portfolio in June 2017, on the condition that the deforestation and degradation assessment would be improved. The Emission Reductions Payment Document (ERPD) included deforestation estimates following a stratified area estimate approach, a degradation estimate based on proxy data (e.g. timber extraction statistics, MODIS burned area, etc.) and an enhancement estimate based on national statistics on areas planted. Food and Agriculture Organization (FAO) has been contracted by World Bank (WB) to support Ghana with the improvement of its activity data, responding to the conditions raised by the Carbon Fund Participants (CFP) and enabling Ghana to assess REDD+ results in 2020. FAO has been supporting the government in developing an improved estimate of activity data for deforestation, forest degradation and enhancement. The updated estimates were reported in a Reference Level amendment for the Ghana Cocoa Forest REDD+ Program.

As part of support from FAO, a virtual knowledge transfer and capacity building training session for Ghana was organized by FAO. The training was held via zoom **from 28th September to 9th October 2020** from **1:30pm - 3:30pm** each day. The main objective of the training was to ensure a complete transfer of capacity to the Ghanaian team. The specific objectives are as follows:

- To independently reproduce the methodology that is detailed in the SOPs to produce emission estimates for reporting to Forest Carbon Partnership Facility (FCPF) by Ghana team.
- To review the technical steps necessary to repeat the procedure for generating emission estimates
- To identify any remaining gaps, which might require further technical support sessions.

The participants of the training session were in-country Measuring, reporting and verification (MRV)/Geographic Information System (GIS) experts, who have been involved in the MRV process.

WEEK 1 (28th September – 2nd October)

The first week was focused on the Activity data (AD), how to generate for a selected area, collect and collate for analysis (sampling design, response design and data management).

DAY 1

The training started with introductions from participants. Yelena took the team through the outline for the training, she mentioned that week 1 would be focused on the sampling design, response design and data management. Yelena was the tutor for week 1 and Marieke for week 2. Yelena presented on the sampling design, she mentioned that Ghana adopted the stratified systematic sampling approach which is due to the heterogeneous nature of the landscape. And also for this exercise, Emission Factors (EFs) were quantified by vegetation zones thus another reason for different intensification sampling (4×4, 2×2 and 1×1). She added that with the sampling design, sample selection needs to be done by one person, a dedicated person in charge of distributing samples and collating all for analysis. Because it allows for a much easier data analysis and no data is lost as a result of mishandling from various people.

There was a practical session on how to extract Hotspot Intervention Area (HIA) information to grid for intensification with QGIS and determining number of samples for duplication (QA/QC). After which a practical task was given to participants.

DAY 2

Training begun with Yelena asking participants progress on the previous day's task. The major challenge participants were facing was how to replicate the practical session on ArcGIS. Notwithstanding, the day's training commenced with a presentation on the response design (refers to what rules have been applied when interpreting the sample plot) and how it can be modified to suite project needs. She mentioned that, going forward Ghana does not need to create a new response design; the one already created can be used any time in the future. She also took participants through the Collect platform where the response design is created and modified. In recalculating the FRL, Ghana adopted the use of IPCC hierarchy classification as a benchmark in the interpretation of plots:

- Settlement = 20%

- Cropland   = 20%

- Forest   = 20%

- Grassland = 20%

- Wetland = 20%

- Otherland = 20%

This is to infer that all plots interpreted, had 20 % of land use classes which preceded over the other at any point in time following the order in which the land uses are listed above.

DAY 3

The training was focused on developing an interpretation key, description of the imagery and how interpreters would classify a plot. Yelena mentioned that to enhance consistency among interpreters, it is necessary to have an interpretation key. In developing an interpretation key, the same plot must be captured on various platforms (google earth, bing map, planet data and google earth engine). An exercise was given to all participants to assess and create interpretation key for 4 samples. Yelena, mentioned that those interpretation keys would be added to the SOPs. She lastly presented on training and calibration before, during and after data collection.

DAY 4

Before the training started, there was a volunteer (Ebenezer, RMSC) who had been able to extract HIA information to grid for intensification with ArcGIS. He took participants through the steps of replicating as done in QGIS on ArcGIS. After that, participants presented on the previous day's exercise (creating interpretation keys for 4 samples). Yelena then took participants through the steps in saving collected data on CE and extracting to various formats for analysis.

DAY 5

Recap of the week's lessons, activity data and data collection process. On sampling design, using QGIS/ArcGIS to extract HIA information to subset. Response design-review of using collect to modify the survey, creating an interpretation key. Data management, exporting/importing data in collect earth, planning rounds of assessment.

WEEK 2 (5th – 9th October)

The second week was focused on the calculations in excel, that is, deforestation estimates and uncertainty, emission factors and their uncertainties and AD and EF and propagate their uncertainties.

DAY 6

Marieke was in charge of week 2, the training started with a presentation on how to calculate AD from an intensified systematic grid. She added that due to the heterogeneous nature of the Ghana landscape, there are some considerations needed to be taken when choosing a sampling size, the rarest of feature, what you want to find amongst others. This sampling size determines the formula to be used for calculation. To calculate the area statistics from an intensified sample, the formula below is used.

The area of variable v:

$$A_v = \sum_{i=1,\,2} \frac{n_{v,i}}{n_i} \times A_i$$

where

- $n_{v,i}$ is the number of sample plots of variable v falling in stratum i,

- n_i is the number of sample plots falling in stratum i,

- A_i is the area of stratum i.

So for example, to calculate the forest area of Moist evergreen 2×2, per the formula, you will need Nvi = number of sample plots that fall in the forest area of moist evergreen (points in the 2×2 grid)

Ni = number of sample plots falling in all of the moist evergreen (2×2 and 4×4)

Ai = total area of the moist evergreen

This method is used for all the vegetation zones

DAY 7

Marieke started the training with a presentation on how to calculate uncertainty for AD. She added that to calculate for the confidence interval around the area of a variable in a vegetation zone, the formula below is used.

$$CI\,(\pm)\,of\,A_{v,e} = 1.64 \times \sqrt{\frac{p_{v,e} \times (1 - p_{v,e})}{(n_e - 1)}} \times A_e$$

where

- $p_{v,e}$ is the estimated probability of variable v in vegetation zone v, calculated as $n_{v,e}/n_e$

- n_e is the total number of samples in vegetation zone e,

- A_e is the total area in vegetation zone e

For example to calculate for the confidence interval around the forest area of the moist evergreen (2×2)

Pve = number of sample plots in the moist evergreen (2×2) divided by total sample plots in the moist evergreen

Ne = total sample plots in the moist evergreen

Ae = total area of the moist evergreen

All these in a square root multiplied by 1.64 would determine the confidence interval around the forest area thus implying the uncertainty.

DAY 8

Before moving on to the day's purpose, there was a recap of what had been done during the week. The day's training followed with a presentation on calculating AD × EF.

DAY 9

The day's activity was on recap of all that has been done in the previous week and present week. Participants asked questions and gained clarity on them. Yelena then asked all to complete their interpretation key and upload them unto the google drive.

DAY 10

Coalition of Rainforest Nations (CfRN) has expressed interest in supporting Ghana calculate the national FRL. Therefore, on the last day, CfRN joined in on the call to discuss together how the recalculation process can be used to calculate a national FRL. Marieke gave a presentation on how feasible Ghana can replicate the process. She mentioned that a revision of current available data would be necessary to aid in deciding what data needs to be collected to scale the current assessment up to the national level. On the EFs it was mentioned that RMSC has undertaken a carbon assessment for the upper region 2016, "2019" which can be used to complement the national FRL process.

Challenges

The major challenge encountered during the training was the limited time allocated and network interruptions.

Conclusion

The knowledge transfer exercise was successful, with the Ghana MRV team well informed to replicate the recalculation with minimal support from FAO.

REPORT ON SAMPLE DATA COLLECTION WITH COLLECT EARTH

Table of Contents

ABBREVIATION

AOI	Area of Interest
CAD	Computer Aided Design
CEO	Collect Earth Online
CFP	Carbon Fund Participant
FAO	Food and Agriculture Organization
FCPF	Forest Carbon Partnership Facility
PI	Photo Interpretation

1. Introduction

Ghana's ER Programme was included in the FCPF portfolio in June 2017, on the condition that the deforestation and degradation assessment would be improved. The ERPD included deforestation estimates following a stratified area estimate approach, a degradation estimate based on proxy data (e.g. timber extraction statistics, MODIS burned area, etc.) and an enhancement estimate based on national statistics on areas planted. Since June 2017, several activities have taken place to address the conditions raised by the Carbon Fund Participants (CFP), including a dense time series analysis of forest reserve areas and an improved stratified area estimate for which a report was prepared in July 2018.

Subsequently, FAO was contracted by the World Bank to support Ghana with the improvement of its Activity Data, responding to the conditions raised by the CFP and enabling Ghana to assess REDD+ results in 2020. The support is to enable the government develop an improved estimates of Activity Data for deforestation, forest degradation and enhancement. The data is to be generated following best practices, with a robust response design, including labelling and classification standard operating procedures to ensure the high quality of the reference data.

In view of this, a technical working group was established in-country to work in hand with the FAO team to achieve the goal of recalculating the FRL. In October, Yelena Finegold, a Remote Sensing expert with the Forestry Department of FAO (Rome) met with the Forestry Commission's MRV sub-working group and technical working group from 28th October 2019 to 1st November 2019 at the Eastern Premier Hotel in Koforidua. This was a maiden meeting organised in line with agreed timelines between Ghana and FAO.

The next line of action included the need for sample data collection by the technical working group in Ghana. This exercise was scheduled for 16th to 20th December 2019 also at Eastern Premier Hotel, Koforidua.

This report presents an overview of the work of the technical working group in Koforidua during 16th – 20th December 2019.

2. Approach adopted for the exercise

The team was tasked to interpret 2,969 plots on Collect Earth, an Open Foris tool developed by FAO for data collection. This work was a continuation of an already existing collect earth project by FAO (Mapathon Data) on 8 x 8 systematic grid. The 2,969 plots comprised a 4 x 4 intensified grid, unassessed 8 x 8 grid and, as well as Test plots.

Prior to the data collection, participants were taken through a presentation on the expected deliverables, the procedures for the data collection and a pilot interpretation of test plot file to demonstrate to members how the data was to be collected in a unified manner by taking into account how to detect the various land use distribution in plots, as listed and illustrated below:

- **Homogenous** – a plot is homogenous when all the plot area (100% or 45-49 points) is marked as one single land use.
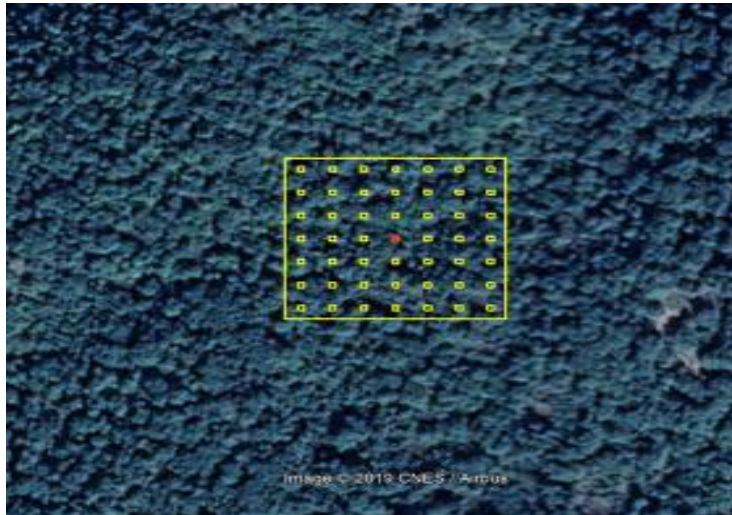
- **Distinct**- plots with more than one land use are distinctly separated
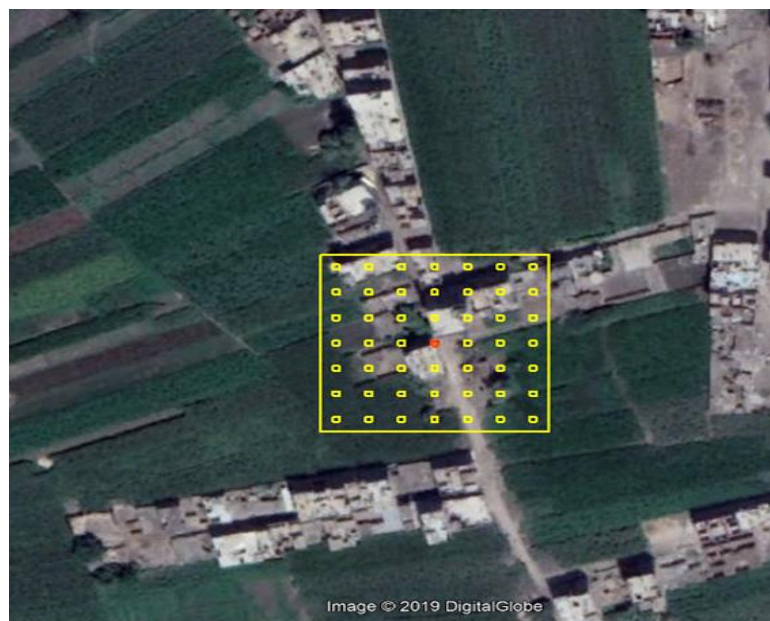
- **Overlapping**- plots with more than 1 land use within the plot that occur in the same area.

The plots were shared amongst the 11 team members with each collecting 270 points. Classification of the plots was done in Collect Earth mainly with the support of Google earth, Bing maps, planet data and Google earth engine. Each plot had 49 points and classified under the main land use systems in Ghana which is also in line with the IPCC Classification hierarchy as follows:

- Settlement = 20%
- Cropland   = 20%
- Forest   = 20%
- Grassland = 20%
- Wetland   = 20%
- Other land = 20%

The entire approached used for this sample data collection was in line with the collect earth manual developed by FAO. As part of measures to conduct a good sample data collection, the use of Photo Interpretation within the Collect Earth Online environment forms part of the preferred option. This is because it takes into account human interpretation based on experience and skills developed by the interpreter overtime. This enabled a smooth and clear data collection process. The following formed part of the planning process which was adpoted by FAO and Ghana to design the entire plot sample:
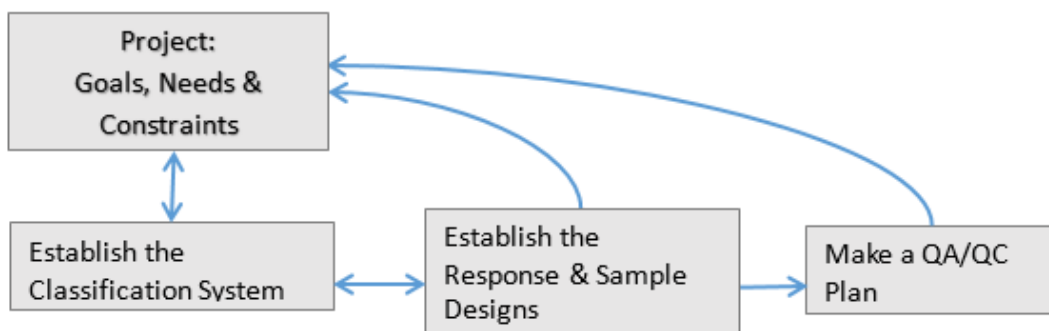
In most cases, there are about eight elements that every skilled interpreter uses in interpretation of an image;

- **The image tone:** The light and darkness of a region within an image

- **Image texture:** the apparent roughness or smoothness of a region within an image.
- **Shadow:** may reveal information about the size and shape of an object which cannot be discerned from an overhead view alone.
- **Pattern:** the arrangement of individual objects in distinctive recurring patterns, such as buildings in an industrial complex or fruit trees in an orchard.
- **Association:** the occurrence of one type of object may infer the presence of another commonly associated object nearby.
- **Shape:** manmade and natural features often have shapes so distinctive that this characteristic alone provides clear identification.
- **Size:** the relative size of an object related to other familiar objects gives the interpreter a sense of scale, which can aid in the recognition of objects less easily recognized.
- **Site:** refers to topographic position. For example, certain crops are commonly grown on hillsides or near large water bodies.

The results of image interpretation are most often delivered as a set of attributed points, lines, and/or polygons in any one of a variety of CAD or GIS data formats.

3. Quality Assurance and Quality Control

It is ideal that every data collected is always of exceedingly high quality, devoid of error and biases. This was therefore achieved, through the process of Quality Assurance and Quality control to trace any error that emerged throughout the data collection process.

In other to avoid any form of enormous and misleading results, especially for the purpose of this work, the interpreters were made to follow the following approach as means of proper scrutinizing the data collection process:

- Interpreters made to test samples that mimicked the work that was to be done; interpreters were introduced to images that were quite similar to the actual work they were tasked to undertake.

- Giving the same image to different interpreters for interpretation; a second interpreter has the opportunity of re-classifying already classified data. That is, by interpreting 10% of the total sample plots that each team member worked on. Thus, after collecting data on 270 plots, each team member also collected data on 10% of the plots assigned to another interpreter, i.e., assessment of the first 30 plots of each 270 shared points. This was to ensure that accurate classification had been done by each interpreter in line with the IPCC classification hierarchy.

- To avoid introducing more bias in the data collection process, different land use were discussed and a guide created for all classes to help participants in their interpretation.

- In situations where an interpreter encountered difficulty with interpretation of a plot, the screen was projected for discussion by the whole group to reach consensus on the classification for the plot.

4. Some challenges encountered with the interpretation of samples

- Quality of the planet data was very bad; almost all available images had cloud cover. The rest of the images were hazy and did not facilitate the interpretation process.

- Difficulty in classifying shaded cocoa from that of forest. Whether the distinct or overlap was selected, the land-use was assigned to the category with majority of number of control points. To change this, one has to choose more control points for the forest area to outweigh the cropland. This leads to bias when there is only about 20% tree cover but has to be made 50 or 60%.

- Some points only had one year of available high resolution imagery for the interpretation of the sample plot this therefore made it difficult for interpreters to give a substantive interpretation within an AOI where a sample plot had been mapped.

The sampled images below illustrate the difficulty most interpreters had relying mostly on Planet imagery and Bing maps; especially in extreme cases where there were little or no historical imagery within the area of interest / selected plot area.

Sample image from planet data
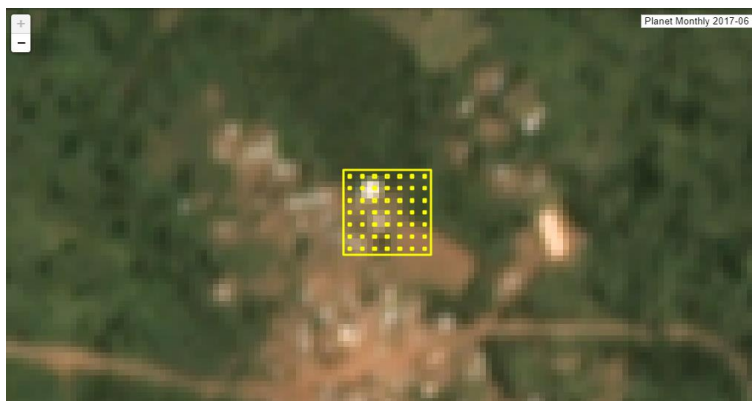
Sample image from Google Earth

Sample image from Bing map.

5. Next Steps.

It is expected that the National REDD+ Secretariat would submit the collected data to the FAO team to carry out further data calibration alongside the processed data from BFAST to generate a change map for Ghana. The feedback from FAO team will therefore determine the way forward.

Methodology for Activity Data

Virtual capacity building on REDD+ FCPF reporting
Week 1: Activity data

Yelena Finegold

FAO Forestry Officer

Virtual capacity building

28 September – 2 October

Outline: Methodology

- Sampling design

- Response design

- Quality management

Sampling design

- Sample design – stratified systematic

- Stratification

- Sample size
  - Planning QA/QC

- Intensification

Deforestation estimate sample only
estimate versus post-stratified estimate
with Hansen data
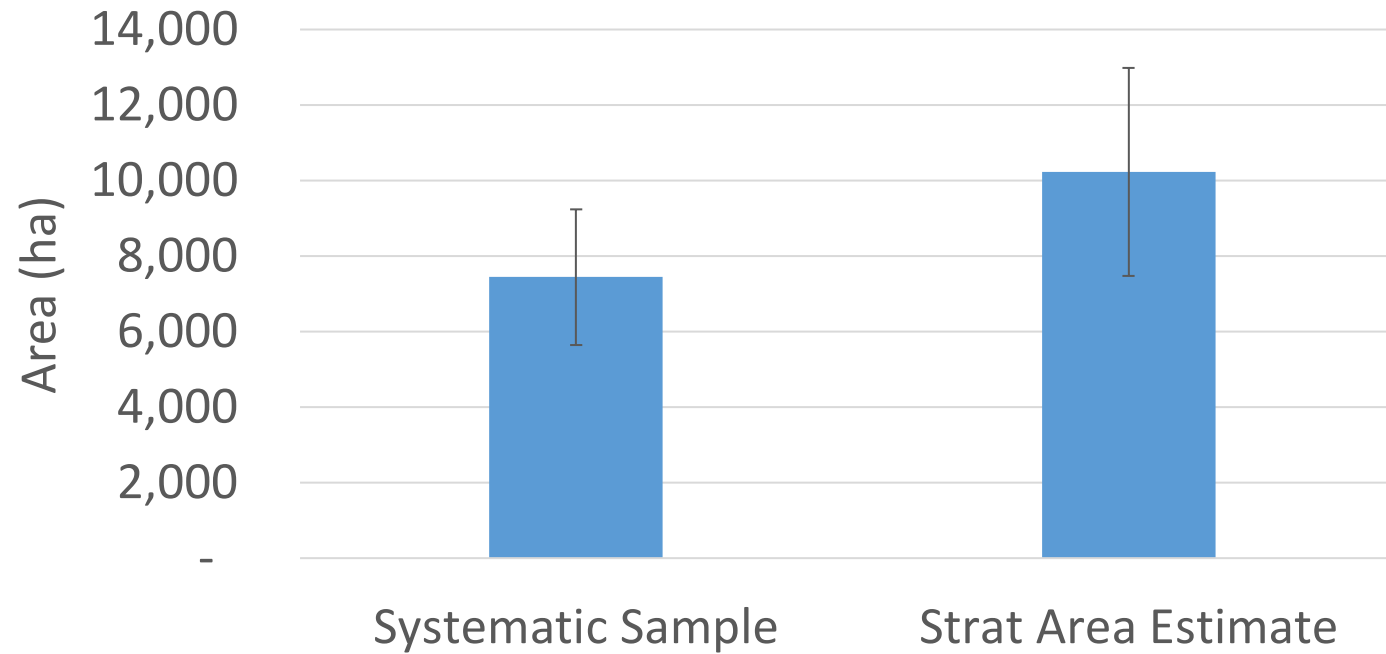
Confidence interval sample only estimate = +/-24%, confidence interval post-stratified samples with hansen map = +/-27%

The map is **not** an efficient stratifier since it increases the CI

How many samples do we need to sufficiently assess change?

