Creating NFI Phase 2 sample location with SFC

2024-02-26

# Context

This technical report covers the establishment of field sample locations for the Improved “National Forest Inventory Plan for Indonesia Forest”. The main document proposes a two-phases systematic stratified design with target sampling sizes for seven strata of dryland forest and one strata of mangrove forest.

Three annexes to the main report describe: 1. Optimal sampling grids for the NFI phase 1 visual interpretation. 1. Results of the NFI Phase 1 visual interpretation. 1. Distribution of the NFI Phase 2 field sample location (this report).

A separate annex is dedicated to Mangrove, Mangrove sampling is not addressed in this report.

# Method

Given a final dataset of NFI phase 1 visually interpreted samples, the general approach to selecting the NFI phase 2 samples was as follows:

1. Determine the distance of each sample along a space filling curve (SFC), following Hilbert’s distance and computed with Python dask-geopandas [library](https://dask-geopandas.readthedocs.io/en/stable/docs/reference/api/dask_geopandas.GeoSeries.hilbert_distance.html)).
2. Select samples along the curve based on a sampling proportion calculated with the number of plots in the pool (i.e. the NFI phase 1 plots that are accessible forest), the number of target plots to sample (i.e. the number of target plots for the Phase 2) and adjustment factors to take into consideration uncertainty around the identification of forest / non-forest in the Phase 1.

in R: sampling\_proportion = ceiling(n\_plot \* (1 + U) \* sfc\_adj / n\_pool)

With a manually adjusted variable to reach the target number of samples in accessible forest and the uncertainty around forest/non-forest identification in the Phase 1.

1. A random starting point is assigned then the selection is systematic following the space filling curve.
2. The adjustment factors are tailored manually to provide a number of forest plots slightly exceeding the target.
3. All natural forest samples are kept, 70 samples are randomly selected among the accessible Trees Outside Forest samples.
4. Separately, additional samples in primary forest are provided for an intensification study on forest types that cover an area too small to be effectively represented in a national level sampling design. These samples can be