Sample size

Simulations with Hansen data to test ideal intensity

Reference period: 2005-2014

Result of 10 iterations 2005-2014

Ha of tree cover loss

Sampling intensity: distance between samples (km)
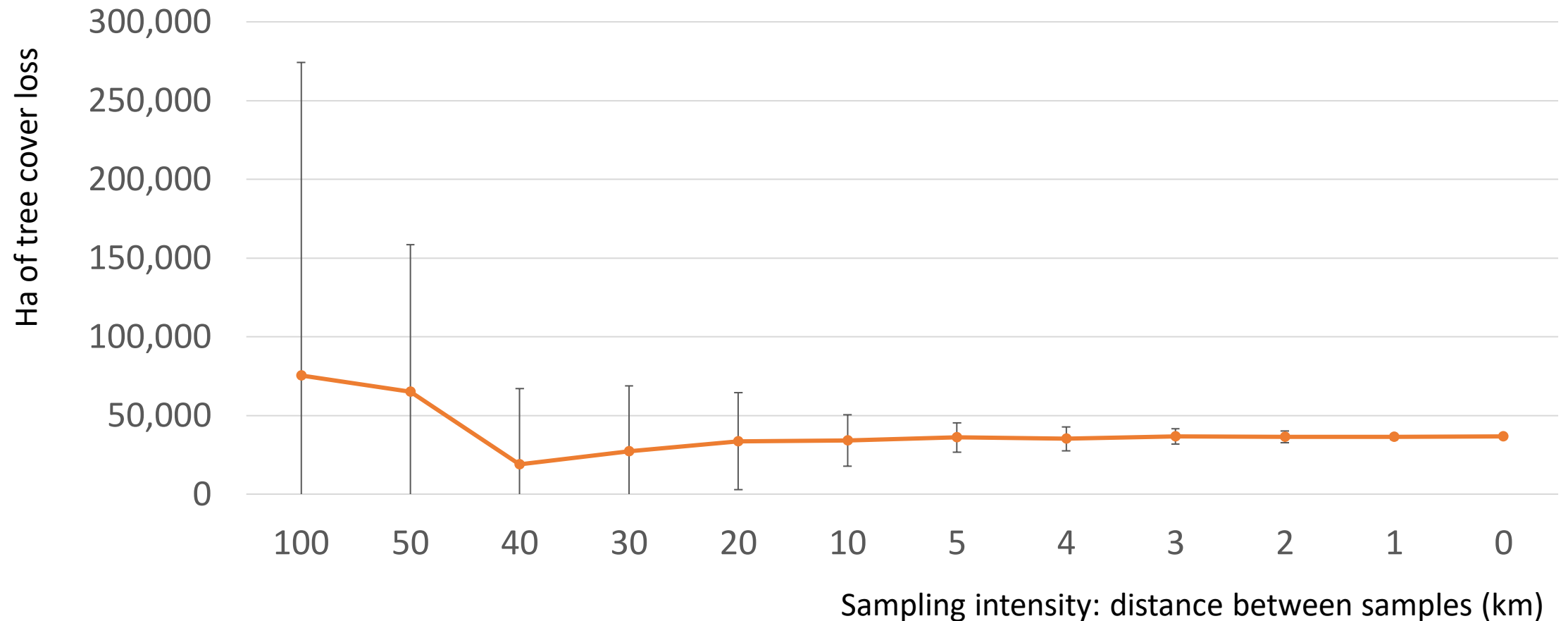
Simulations with Hansen data to test ideal intensity

Reference period: 2017-2018

Result of 10 iterations 2017-2018
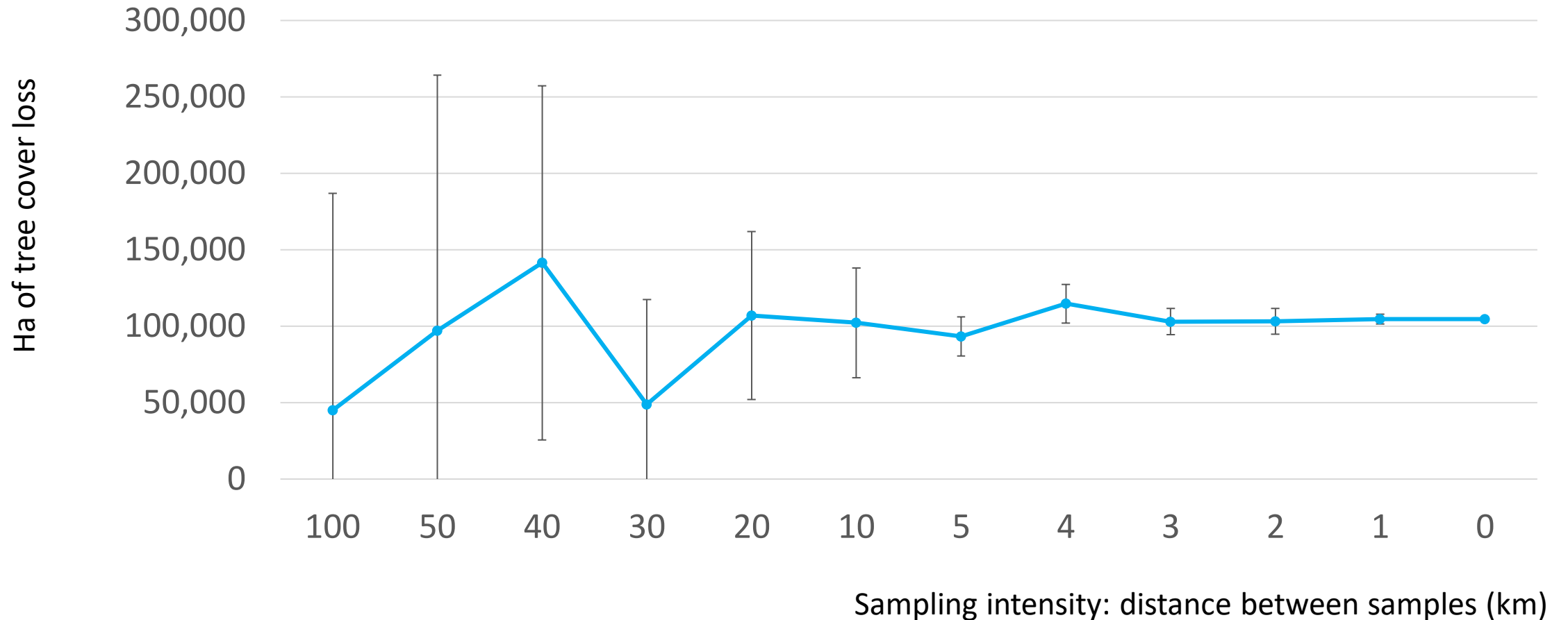
Some considerations

- Tree cover loss in Global Forest Change data is 6 times larger area than deforestation (meaning sampling density needs to be higher)
- We need to calculate results by forest type (meaning sampling density needs to be higher)
- So we need a larger density….

...but the increase in density results in an exponential increase in # sample points

Sampling density and total number of samples
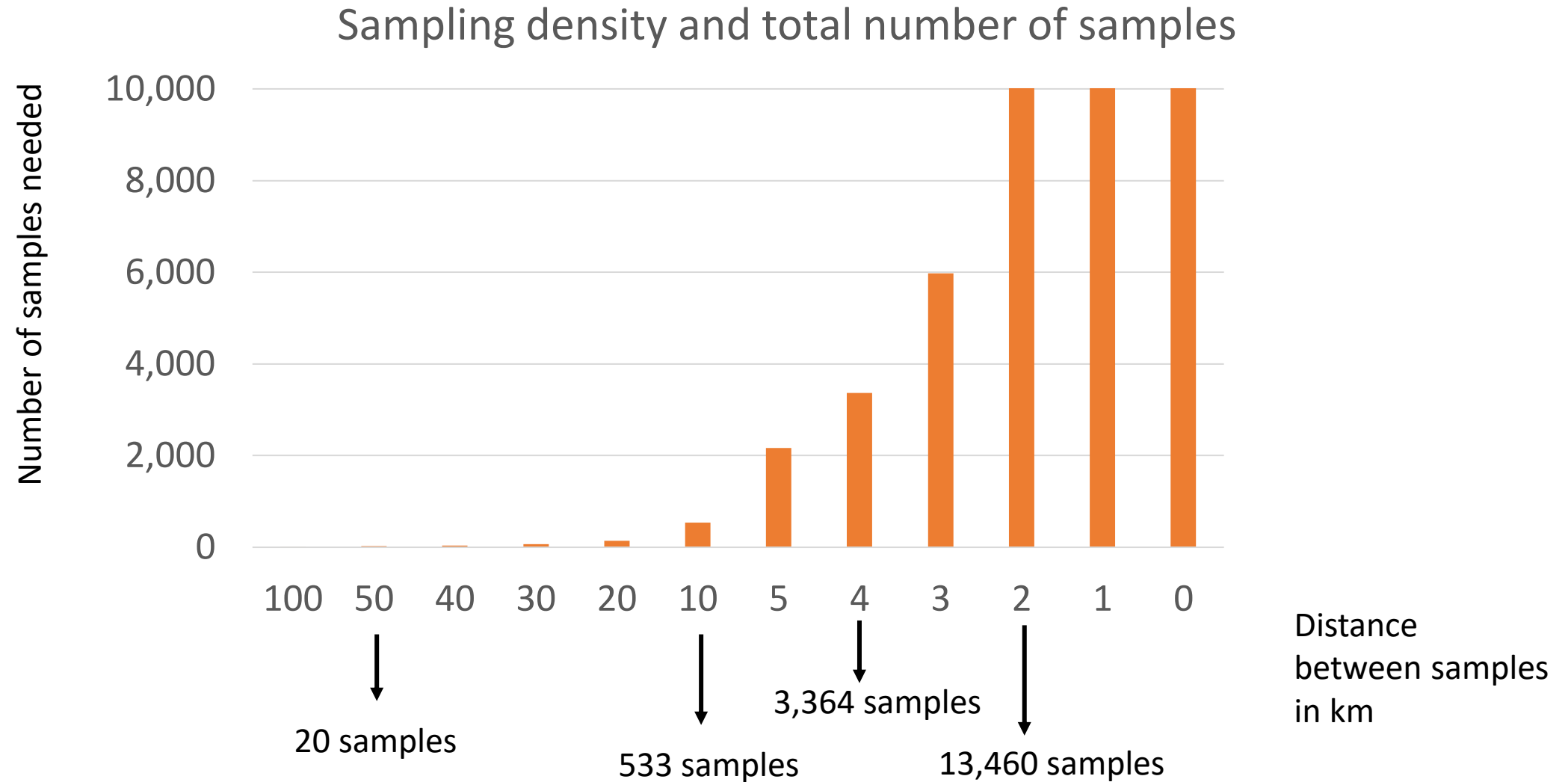
Stratification for intensification

- Using Ghana's forest mask
  - (intensification to 2 x 2 km)
- Using Upland evergreen as a stratum
  - (intensification 1 x 1 km)
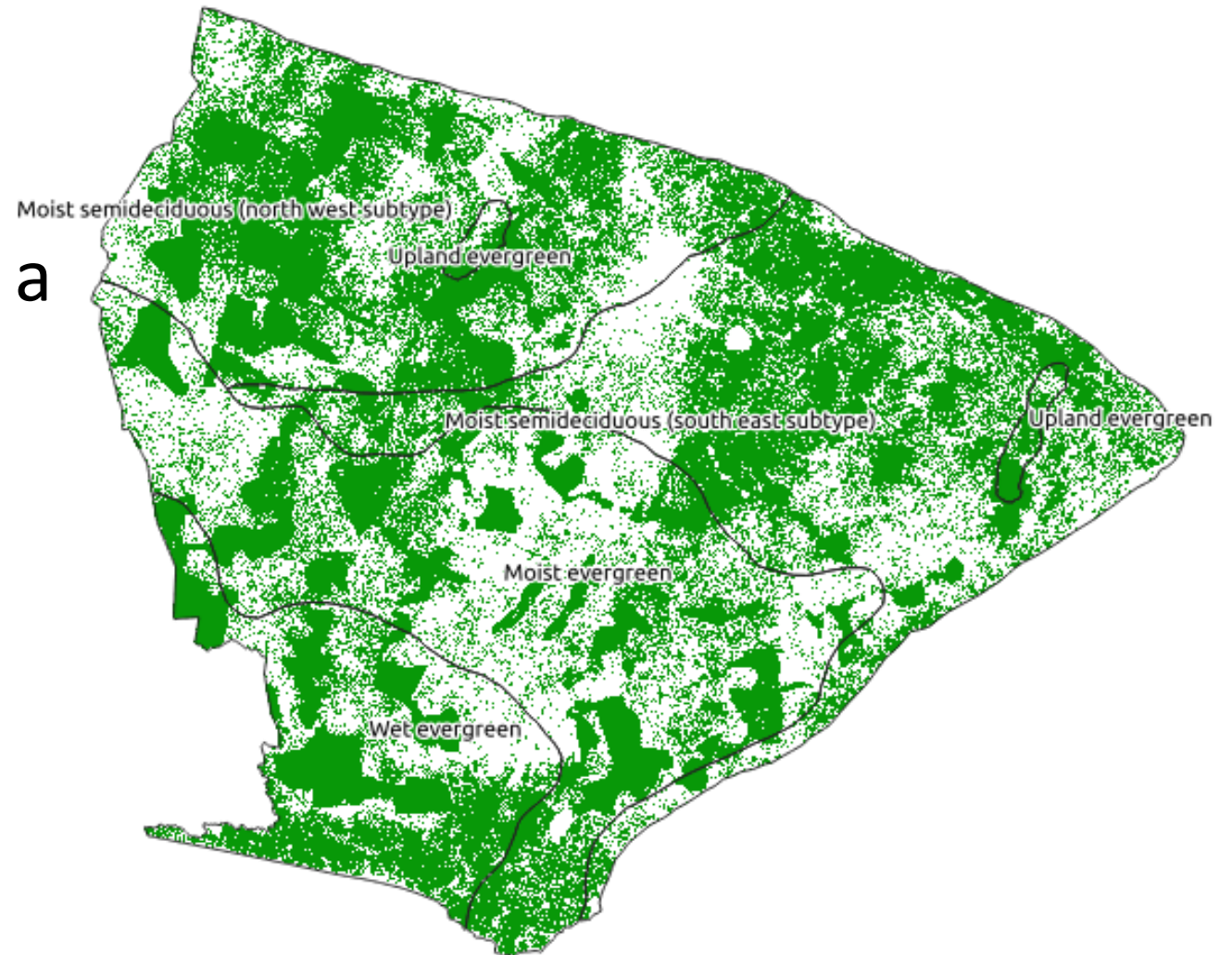
8 x 8 grid - national scale
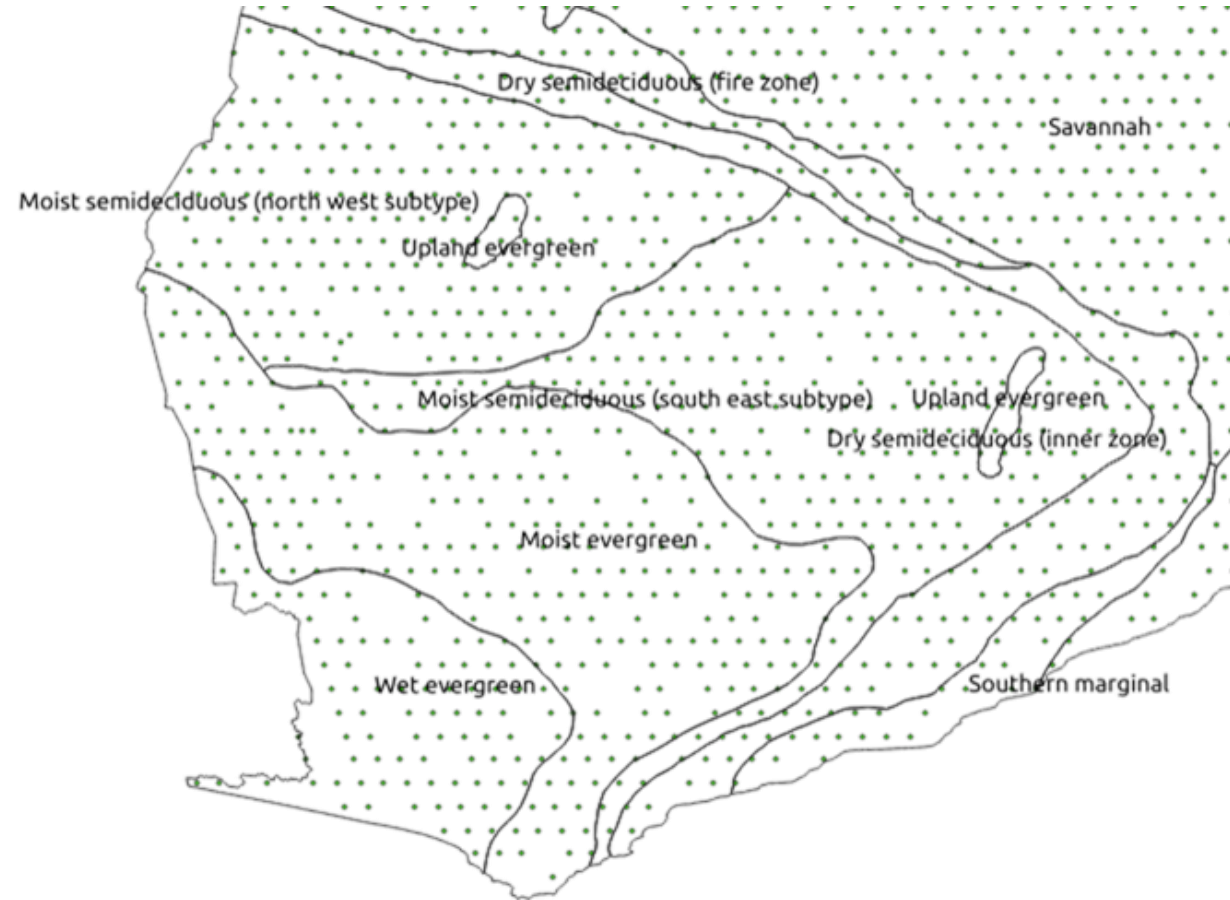
4 x 4 intensification in GCFRP
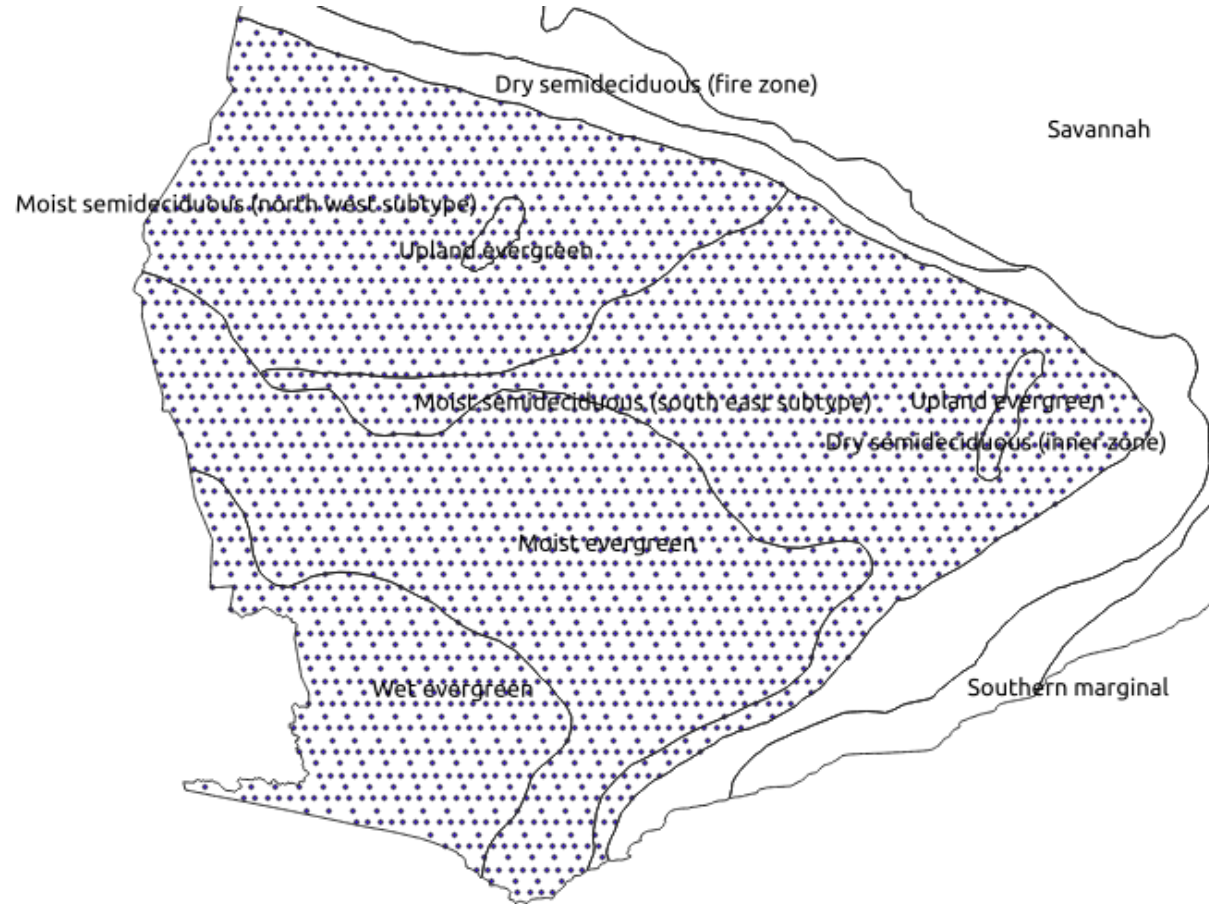
2 x 2 intensification in GCFRP forest
1 x 1 intensification in upland evergreen
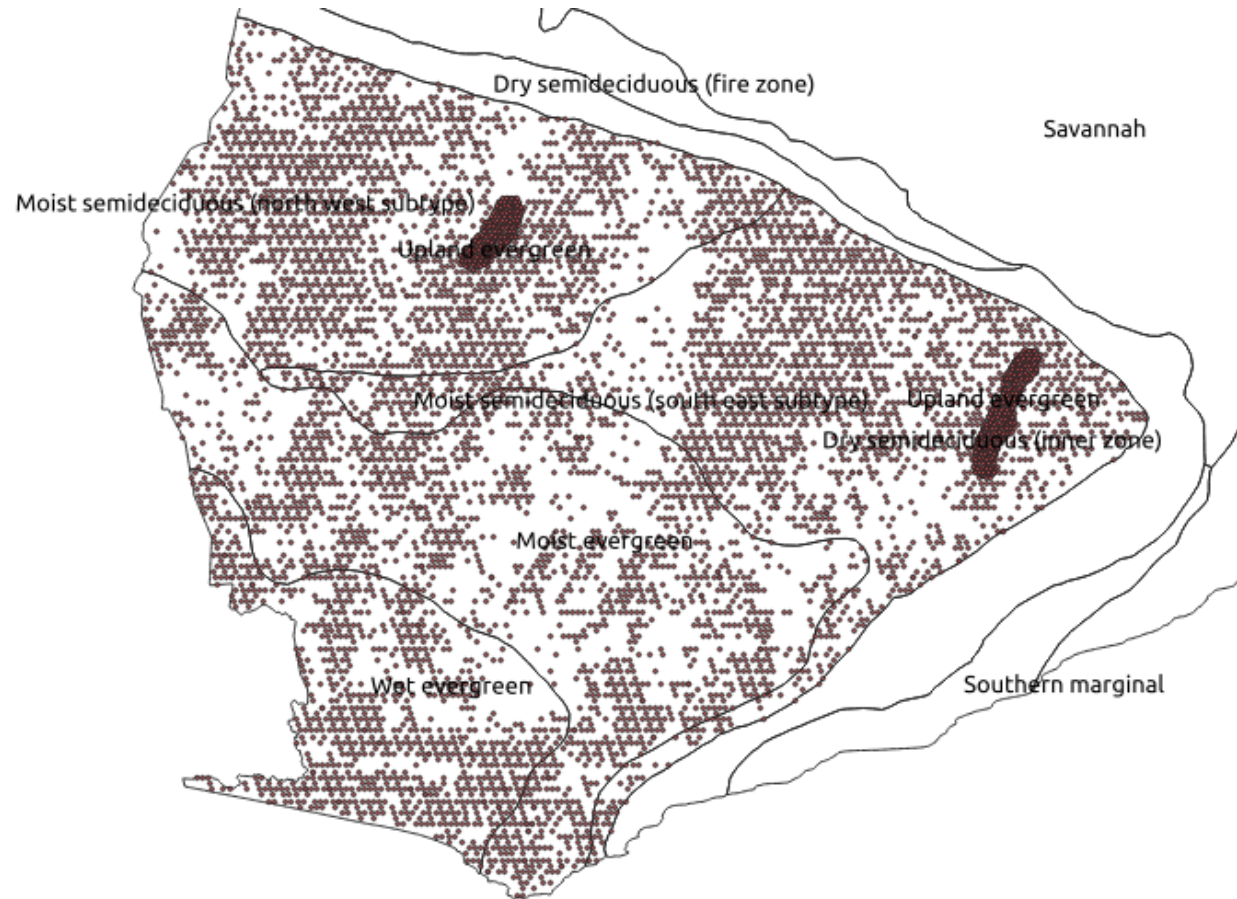
All samples
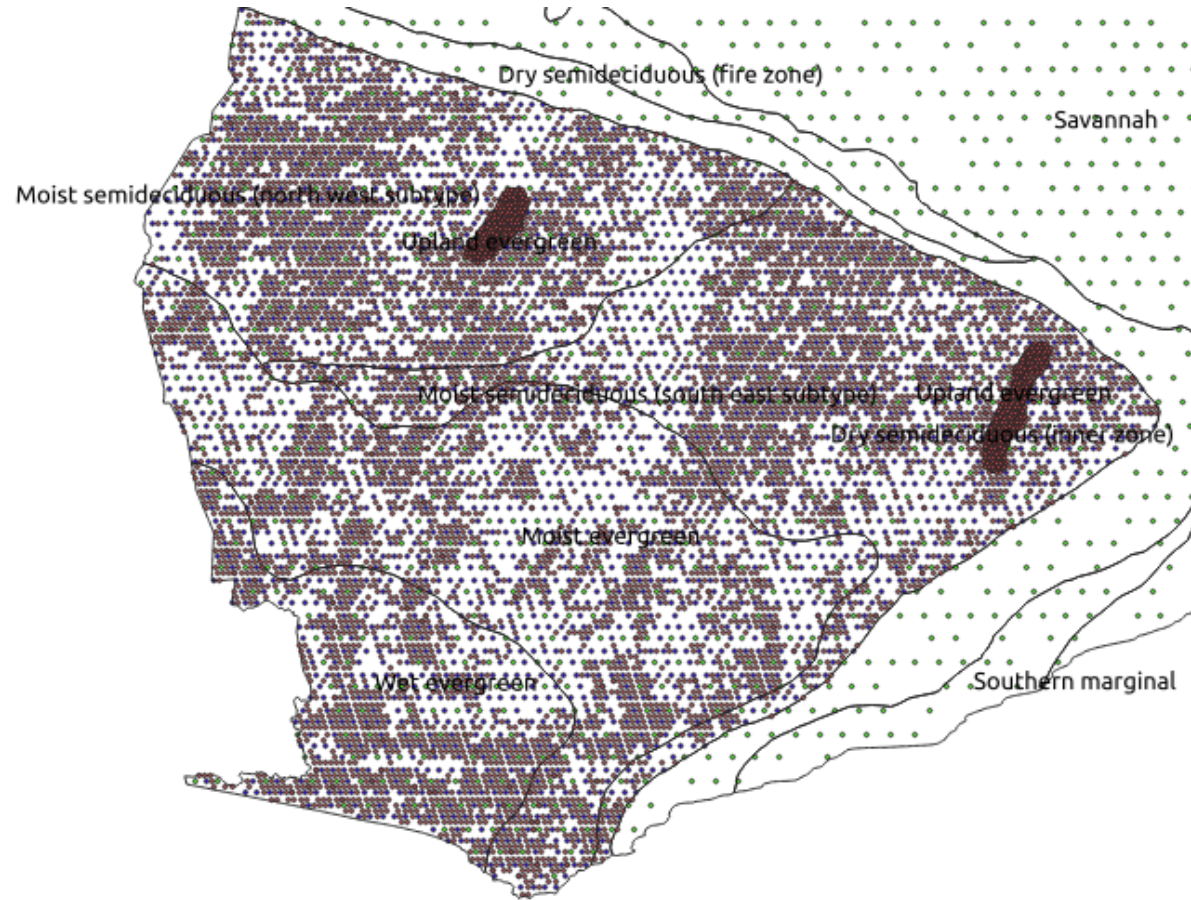
How to assess the samples?

Response design

Spatial support unit
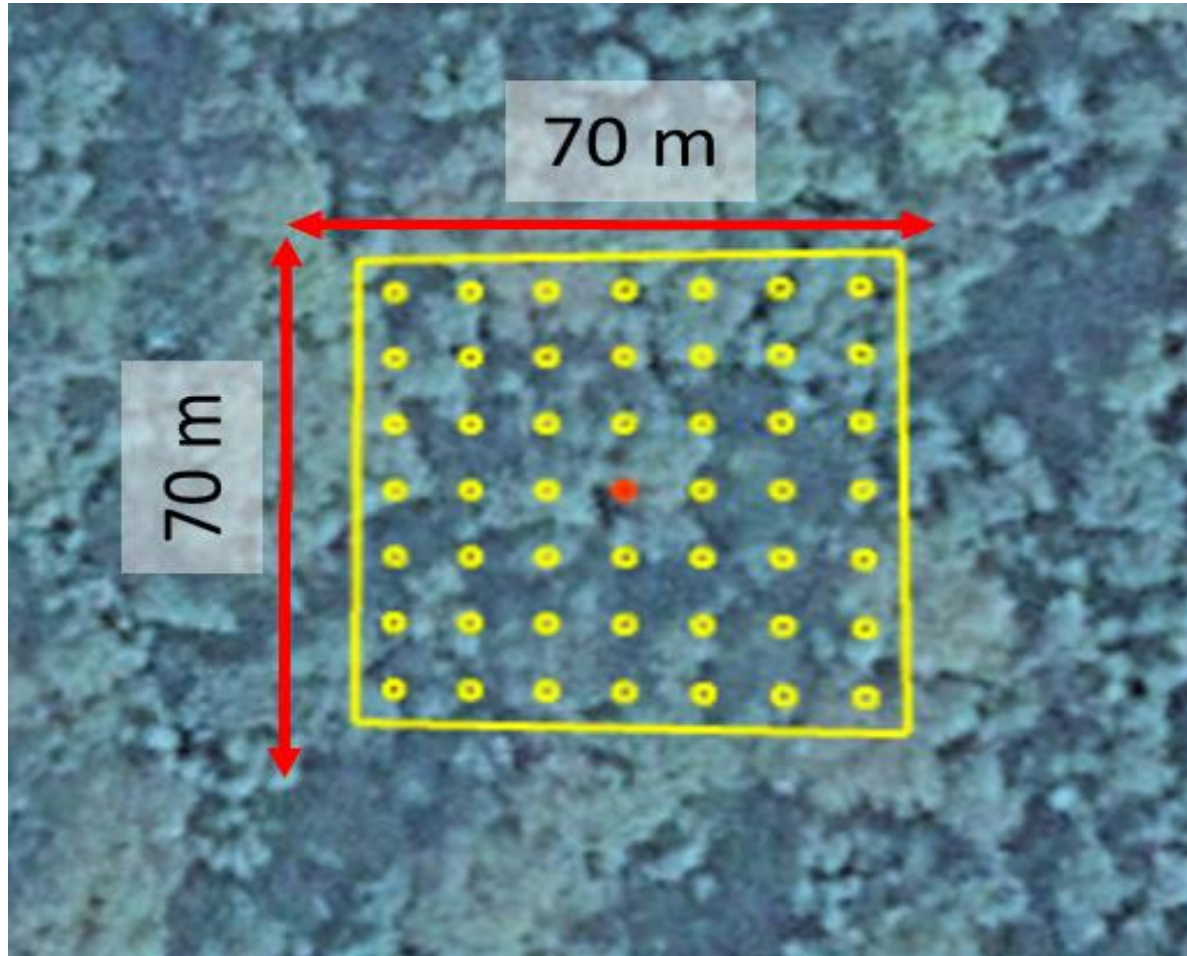
Classification system
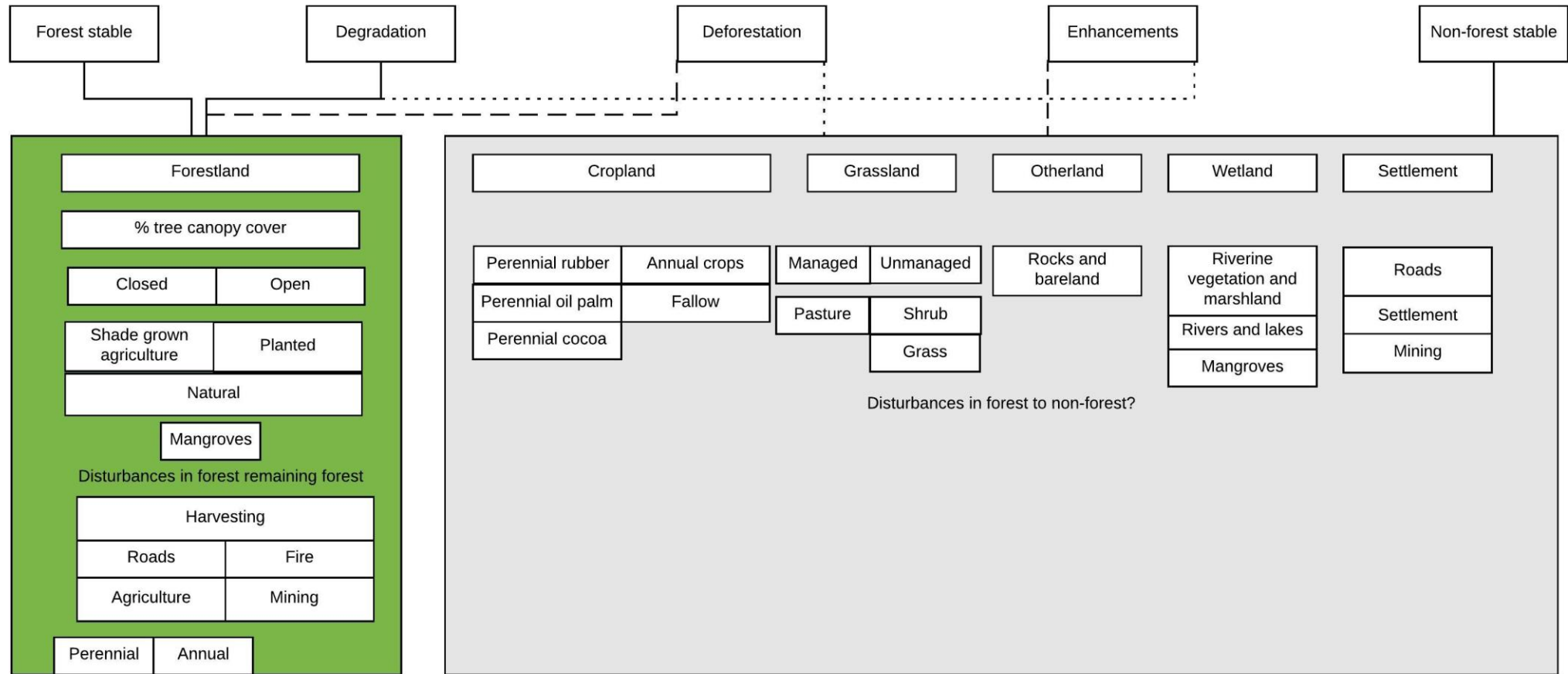
Response design and labeling
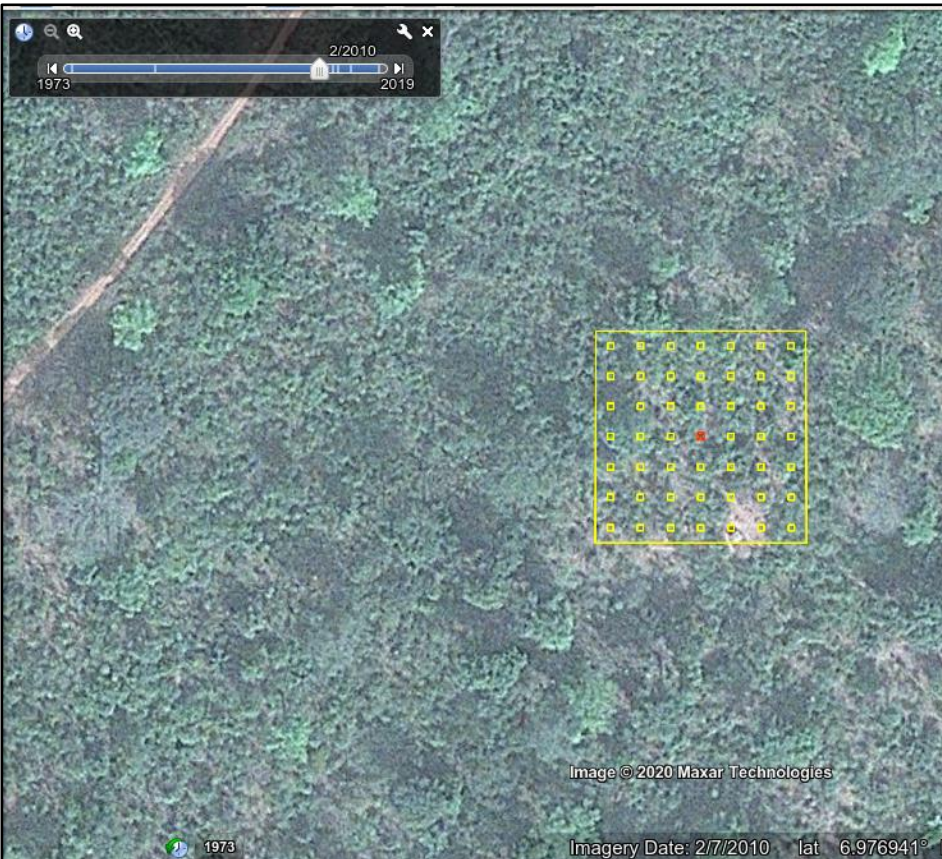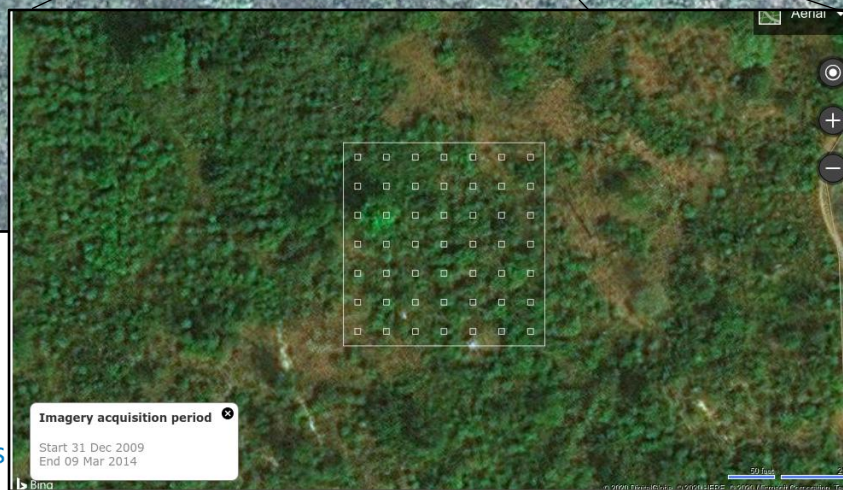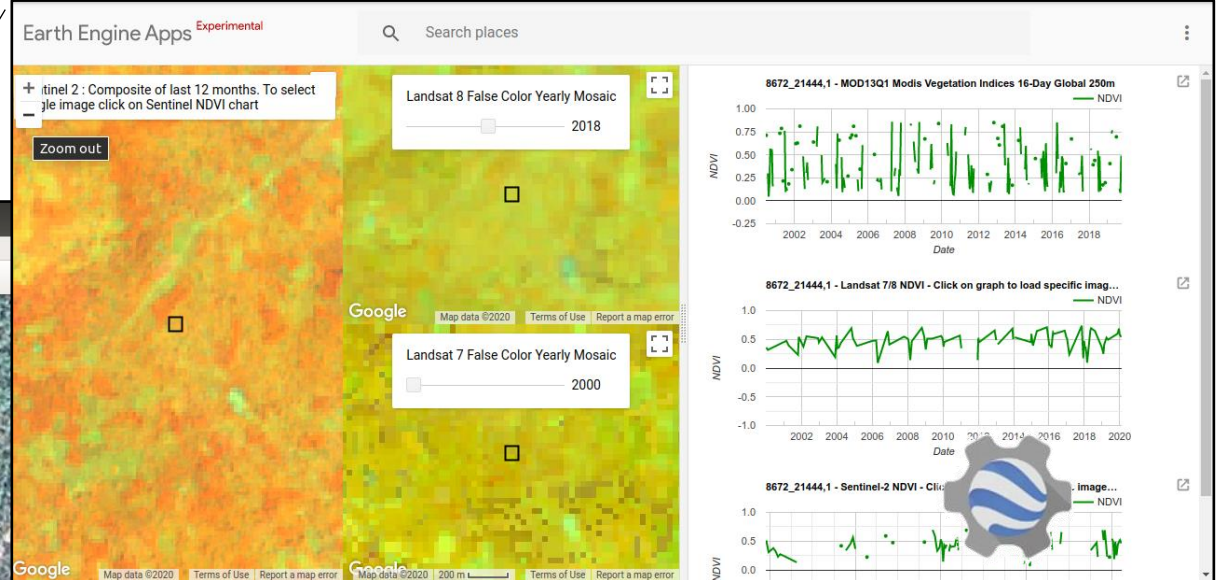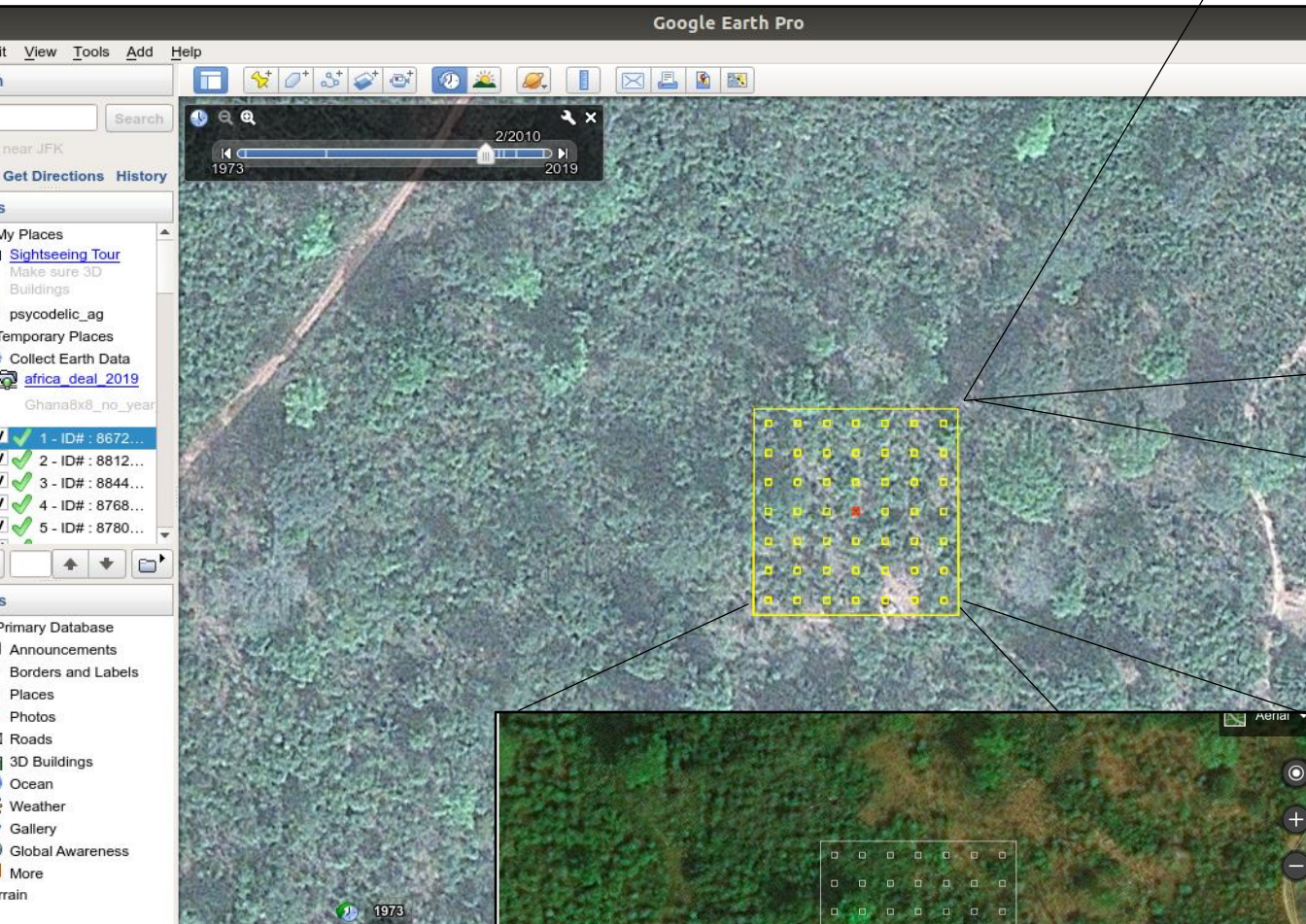
How to assess the quality of the samples?

Quality management

Training and team work

QA & QC planning

- Goal - Ensure data quality
- Planning must start at the sample design phase
- Quality Assurance (QA) – The process established at the beginning of a project and followed throughout
- Quality Control (QC) – The discrete tasks that are executed to ensure that certain standards are met
- Assess resources and evaluate methods to make a QA/QC plan
- Document procedures in standard operating procedures (SOPs)

Virtual capacity building on REDD+ FCPF reporting
Week 1: Activity data

Yelena Finegold

FAO Forestry Officer

Virtual capacity building

28 September – 2 October

Mic check and Zoom housekeeping

Objectives

- Ghana gets a full understanding of the data collection and analysis
- Ghana will be able to repeat the calculations with minimal guidance from FAO

Don't forget:

➢ Let us know what you want us to focus on

➢ We are flexible and rely on your feedback!

Calculation chain

Sampling design

How to create your sample (size, plot location)?

→

Response design

What information to collect? How to interpret satellite imagery?

→

Data management

How to get all data from different interpreters into a single database

↓

Calculate emissions (AD x EF) and propagate uncertainties

Review all excel calculations

←

Calculate emission factors and their uncertainties

Review all excel calculations

←

Calculate deforestation estimates and uncertainty

How to extract from all those observations the relevant averages?

This week

- Daily calls from 1.30pm – 3.30 pm Accra time
- Monday and Wednesday will include practical guidance on the subjects
- Tasks will be assigned to review SOPs and complete practical exercises
- Tuesday, Thursday and Friday will be an open format session to address questions and needs from participants

Monday: Sampling and response design

- Review of concepts
  - Sampling design
  - Response design
- Data requirements: shared Google Drive
  - Sampling intensification files
  - HIA shapefile
  - Collect earth project
- Review of SOPs 1 and 2

Tuesday: practical session on sampling and response design

- Review questions
- Technical support
- Review comments on SOPs 1 and 2
- Transfer of Planet data
- Optional: creating offline time series clips for visual interpretation from downloaded Planet data

Wednesday: Interpretation key, data management and QA/QC

- Review of concepts
  - Response design: Interpretation key
  - Training
  - Data management
  - Data quality assurance and control
- Data requirements
  - Assessing samples in Collect Earth
- Review of SOPs 2 and 3
- Mid-week check-in of needs and format

Thursday: practical session on Interpretation key, data management and QA/QC

- Review questions
- Review interpretation key
- Training questions
- Review comments on SOPs 3 and 4

Friday: practical session on Interpretation key, data management and QA/QC

- Review questions

- Data management

- QA/QC exercise

- Review comments on SOPs 3 and 4

Virtual capacity building on REDD+ FCPF reporting
Week 1: Activity data

Yelena Finegold

FAO Forestry Officer

Virtual capacity building

28 September – 2 October

Calculation chain

Sampling design

How to create your sample (size, plot location)?

Response design

What information to collect? How to interpret satellite imagery?

Data management

How to get all data from different interpreters into a single database

Week 1

Calculate emissions (AD x EF) and propagate uncertainties

Review all excel calculations

Calculate emission factors and their uncertainties

Review all excel calculations

Calculate deforestation estimates and uncertainty

How to extract from all those observations the relevant averages?
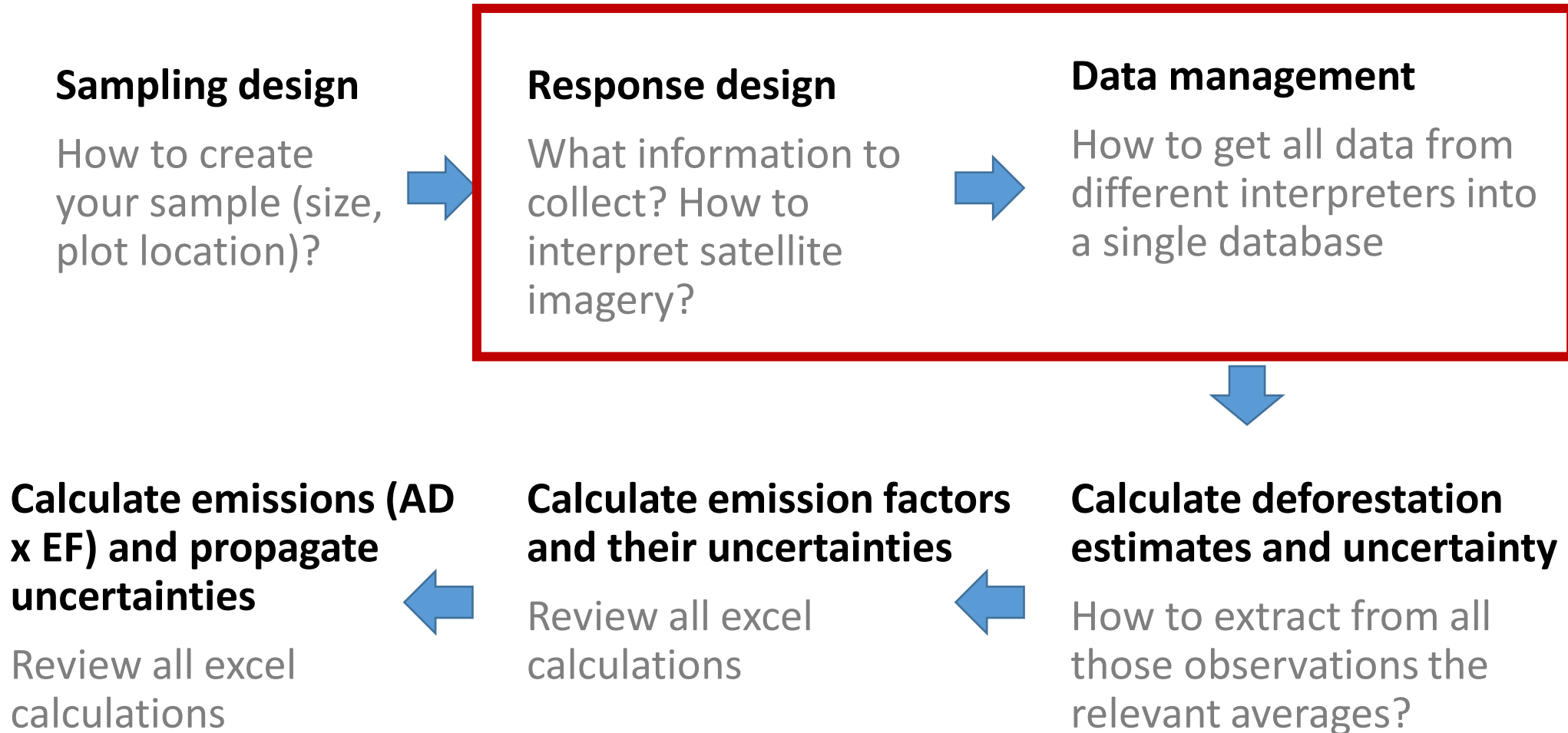
Week 2

Why data collection?

All maps have errors...

All samples have errors... ➡ that are mostly avoidable, with good quality management practices

consistency between interpreters and over time

consistency between interpreters and over time

Training & Calibration

- review response design and labelling protocol

- review interpretation key

- review and discuss example plots as a group

- assess plots individually and compare results between interpreters

  - discuss discrepancies in classification and come to a consensus classification

- modify the labelling protocol as needed

Earth Engine (Landsat 7 - 2008)

Google Earth Historical Image

Bing Maps (Free)

Google Earth Latest Image

Earth Engine (Sentinel 2)

MOD13Q1 Vegetation Indices 16-Day Global 250m

NDVI

10,000

7,500

5,000

2,500

0

-2,500

2001 2003 2005 2007 2009 2011 2013 2015 2017

Date

Google Earth Engine Code Editor

Earth Engine (Landsat 7 - 2009)

Elementos(A) Elementos Anteriores Elementos(B)
Perturbações Erosão Uso da Terra Produtividade Agrícola
UNCCD

Imagem de alta resolução disponível

2016

Vegetação

Tipo	Cobertura
Árvore	0%
Arbusto	0%
Palmeira	0%
Cultura agrícola	90-100%

Número de Árvores

0

Número de Arbustos

0
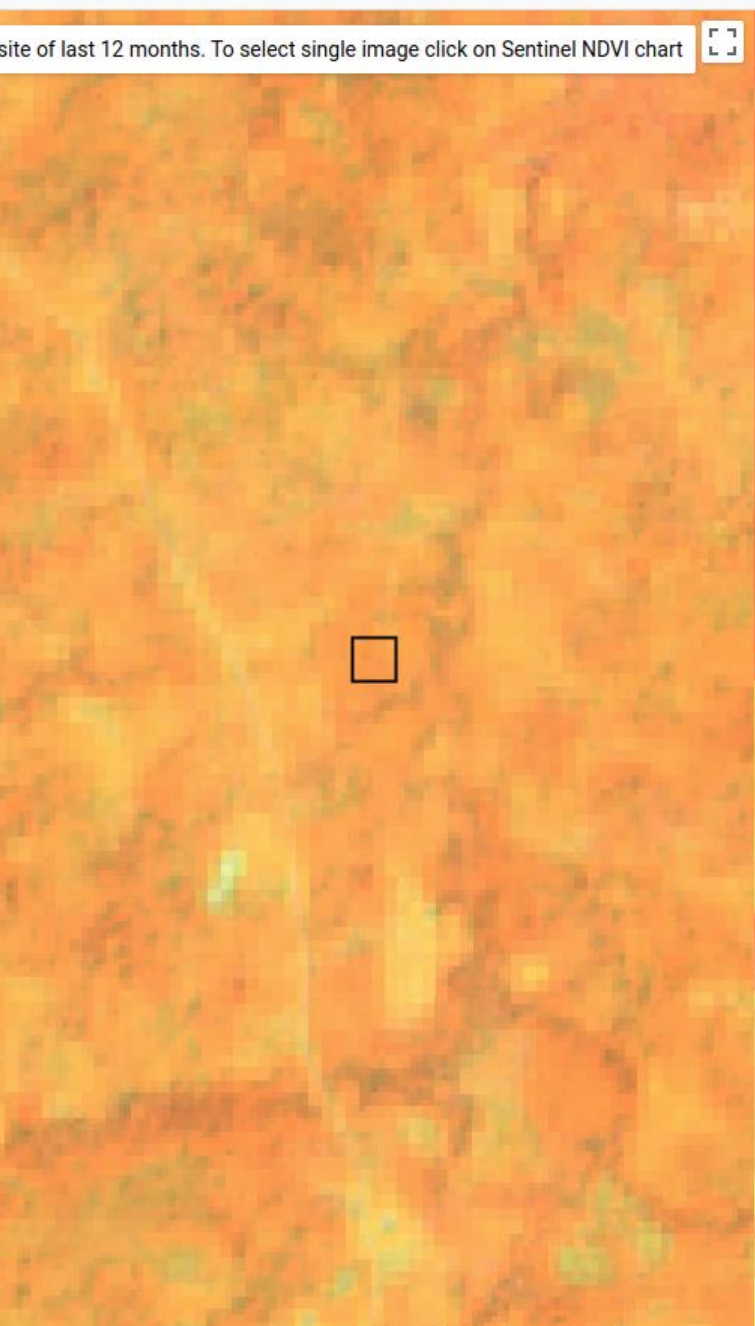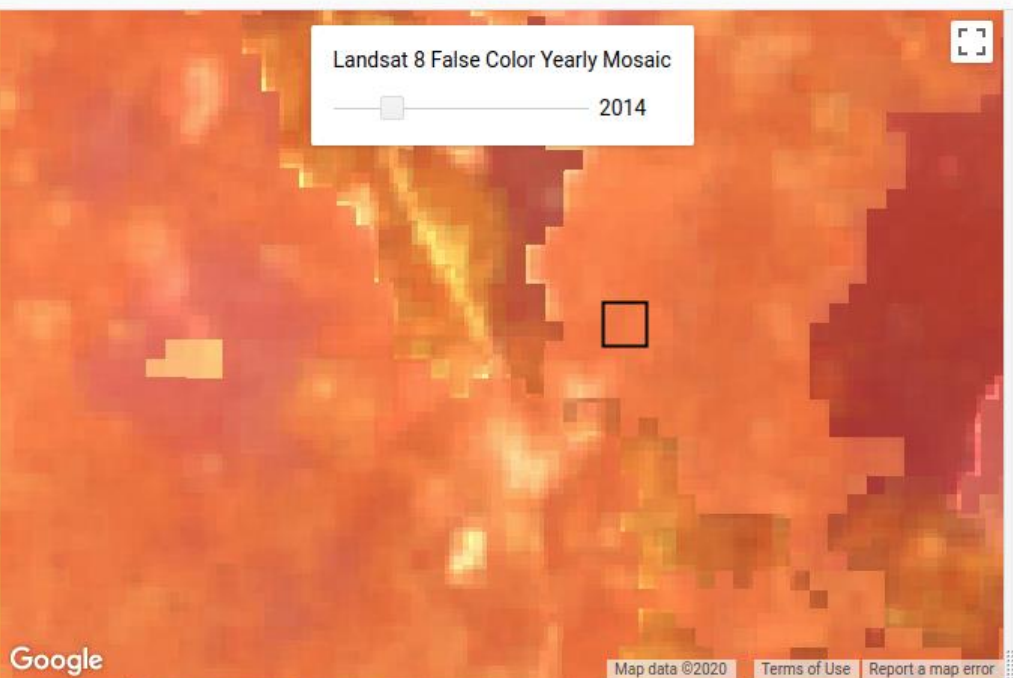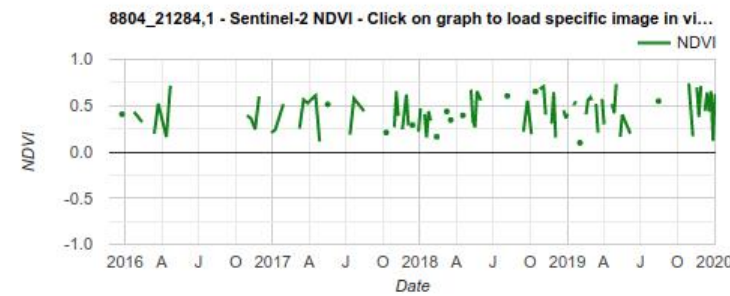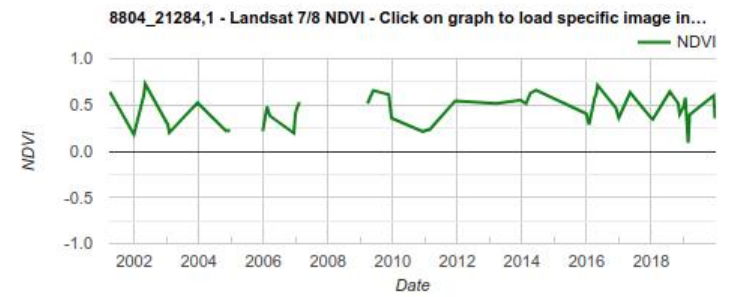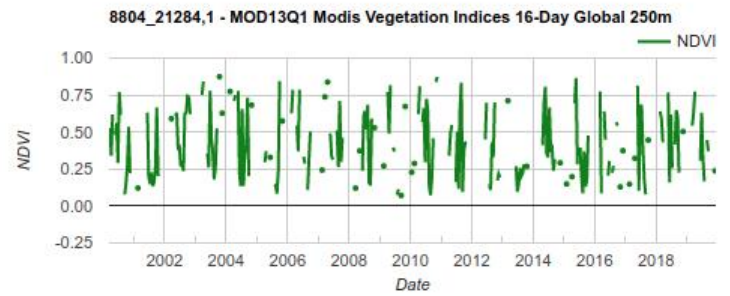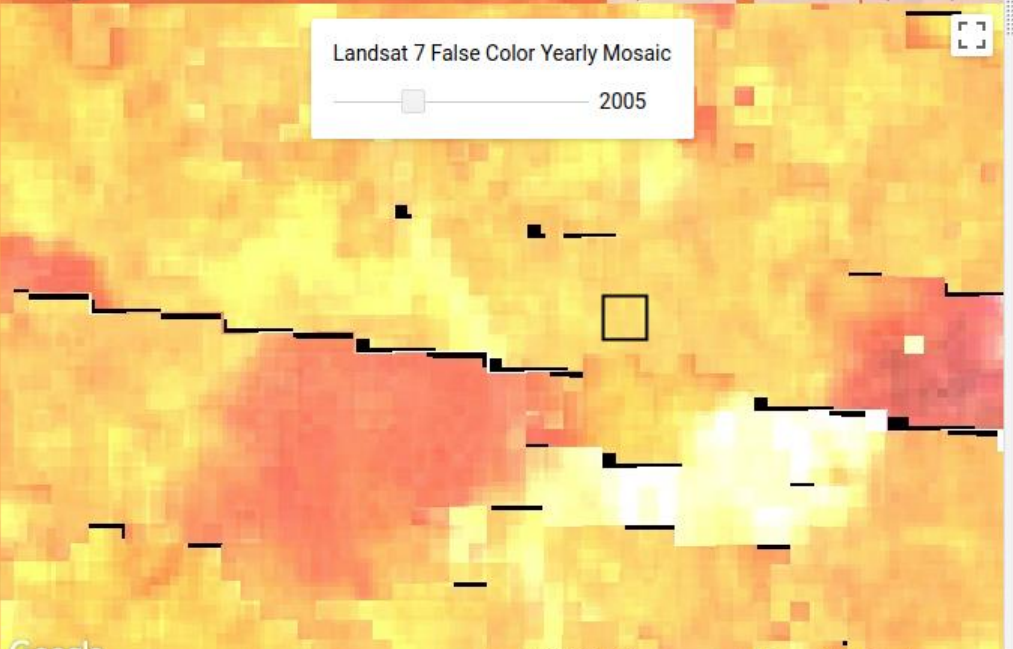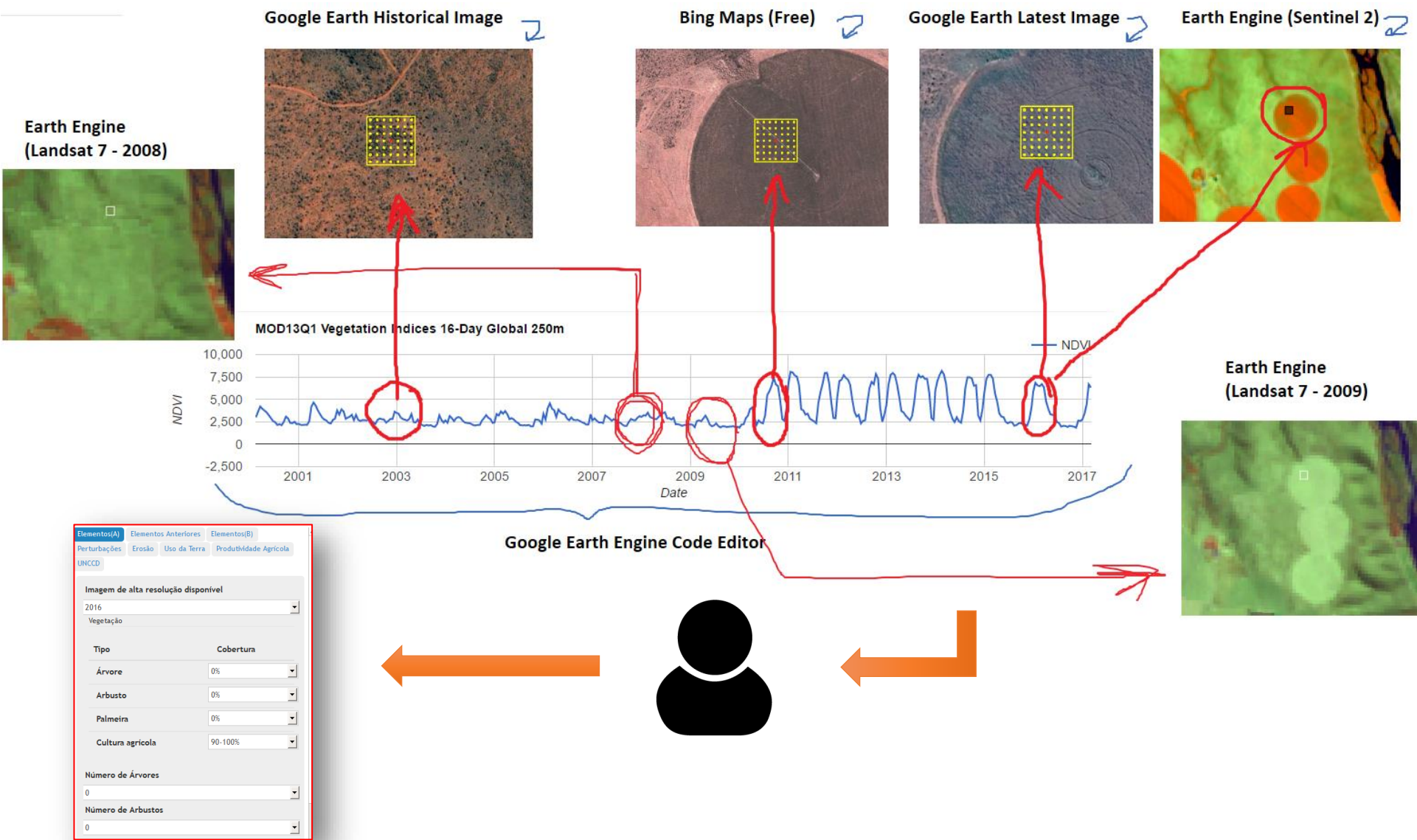
False Color Guide

In Sentinel-2:

- The color orange in false color images indicates vegetation.

- Textured dark orange is an indication of trees in the area.

Less Vegetation
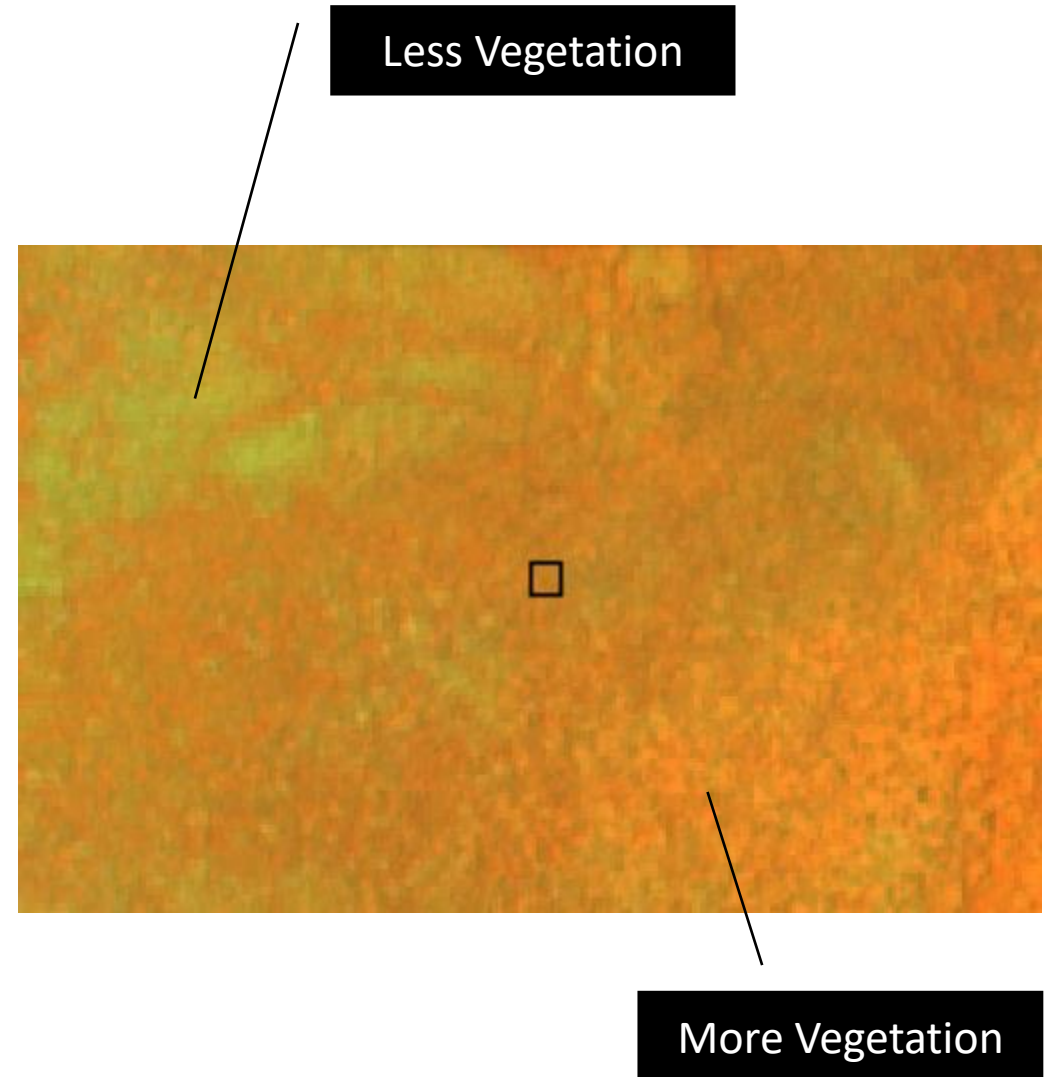
More Vegetation

NDVI Chart Guide

- The Sentinel-2 chart is interactive. You can click on the green line to change the image.

Plot - Sentinel-2 NDVI - Click on graph to load specific image in visor
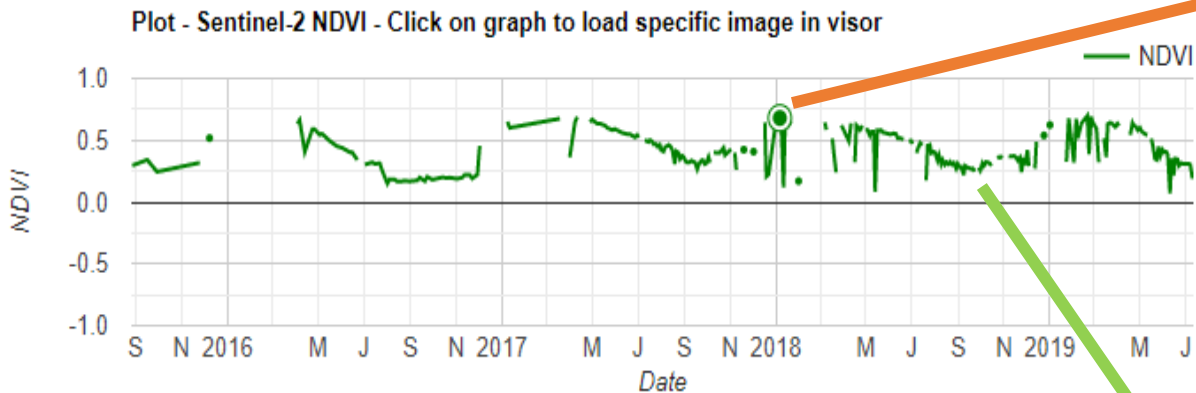
- The high values mean high vegetation (orange in the image)

- The low values are dry or no vegetation (pale green in the image)

Wet Season
A Lot of vegetation (shown in orange)
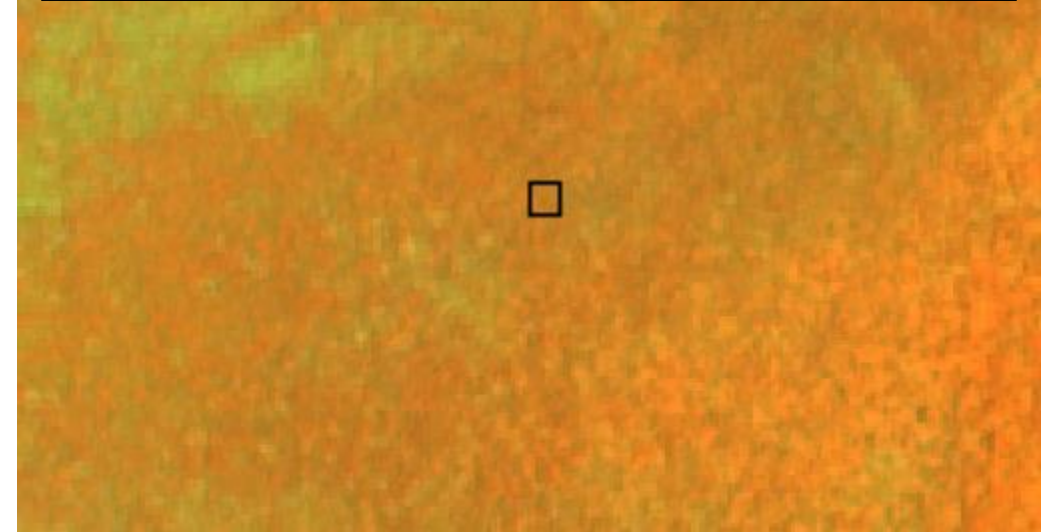
Dry Season
Very little vegetation (dry vegetation shown as pale green in the false color image below)
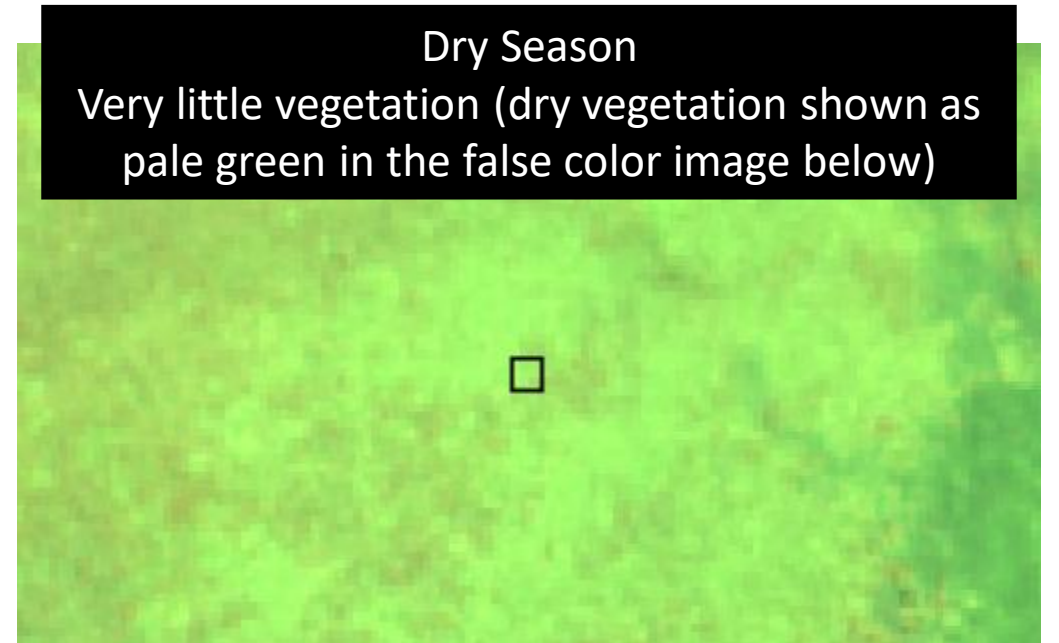
Same Plot

Croplands in the false color

Crop fields often have a geometric shape.

They can be bright orange (when growing) or pale green (when off season).

Round crop fields in the forest are normally associated with Shifting cultivation (not in the image).

Crops

Crops

Crops

Tree cover

Tree cover

Land Use Change (LUC)

Plot - MODIS NDVI 16-Day + Burned Area MODIS MCD64A1

— Fire — NDVI
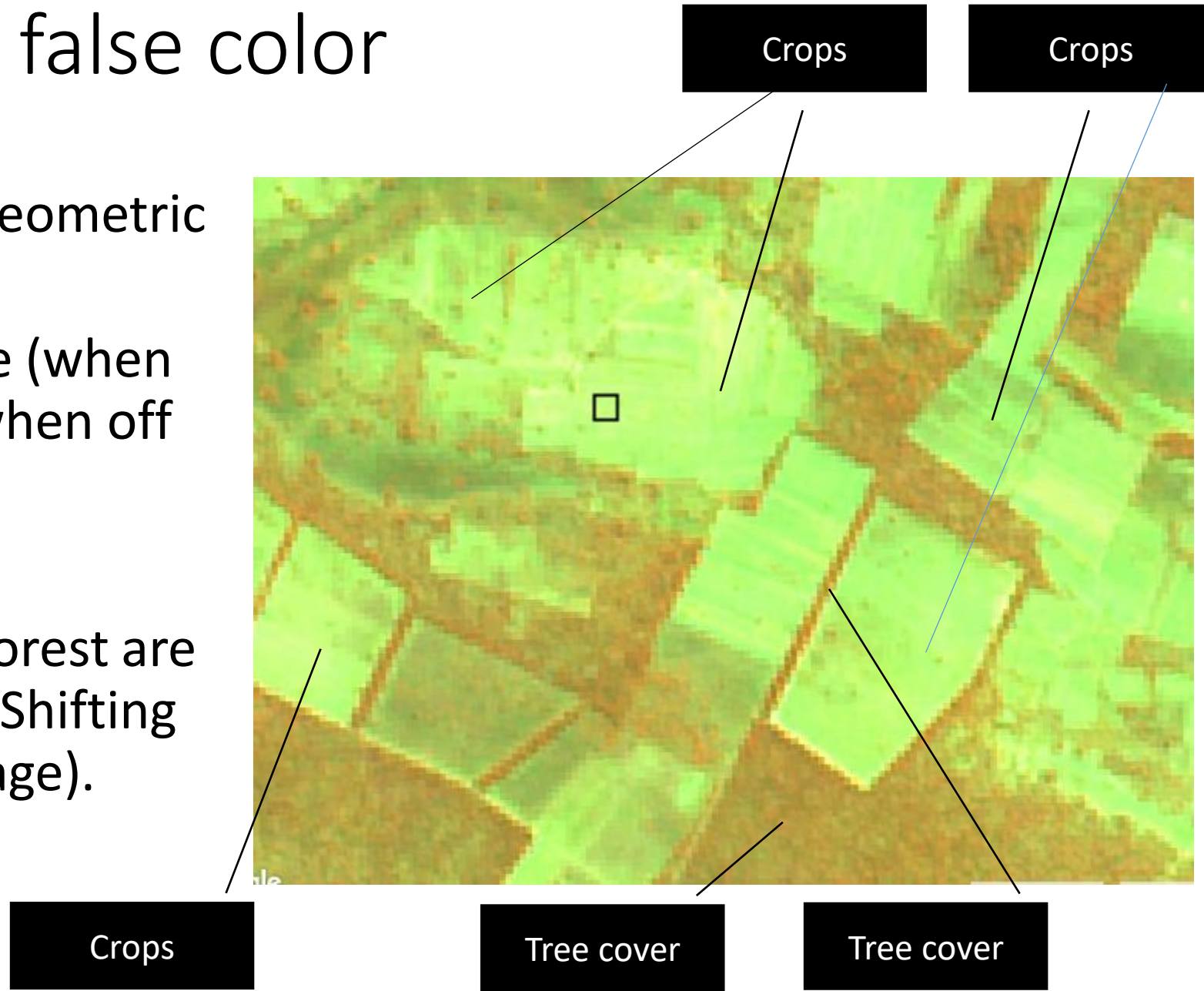
LUC is not always easy to see in the chart. Sometimes the forest seasonality matches also the cropland seasonality (both rely on rainfall).

Landsat 8 Greenest Pixel False color

2015

Use the slider to see the plot in 2000 > 2018

Landsat 7 Greenest Pixel False color

2000

2018: Cropland

2000: Forest
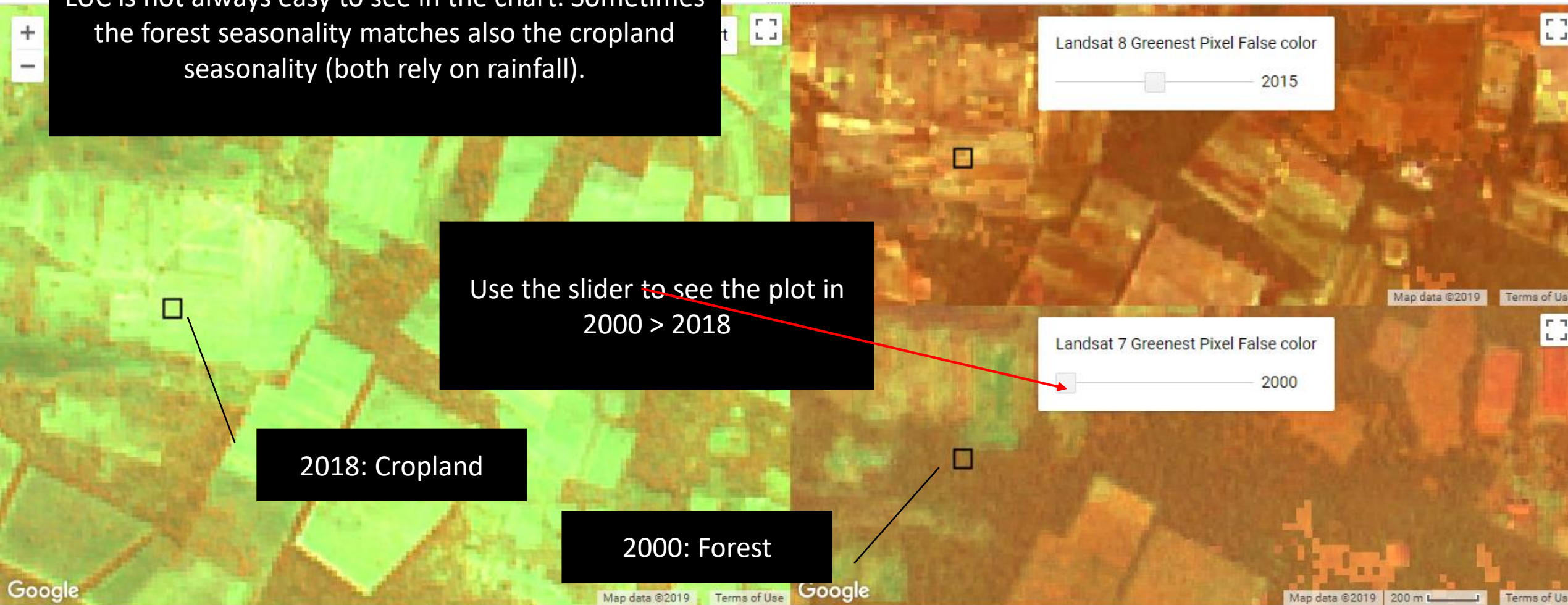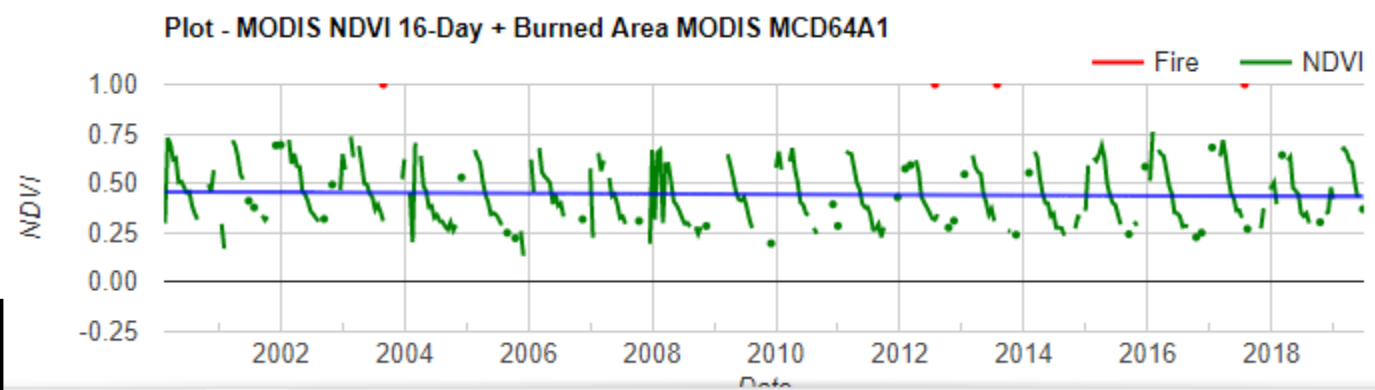
Land Use Change

Before (2000)
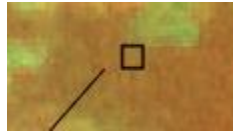
Change happened in 2001.
Forest to Cropland.
You can notice the square of removed vegetation forming around the plot.

Masked pixels. It can complicate the interpretation.

Sentinel 2 : Composite of last 12 months. To select single image click on Sentinel NDVI chart

Landsat 8 Greenest Pixel False color
2015

Landsat 7 Greenest Pixel False color
2001
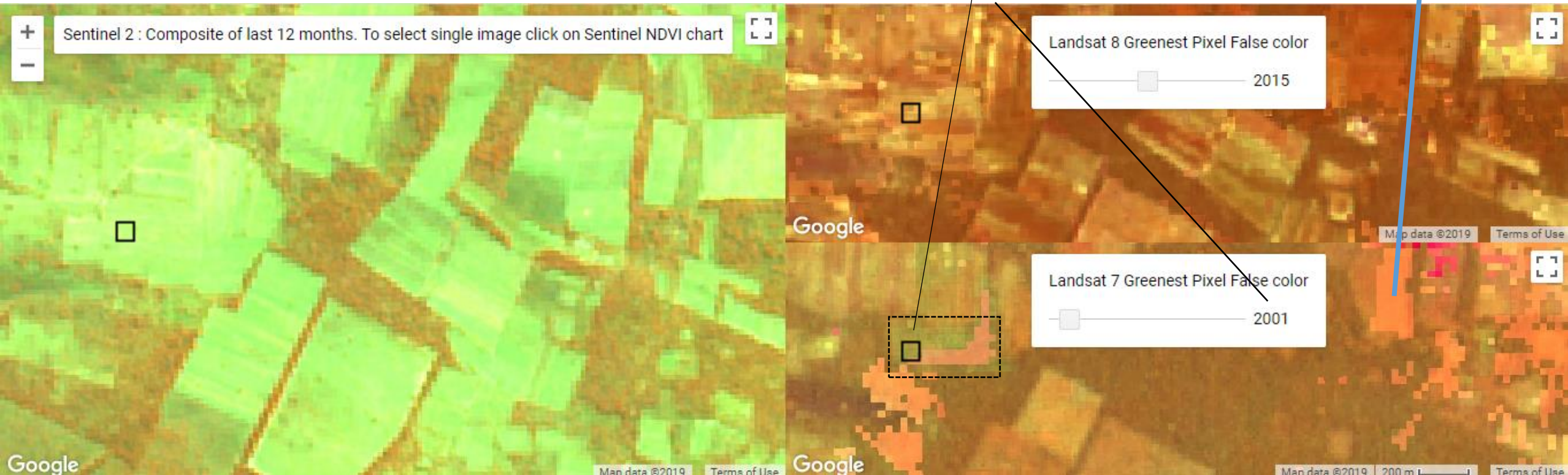
Data resoltuion and availability affects interpreters' agreement
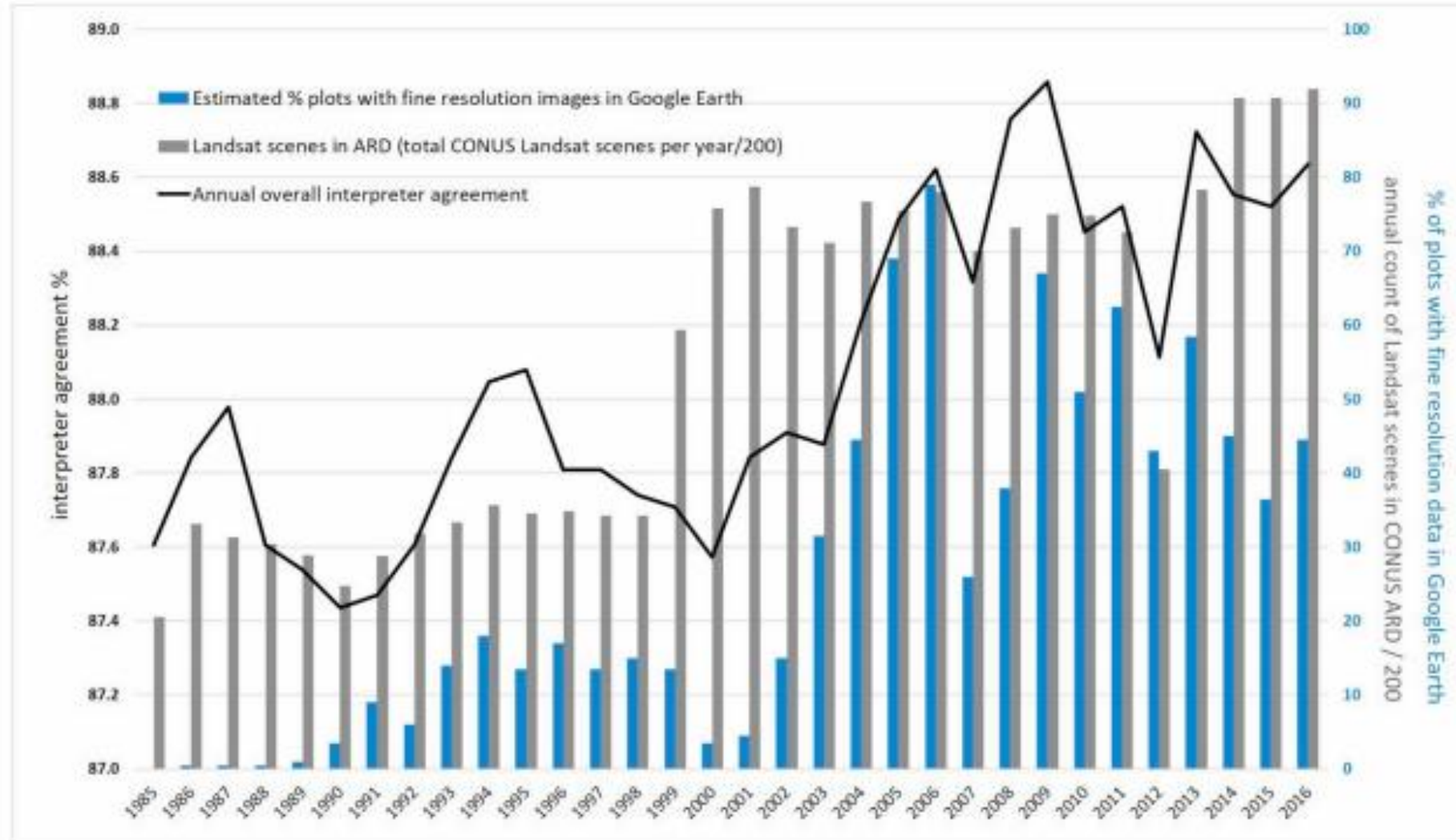
Fig. 5. Overall interpreter agreement (%) and Landsat and Google Earth fine resolution imagery availability 1985–2016.

training and calibration before, during and after data collection

Data collection

- split sample randomly between interpreters
- possible for all interpreters to check all samples (resource intensive)
- avoid mental fatigue, especially for inexperienced interpreters - schedule breaks in data collection

Data collection

- self-checks - recheck own samples

- duplicate samples- multiple interpreters check the same samples
  - depending on # of samples and resources:

- random checks (by supervisor) and review errors with the group

- hot and cold checks

Post data collection

- ancillary data check - use other data sources to flag plots for reassessment

- logical checks - data collected reflects possible conditions

- recheck plots marked with low confidence

options for rechecking: self-check, other interpreter, group, supervisor, field data, external

Reporting consistency

- percent agreement in the duplicate samples

- percent of plots with low confidence

- additional analysis... such as geographical bias, bias

  by data sources
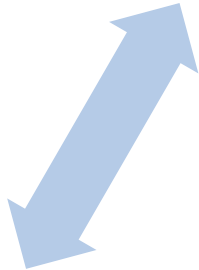
Virtual capacity building on REDD+ FCPF reporting
Week 1: Activity data

Yelena Finegold

FAO Forestry Officer

Virtual capacity building

28 September – 2 October

Monday: Sampling and response design

- Review of concepts
  - Sampling design
  - Response design
- Data requirements
  - Sampling intensification files
  - HIA shapefile
  - Collect earth project
- Review of SOPs 1 and 2

Tuesday: practical session on sampling and response design

- Review questions
- Technical support
- Review comments on SOPs 1 and 2
- Transfer of Planet data
- Optional: creating offline time series clips for visual interpretation from downloaded Planet data
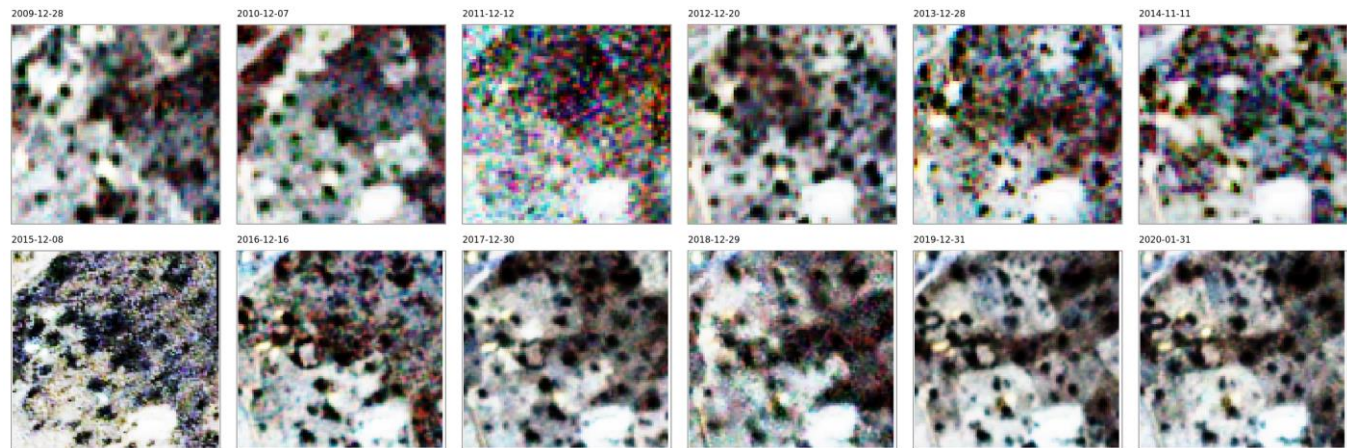
Review shared folder and prerequisites

[Google Drive link: https://docs.google.com/document/d/1DQq5StnIW6dGFniwKeUrLOJ8N95hjytlKjdzRfbFLoY/edit?usp=sharing](https://docs.google.com/document/d/1DQq5StnIW6dGFniwKeUrLOJ8N95hjytlKjdzRfbFLoY/edit?usp=sharing)

Please make sure the following programs are installed on your computer:

- Zoom
- Google Chrome
- Collect Earth
- Collect
- QGIS or ArcGIS
- FileZilla
- Excel or another spreadsheet program

Please make sure you have log in access to the follow platforms:

- Google Earth Engine
- SEPAL

What is a SOP?

- A **standard operating procedure** (**SOP**) is a set of step-by-step instructions compiled to help staff carry out **complex** routine operations.

- SOPs aim to achieve efficiency **quality output, consistency** and adherence to **reporting requirements**.

- SOPs are accompanied by template forms that are filled out for each iteration of activity data.

Reviewing SOPs and accompanying forms

- Main inputs for SOPs:
  - Roles: who should be involved in different steps of the process?
  - Storage of materials: where are the forms, relevant data and documents stored?
  - Are any steps missing?
  - Review and complete forms

Practical exercise
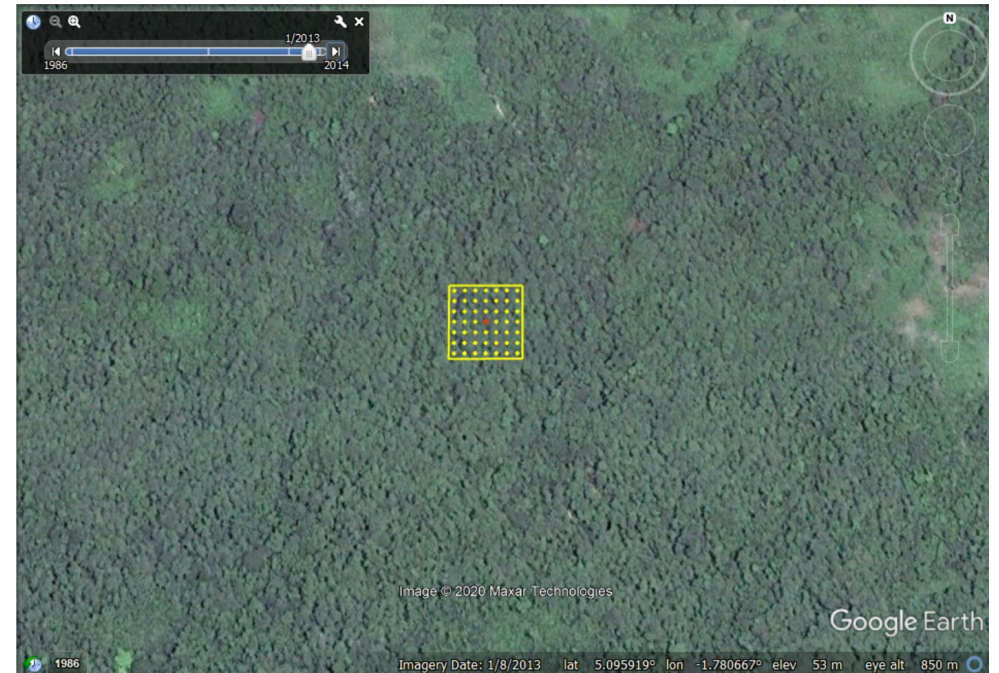
Sample design and selection

- The sample design is determined by the coordinator and taking into account decisions from a facilitated group discussion of experts
- Samples for duplication are determined before interpretation
- The sample selection is completed by *one person* for consistency
- The sample selection can be quality checked by another person(s)
- The selected samples are distributed to interpreters

Sampling design: data requirements

- Activity data/Sample intensification/ Grid_Ghana_1000m_intensification*

- Activity data/Additional data/HIA/merged/HIA_merged*

Practical exercises

- Intensify sampling for 1 HIA to 2 x 2 km
  - Identify HIA information in the grid
  - Subset sampling grid to 1 HIA
  - Remove existing samples
  - Observe the number of rows in the newly subset data
  - Determine the number of plots used in a pilot phase in training
  - Determine the number of duplicate samples for QA/QC
  - Duplicate and change the 'round' ID column value for duplicates
  - Randomize the rows
  - Split into multiple files according to the number of assessors
  - Import the files into Collect Earth
  - Optional: load the files into the grid in Collect

Extracting HIA information to grid for intensification

- Open grid in GIS platform (QGIS or ArcGIS, instructions are in QGIS)
- Open HIA layers in QGIS
- *Optional – merge all HIA layers into 1 shapefile*
  - *In QGIS, to merge, select Vector -> Data management tools -> Merge vector layers*
- Extract polygon information to points
  - In QGIS select Processing -> Toolbox
  - SAGA -> Vector point tools -> Add polygon attributes to points
- *Optional – Extract raster information to points*
  - *In QGIS select Processing -> Toolbox*
  - *SAGA -> Vector <-> raster -> Add raster values to points*
- *Optional – Select the points in the grid using an expression*
  - *In QGIS select Features by expression*
  - *Example expression: "grd_8_G" = 'true'*
- Export data as a CSV

Extracting HIA information to grid for intensification

- Open CSV with extract HIA information in Excel
- Copy the column names from the column names file
- Apply different filters to subset the data
- Questions:
  - How many samples are in the 1 x 1 grid in all of Ghana?
  - How many samples are in the 1 x 1 grid in the GCFRP area?
  - How many samples are in the 1 x 1 grid in the Juaboso Bia HIA?
  - How many samples have already been assessed in the Juaboso Bia HIA?
- Create a new CSV with samples in the 2 x 2 grid that have not been assessed in the Juaboso Bia HIA
  - Fill in the sampling density column with '2'
  - How many samples are in this subset?

Determining number of samples for duplication – QA/QC

- How many samples will be assessed by more than one person?
- Will all interpreters reassess the samples? How many interpreters are there?
- Randomly select 10% of the sample size for reassessment.
    - Add a new column in the CSV called 'random numbers'
    - Add the formula =RAND() in all the cells in this column to generate random numbers
    - Sort the spreadsheet using the random numbers column
    - Calculate how many samples need to be selected.
        - How many samples are selected?
    - Select that number of samples and paste in a new sheet
    - Change the values of the 'round' column to 2
    - Paste these columns back to the original spreadsheet
        - How many samples are in the database?
    - Use the random number column or the ID column to reshuffle the samples so the duplicates are not all next to each other
    - Save as a CSV

Importing samples in Collect Earth

- Open Collect Earth ** update to latest version
- Split the CSV among participants
  - Determine the number of interpreters
  - Go to Tools -> Utilities -> Divide large CSV plot files
  - Split the files by number of interpreters
- Go to Properties -> Sample data -> Browse and select the newly created intensified sample file and Save and apply changes
- Explore all Collect Earth Options

Exploring Collect, a demo

Select a coordinator to upload their samples to the shared folder

Ghana national FRL

Discussion

Tentative agenda

➢ Revision of current available data

➢ What data needs to be collected to scale the current assessment up to the national level?
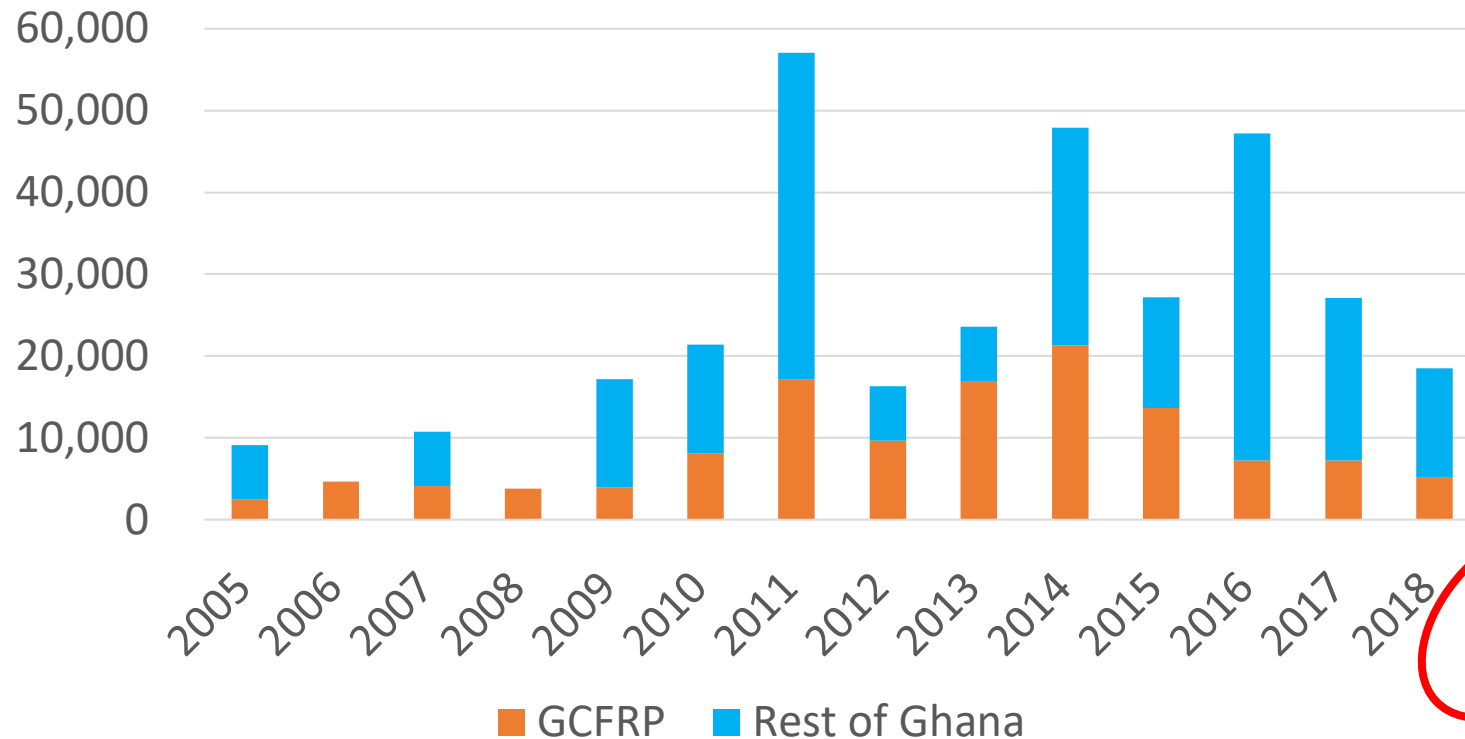
Current assessment

GCFRP landscape

Deforestation (in ha)
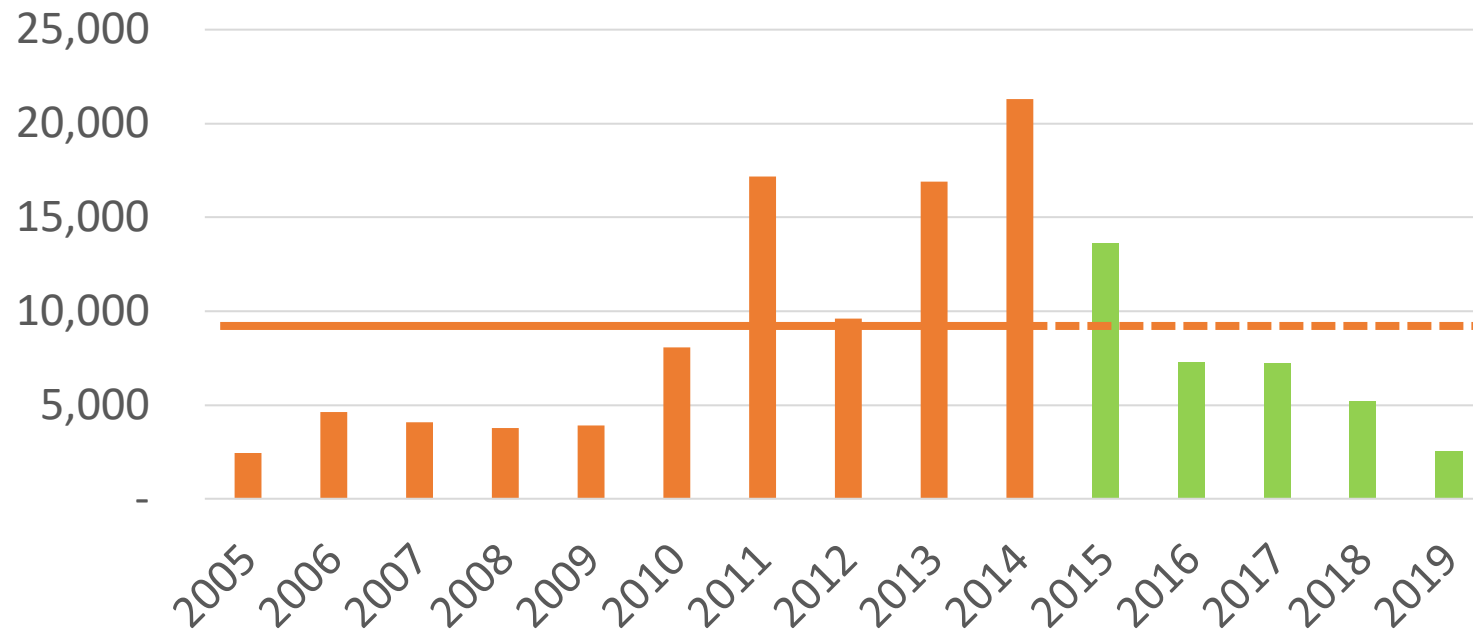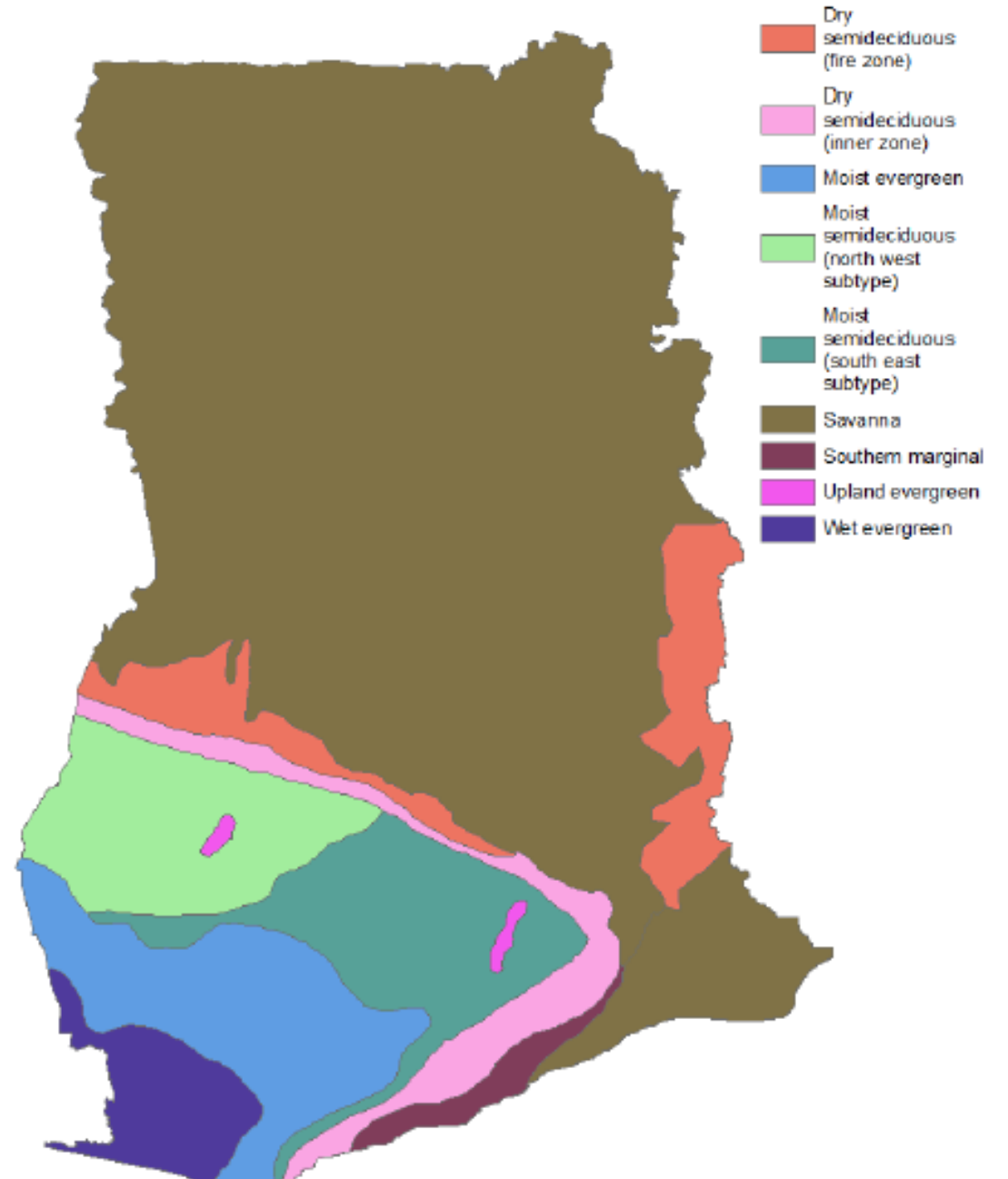
Current assessment

National scale

Deforestation in Ghana (in ha)

Current assessment

Remainder of Ghana

	Total area	Plots	Forest cover	Deforestation rate
Dry semideciduous (inner zone)	920,096	139	29%	0.24%
Dry semideciduous (fire zone)	1,358,993	203	37%	0.13%
Savannah	15,458,513	2,327	20%	0.33%
Southern marginal	263,394	38	21%	1.25%
		2,707		

Dry semideciduous (fire zone)

Dry semideciduous (inner zone)

Moist evergreen

Moist semideciduous (north west subtype)

Moist semideciduous (south east subtype)

Savanna

Southern marginal

Upland evergreen

Wet evergreen

# What data needs to be collected to scale up consistently to the national level?

➤ Calculate EFs for the 4 vegetation zones outside the cocoa landscape

- Use field observations + gap filling 2012 map for the classification of the plot measurements (open/closed forest; annual/perennial cropland, grassland, settlement)
- Calculate average carbon contents for above mentioned classes

➤ Re-assess stable forest plots   549 plots

- Assess degradation year
- Assess deforestation (2019 + use Planet data if available?)
- Assess post-deforestation land-use
- Assess pre- and post-degradation canopy cover & pre-deforestation canopy cover

➤ Re-assess deforestation plots   31 plots

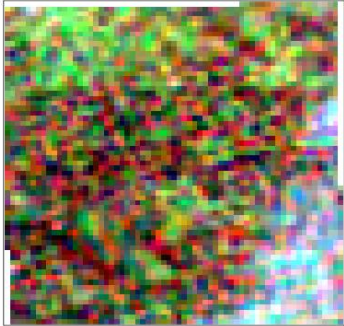- Assess pre-deforestation canopy cover (+ use Planet data if available?)

Additional decisions to be made

➢ Objective?

➢ Reference period?

➢ Sampling design?

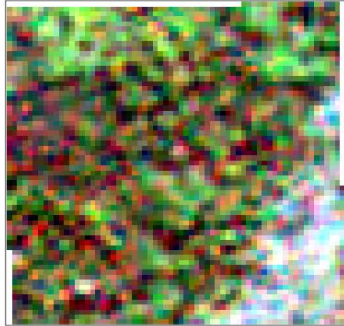➢ Response design?

➢ Same definitions?

Deforestation in Ghana (in ha)
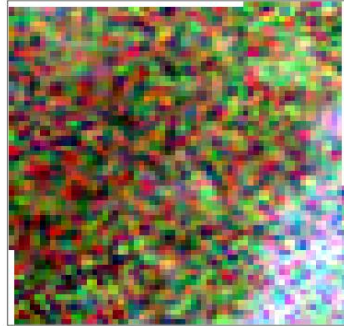
■ GCFRP   ■ Rest of Ghana
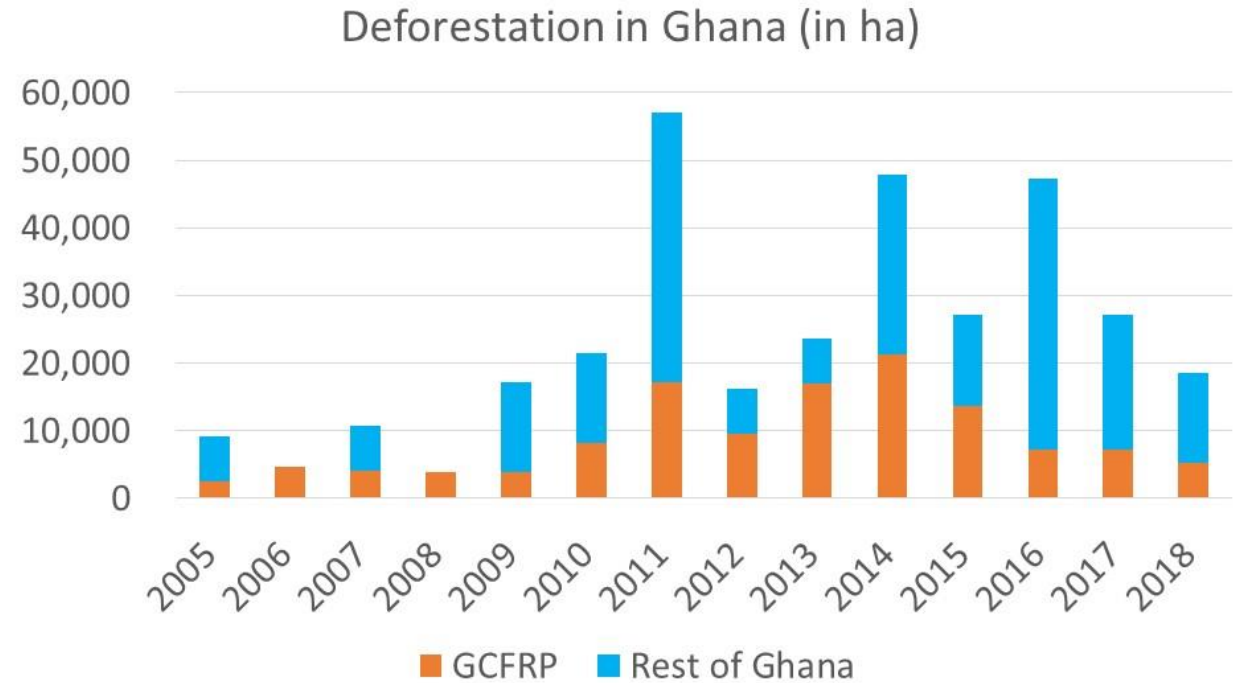
Downloaded Planet data - offline
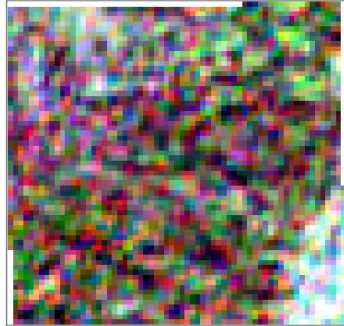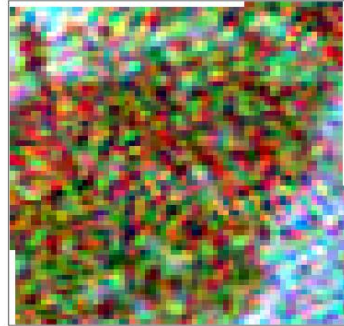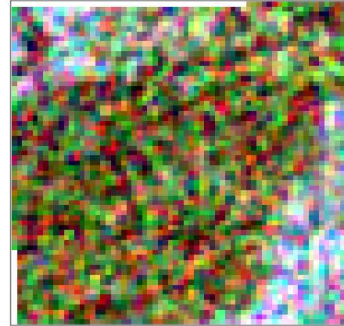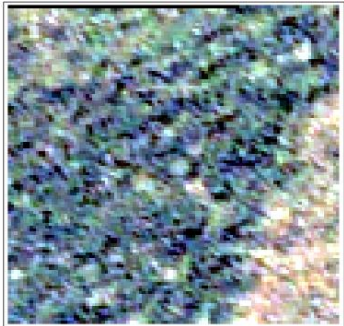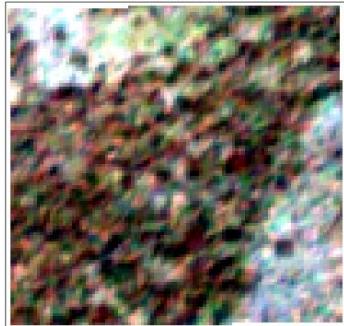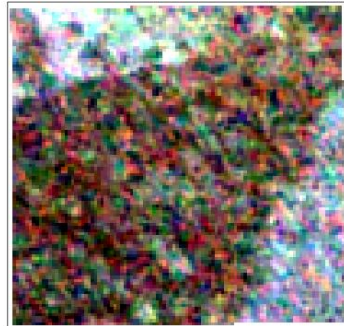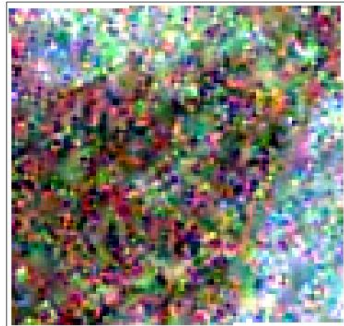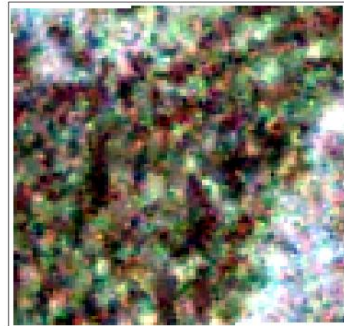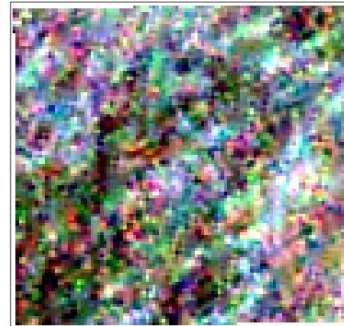
sample_8632_21136

Week 2 Day 1:
How to calculate AD from an intensified systematic sample

Marieke Sandker

FAO Forestry Officer

Virtual capacity building

5 – 9 October

Calculation chain

Sampling design

How to create your sample (size, plot location)?

Response design

What information to collect? How to interpret satellite imagery?

Data management

How to get all data from different interpreters into a single database

Calculate emissions (AD x EF) and propagate uncertainties

Review all excel calculations

Calculate emission factors and their uncertainties

Review all excel calculations

Calculate deforestation estimates and uncertainty

How to extract from all those observations the relevant averages?

Tentative agenda

- Monday: how to calculate deforestation/degradation from the data collected

- Tuesday: combine AD x EF

- Wednesday: calculate associated uncertainty

- Thursday: Final outstanding issues: QA data analysis; SOP follow-up? Testing mapping approaches for HIAs?

- Friday: discuss national assessment (sampling design, response design)

Calculating area statistics from a systematic sample
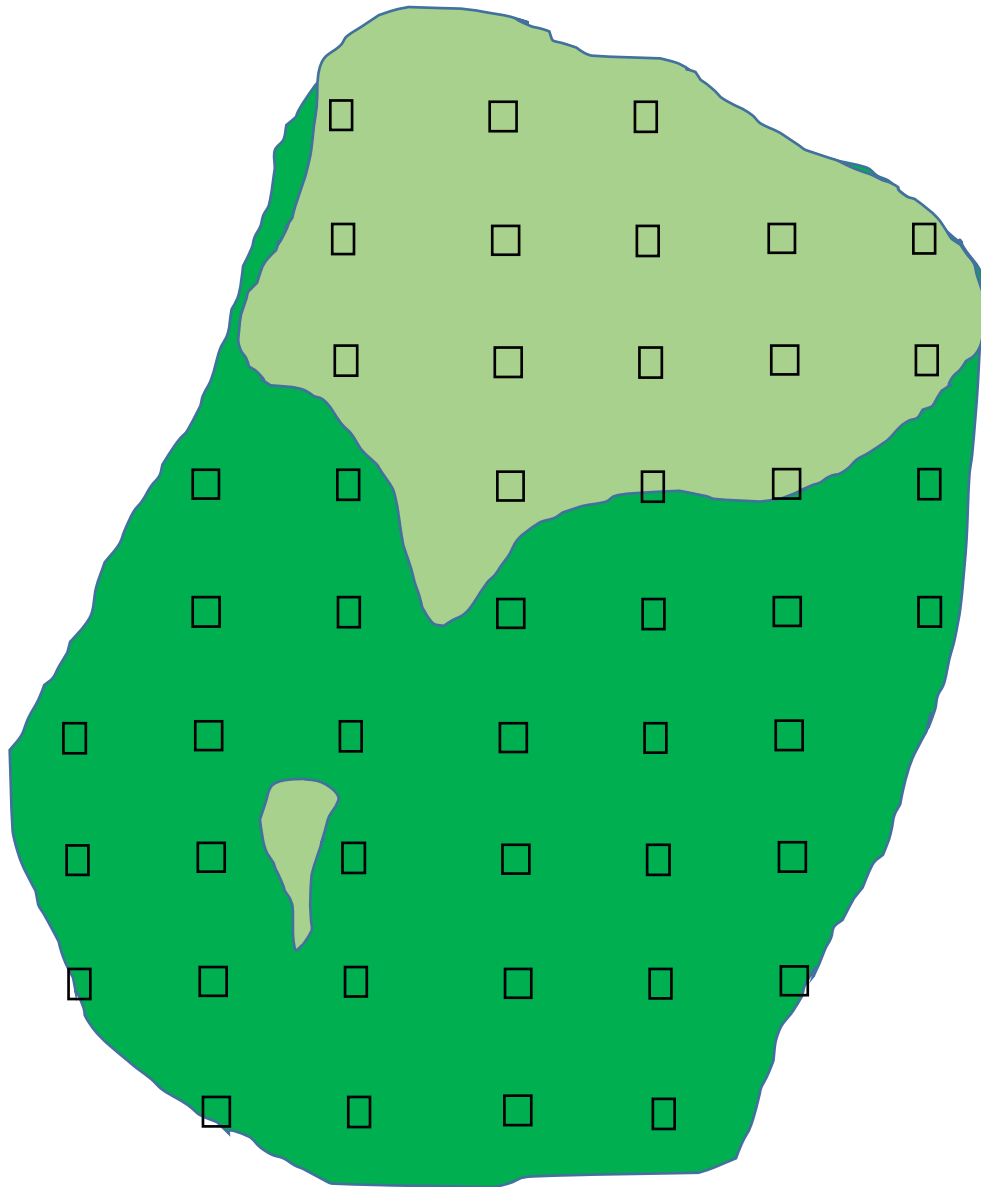
Let's use a simple theoretical example

➢ Total area of interest: 5,914,425 ha

➢ Total sample size: 47 plots

Now I want to calculate the area of forest from my sample

Each sample unit represents an area of: 5,914,425 ha / 47 plots= 125,839 ha/plot

31 of 47 sample plots are forest

Forest area = 31 plots x 125,839 ha/plot = 3,901,004 ha

This can also be calculated as: 31/47 x 5,914,425 ha = 3,901,004 ha

What is the associated formula?

The area of variable v (A_v):

$$A_v = \frac{n_v}{n} \times A$$

where

- n_v is the number of sample plots of variable v,

- n is the total number of sample plots,

- A is the area of the landscape.

Calculating area statistics from an intensified systematic sample
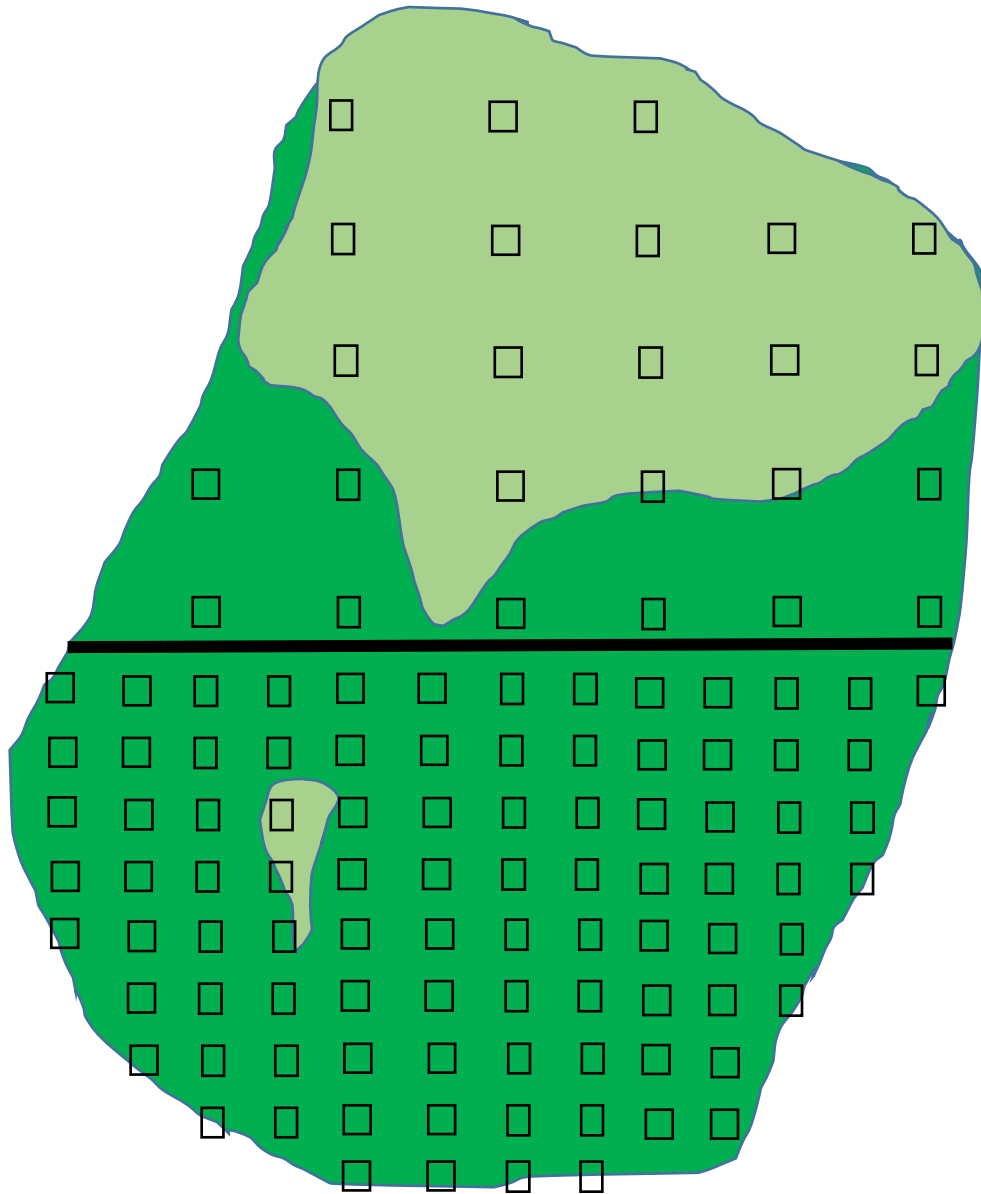
- Total area of interest: 5,914,425 ha
- Total sample size = 116

- Area in stratum 1 = 2,957,000 ha
- Area in stratum 2 = 2,957,425 ha
- Sample size stratum 1 = 25
- Sample size stratum 2 = 91

Now I want to calculate the area of forest again:

Stratum 1: 9/25 x 2,957,000 ha = 1,064,520 ha

Stratum 2: 90/91 x 2,957,425 ha = 2,924,926 ha

Forest area = 1,064,520 ha + 2,924,926 ha = 3,989,446 ha

What is the associated formula?

9/25 x 2,957,000 ha + 90/91 x 2,957,425 ha

The area of variable v:

$$A_v = \sum_{i=1,2} \frac{n_{v,i}}{n_i} \times A_i$$

where

- $n_{v,i}$ is the number of sample plots of variable v falling in stratum i,

- n_i is the number of sample plots falling in stratum i,

- A_i is the area of stratum i.

Simulations with Hansen data to test ideal intensity

Reference period:

Result of 10 iterations 2005-2014

Ha of tree cover loss

Sampling intensity: distance between samples (km)
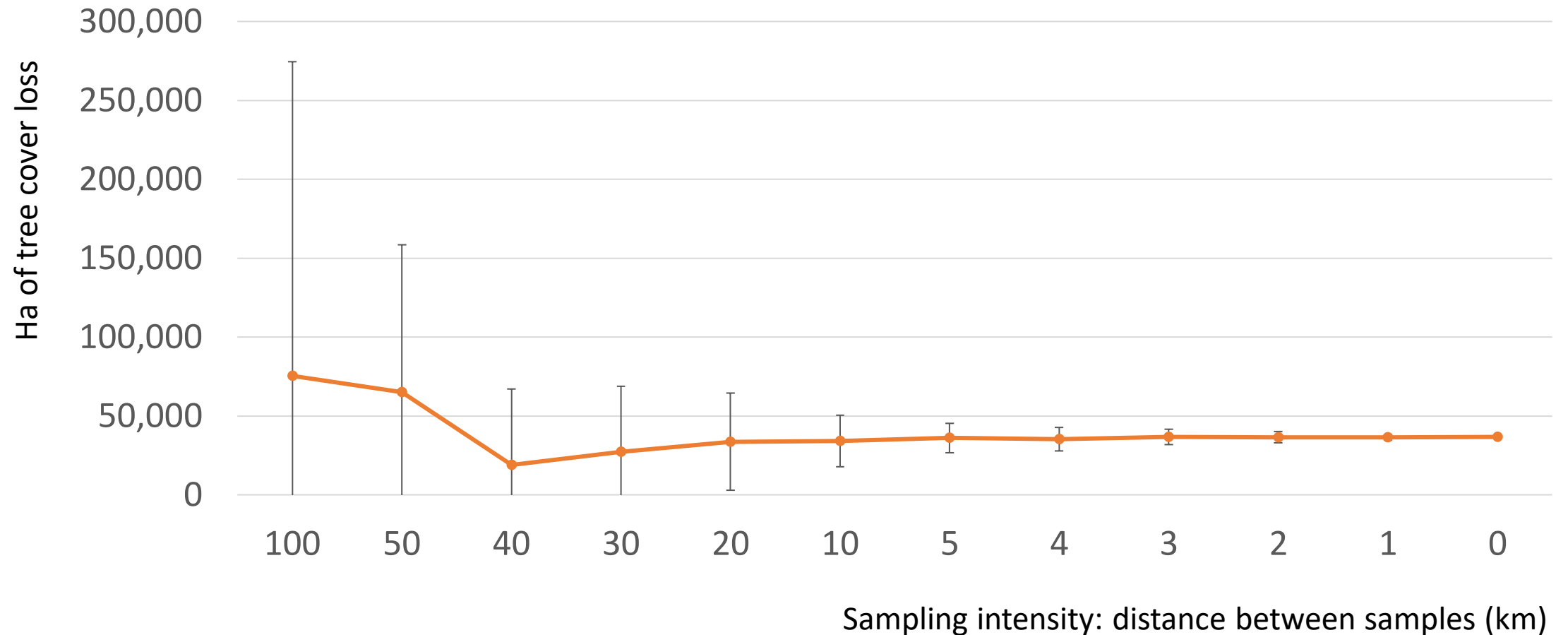
What do our strata look like?

Moist semideciduous (north west subtype)

Upland evergreen

Moist semideciduous (south east subtype)

Upland evergreen

Moist evergreen
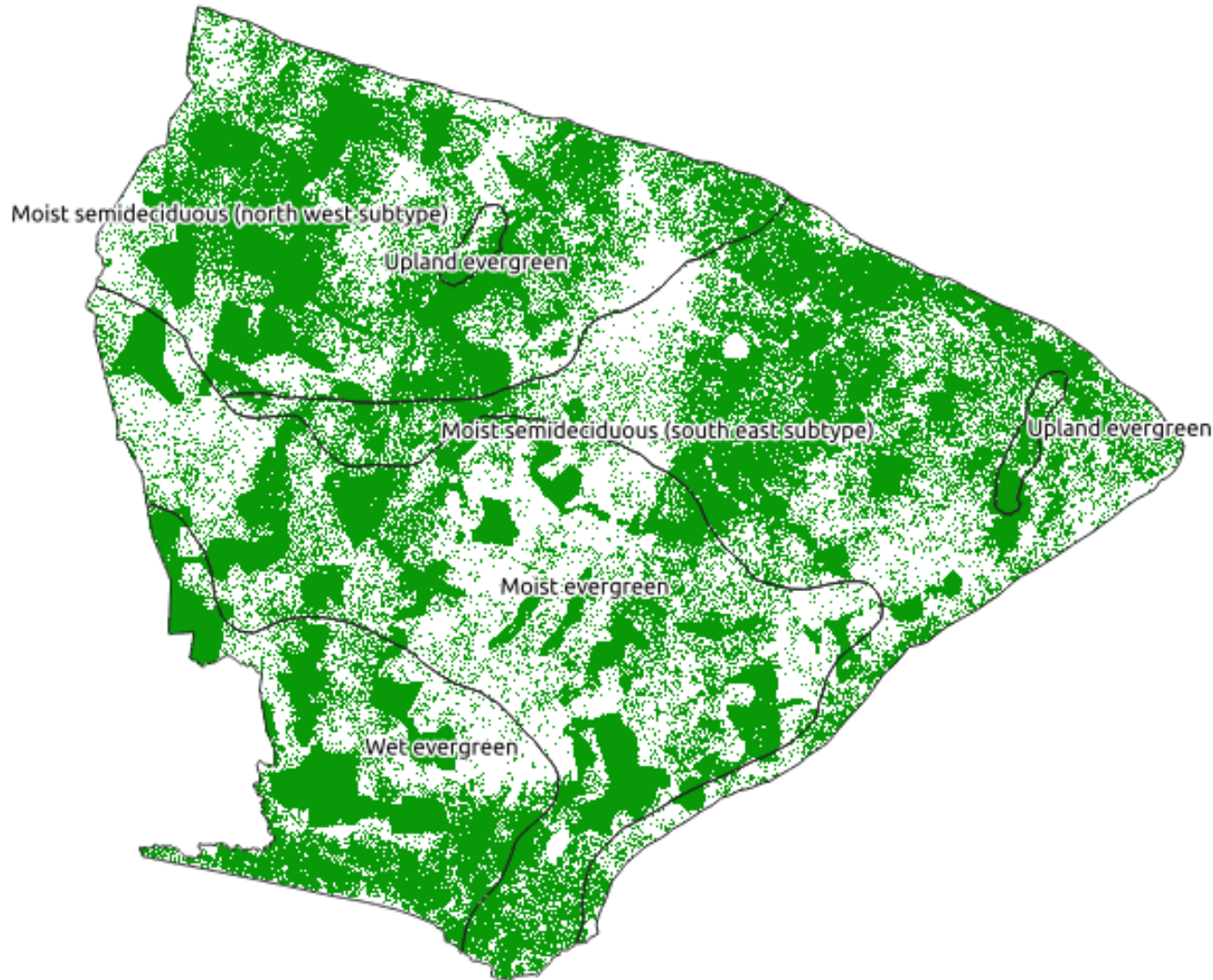
Wet evergreen

➢ We use 3 sampling strata:
  Outside forest mask (4 x 4 km grid)
  Inside forest mask (2 x 2 km grid)
  Upland evergreen (1 x 1 km grid)

➢ And 4 post-strata:
  Moist evergreen
  Moist Semi deciduous NW
  Moist Semi deciduous SE
  Wet evergreen

What are our strata?

Moist evergreen 2x2

Moist evergreen 4x4

Moist SemiD NW 2x2

Moist SemiD NW 4x4

Moist SemiD SE 2x2

Moist SemiD SE 4x4

Wet evergreen 2x2

Wet evergreen 4x4

Upland evergreen

Now let's calculate the basic numbers needed in excel

2 sample plots on the 8x8 grid that fall in vegetation zones

vegetation zone map

Forest mask
(LULC maps)

Tentative agenda

- Monday: how to calculate deforestation/degradation from the data collected

- Tuesday: calculate associated uncertainty

- Wednesday: combine AD x EF

- Thursday: Final outstanding issues: QA data analysis; SOP follow-up? Testing mapping approaches for HIAs?

- Friday: discuss national assessment (sampling design, response design)

What is a confidence interval?
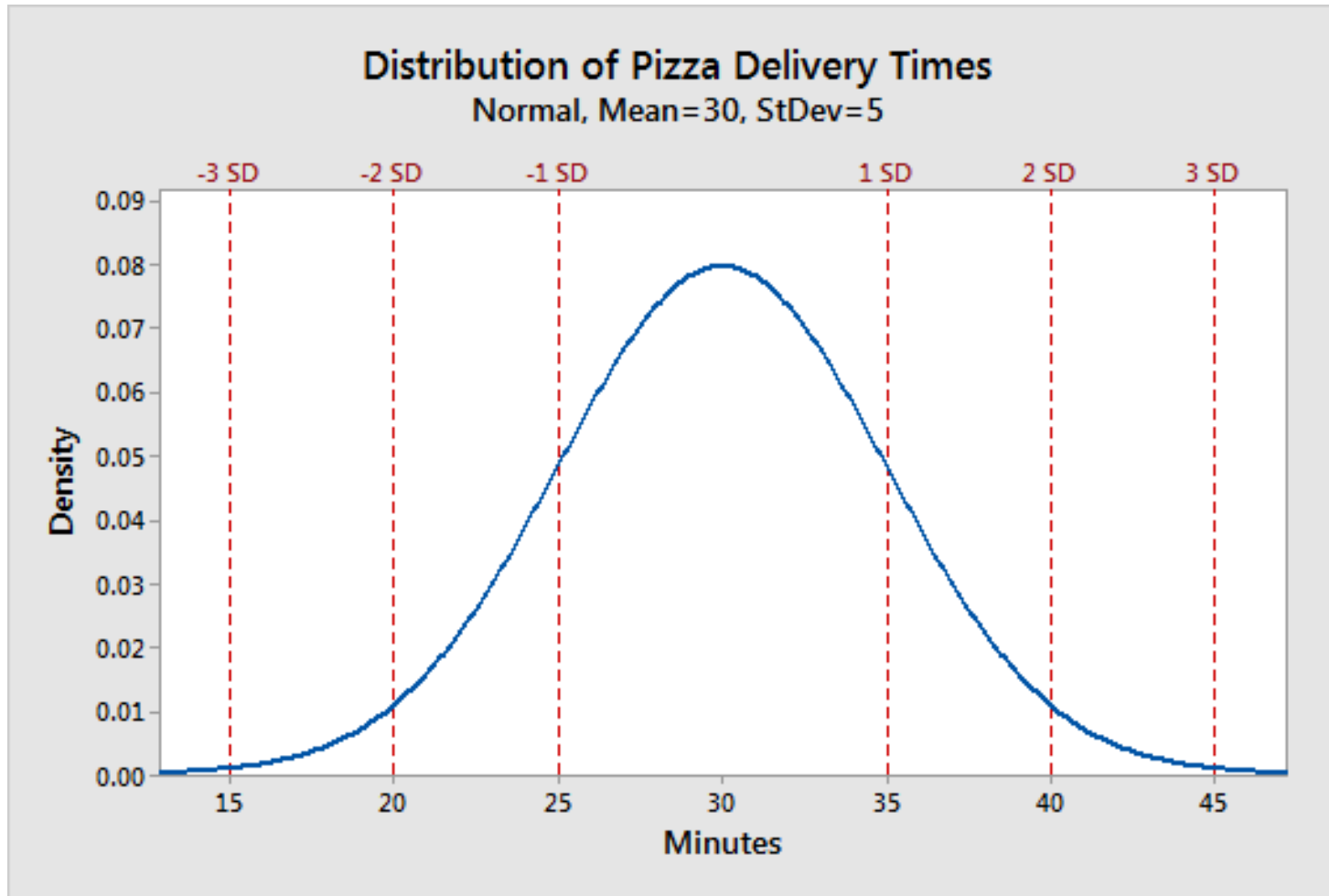
Distribution of Pizza Delivery Times
Normal, Mean=30, StDev=5

I am 95% certain that my pizza will be delivered between 20-40 minutes, or my estimated delivery time = 30 ± 10 min

Confidence interval around Area of variable v

$$CI\ (\pm)\ of\ A_v\ =\ 1.64\ \times\ \sqrt{\frac{p_v \times (1 - p_v)}{(n\ \ - 1)}}\ \times\ A$$

where

- p_v is the estimated probability of variable v, calculated as nv/n,

- n is the total number of sample plots,

- A is the area of the landscape.

Calculating area statistics from a systematic sample

Let's use a simple theoretical example
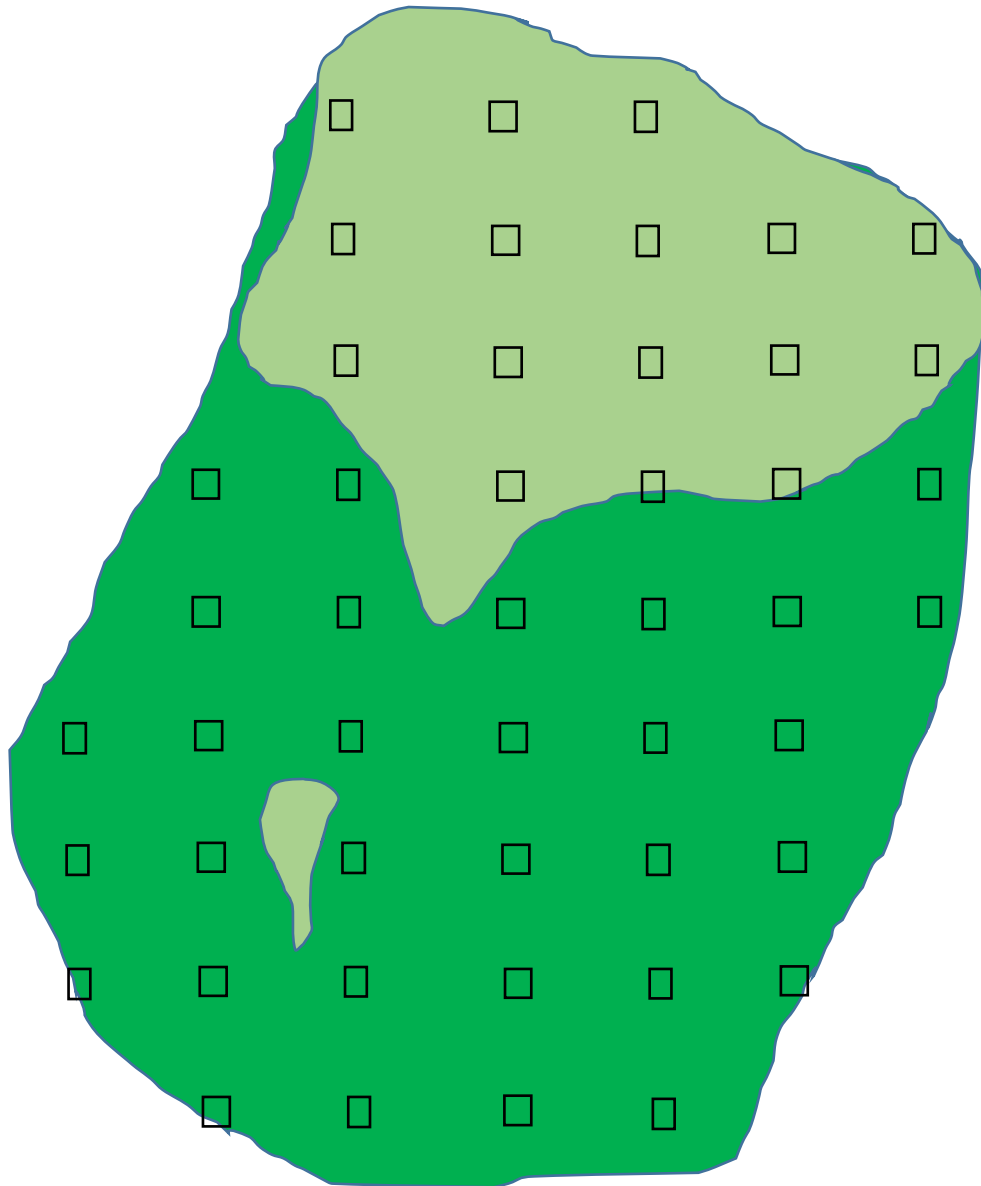
➤ Total area of interest: 5,914,425 ha

➤ Total sample size: 47 plots
➤ 31 forest plots

Calculating area statistics from an intensified systematic sample
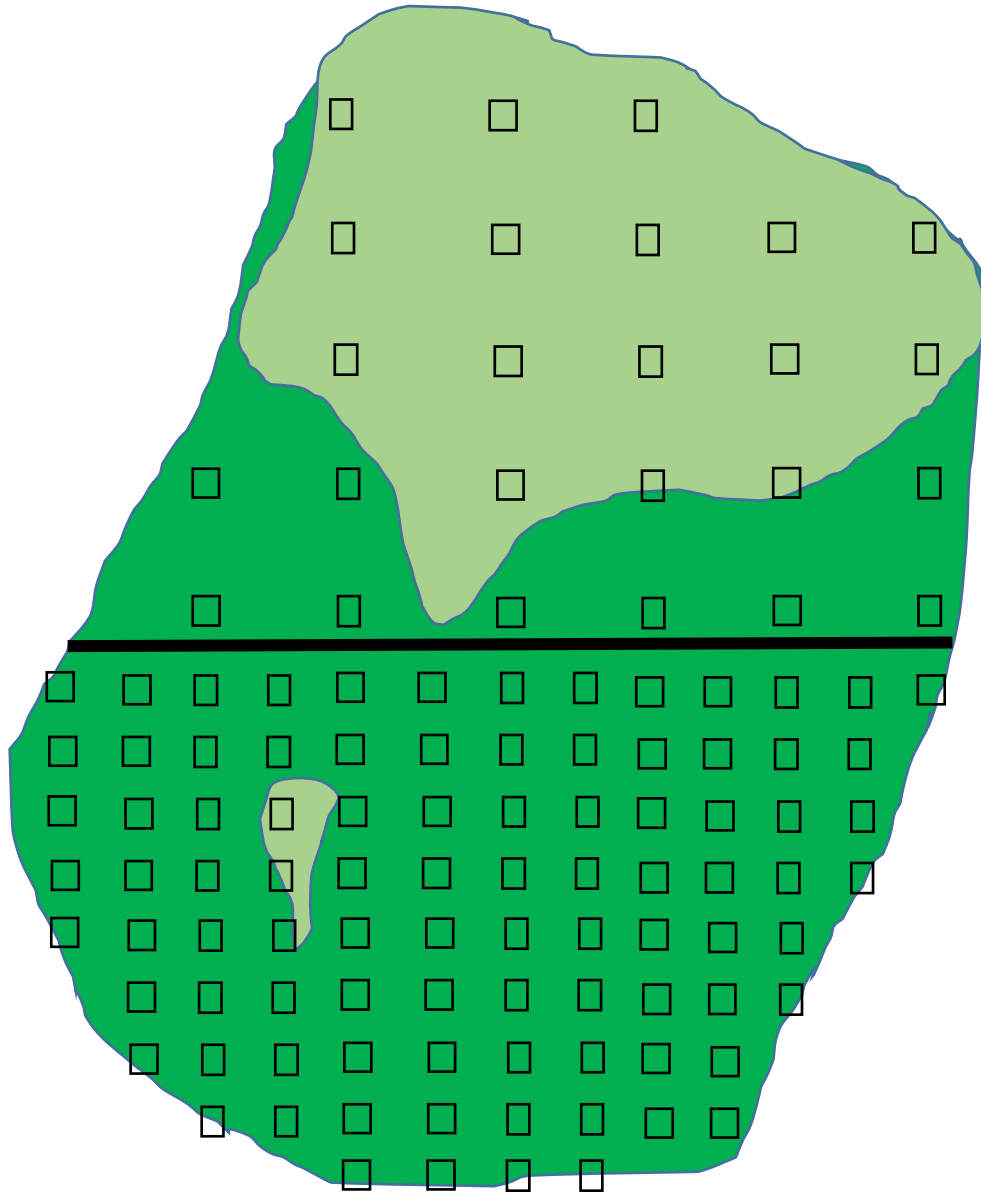
➢ Total area of interest: 5,914,425 ha
➢ Total sample size = 116

➢ Area in stratum 1 = 2,957,000 ha
➢ Area in stratum 2 = 2,957,425 ha
➢ Sample size stratum 1 = 25 ; 9 forest
➢ Sample size stratum 2 = 91 ; 1 forest

$$CI\ (\pm)\ of\ A_v\ =\ 1.64\ \times\ \sqrt{\frac{p_v \times (1 - p_v)}{(n\ -1)}}\ \times\ A$$

$$CI\ (\pm)\ of\ A_v\ =\ 1.64\ \times\ \sqrt{\frac{p_v \times (1 - p_v)}{(n\ -1)} \times A^2}$$

Variance

$$V(A_{v,e}) = \sum_i \frac{p_{v,e,i} \times (1 - p_{v,e,i})}{(n_{e,i} - 1)} \times A_{e,i}^2$$

where

- $P_{v,e,i}$ is the estimated probability of variable v, in vegetation zone e and stratum (grid density) i, calculated as $n_{v,e,i}/n_{e,i}$,

- $n_{e,i}$ is the total number of sample plots in vegetation zone e and stratum i,

- $A_{e,i}$ is the area in vegetation zone e and stratum i.

Ghana's deforestation calculation

➢ Deforestation by vegetation type & open/closed forest

➢ Initial calculation:

The area of variable v in vegetation zone v forest type c:

$$A_{v,e,c} = \sum_{i=1,2} \frac{n_{v,e,i}}{n_{e,i}} \times A_{e,i} \times ratio\ open/closed$$

where

- $n_{v,e,i}$ is the number of sample plots of variable v, in vegetation zone e, falling in stratum i,

- $n_{e,i}$ is the number of sample plots in vegetation zone e, falling in stratum i,

- $A_{e,i}$ is the area of vegetation zone e stratum i.

Ghana's deforestation calculation

The area of variable v:

$$A_{v,e,c} = \sum_{i=1,2} \frac{n_{v,e,i}}{n_{e,i}} \times A_{e,i} \times ratio \frac{open}{closed}^{e,\,i}$$

$$A_{v,e,c} = \sum_{i=1,2} \frac{\cancel{n_{v,e,i}}}{n_{e,i}} \times A_{e,i} \times \frac{n_{v,e,i,c}}{\cancel{n_{v,e,i}}}$$

$$A_{v,e,c} = \sum_{i=1,2} \frac{n_{v,e,i,c}}{n_{e,i}} \times A_{e,i}$$

Ghana's deforestation calculation

Numerical example:

➢ in 2005-2014 we have 7 plots of deforestation that fall on moist evergreen 2x2km

➢ Expansion factor (total area of the stratum/total plots in the stratum)
   = 641 ha (886,983/1,384)

➢ 2 of these fall in open forest, 5 in closed

➢ initial calculation: 7 x 641 = 4,486 ha deforestation (in moist ever 2x2)

4,486 ha x 2/7 = 1,282 ha deforestation in open forest (in moist ever 2x2)

➢ new calculation: 2 x 641 = 1,282 ha deforestation in open forest (in moist ever 2x2)

Week 2 Day 3:
AD x EF
Calculating emissions

Marieke Sandker

FAO Forestry Officer

Virtual capacity building

5 – 9 October

Tentative agenda

- Monday: how to calculate deforestation/degradation from the data collected

- Tuesday: calculate associated uncertainty

- **Wednesday: combine AD x EF**

- Thursday: Final outstanding issues: QA data analysis; SOP follow-up? Testing mapping approaches for HIAs?

- Friday: discuss national assessment (sampling design, response design)

Recap – how to calculate AD

*The area of variable v (A_v) in a **simple** systematic sample:*

$$A_v = \frac{n_v}{n} \times A$$

where

- n_v is the number of sample plots of variable v,

- n is the total number of sample plots,

- A is the area of the landscape.

Recap – how to calculate AD

*The area of variable v (A_v) in an **intensified** systematic sample:*

$$A_v = \sum_{i=1,2} \frac{n_{v,i}}{n_i} \times A_i$$

where

- $n_{v,i}$ is the number of sample plots of variable v falling in stratum i,

- n_i is the number of sample plots falling in stratum i,

- A_i is the area of stratum i.

Recap – Ghana's GCFRP AD

Ghana has assessed deforestation:

➢ Per vegetation zone

e:	i:
Moist evergreen	2x2
Moist evergreen	4x4
Moist SemiD NW	2x2
Moist SemiD NW	4x4
Moist SemiD SE	2x2
Moist SemiD SE	4x4
Wet evergreen	2x2
Wet evergreen	4x4
Upland evergreen	1x1

Area of variable v in vegetation zone e ($A_{v,e}$)

$$A_{v,e} = \sum_{i=1,2} \frac{n_{v,e,i}}{n_{e,i}} \times A_{e,i}$$

where

- $n_{v,e,i}$ is the number of sample plots of variable v, in vegetation zone e, falling in stratum i,

- $n_{e,i}$ is the number of sample plots in vegetation zone e, falling in stratum i,

- A_i is the area of vegetation zone e, stratum i.

Recap – Ghana's GCFRP AD

Ghana has assessed deforestation:

➢ Per vegetation zone

➢ For open and closed forest

We disaggregate AD to this level because the carbon contents differs per vegetation zone and canopy cover

Area of variable v in vegetation zone e for canopy cover c ($A_{v,e,c}$)

$$A_{v,e,c} = \sum_{i=1,2} \frac{n_{v,e,c,i}}{n_{e,i}} \times A_{e,i}$$

where

- $n_{v,e,i}$ is the number of sample plots of variable v, in vegetation zone e, with canopy cover c, falling in stratum i,

- $n_{e,i}$ is the number of sample plots in vegetation zone e, falling in stratum i,

- A_i is the area of vegetation zone e, stratum i.

Recap – what is a confidence interval
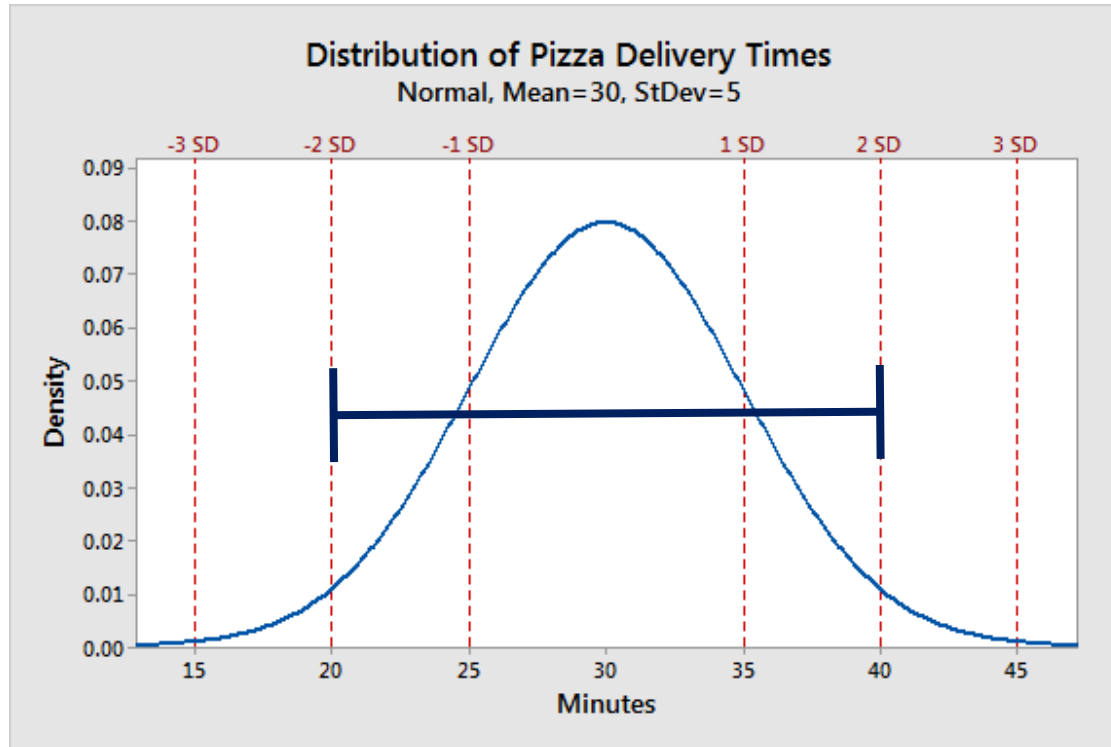
I am 95% certain that my pizza will be delivered between 20-40 minutes, or my estimated delivery time = 30 ± 10 min

Recap – how to calculate CI

*The CI around the area of variable v (A_v) in a **simple** systematic sample:*

$$CI\ (\pm)\ of\ A_v\ =\ 1.64\ \times\ \sqrt{\frac{p_v \times (1 - p_v)}{(n\ -\ 1)} \times A^2} \qquad = \qquad CI\ (\pm)\ of\ A_v\ =\ 1.64\ \times\ \sqrt{V}_{Av}$$

where

V_{Av} is the variance of A_v

- p_v is the estimated probability of variable v, calculated as nv/n,

- n is the total number of sample plots,

- A is the area of the landscape.

Recap – how to calculate CI

*The CI around the area of variable v (A$_v$) in an **intensified** systematic sample:*

$$CI\ (\pm)\ of\ A_v\ =\ 1.64\ \times\ \sqrt{V}_{Av}$$

$$V(A_v) = \sum_i \frac{p_{v,i} \times (1 - p_{v,i})}{(n_i - 1)} \times A_i^2$$

where

- P$_{v,i}$ is the estimated probability of variable v, in stratum i, calculated as n$_{v,i}$/n$_i$,

- n$_i$ is the total number of sample plots in stratum i,

- A$_i$ is the area in stratum i.

Recap – how to propagate uncertainty

For an addition (or subtraction):

$$CI_{abs} (\pm) \text{ of Av} = \sqrt{CI_{abs\ 2x2}^2 + CI_{abs\ 4x4}^2} \qquad in\ ha$$

For today we also will use:

These are equations 3.1 and 3.2 in IPCC

For a multiplication (e.g. AD x EF):

$$CI_{rel} (\pm) \text{ of } Emissions = \sqrt{CI_{rel\ AD}^2 + CI_{rel\ EF}^2} \qquad in\ \%$$

Today: AD x EF

The AD spreadsheet we created gives us the following outputs:

Open forest - annual deforestation (ha/year)

	Wet Evergreen	Moist Evergreen	Moist Semideciduous SE	Moist Semideciduous NW	Upland Evergreen
2005-2014	182	768	1,950	1,840	15.97
2015	607	1,282	1,860	3,070	160
2016	1,217	641	1,860	1,832	-
2017	1,217	-	1,283	1,238	160
2018	-	1,279	641	619	-
2019	-	641	1,283	-	-

Closed forest - annual deforestation (ha/year)

	Wet Evergreen	Moist Evergreen	Moist Semideciduous SE	Moist Semideciduous NW	Upland Evergreen
2005-2014	304	1,728	1,078	1,171	160
2015	607	3,202	641	1,857	319
2016	-	-	-	1,238	479
2017	-	1,282	641	1,238	160
2018	-	-	1,283	1,213	160
2019	-	-	-	619	-

EF calculation

We need EFs for the following classes:

	Wet Evergreen	Moist Evergreen	Moist Semideciduous SE	Moist Semideciduous NW	Upland Evergreen
Open forest					
Closed forest					

EFs are calculated as follows for deforestation:

➢ Carbon contents forest – carbon contents replacing land-use

➢ Carbon contents forest = AGC + BGC + DW + L + delta SOC

EFs are calculated as follows for deforestation:

➢ (AGC+BGC+DW forest) x average relative canopy cover reduction

Emissions

Per vegetation zone, per open/closed forest, per year:

AD x EF

Calculation chain

Sampling design

How to create your sample (size, plot location)?

Response design

What information to collect? How to interpret satellite imagery?

Data management

How to get all data from different interpreters into a single database

Week 1

Calculate emissions (AD x EF) and propagate uncertainties

Review all excel calculations

Calculate emission factors and their uncertainties

Review all excel calculations

Calculate deforestation estimates and uncertainty

How to extract from all those observations the relevant averages?

Week 2

Calculation chain

Next steps

- Calculation of AD and EF
- Discussion of sampling design for planned intensification (national scale)
- SOP follow up
- Planet data transfer
- Testing mapping approaches for HIAs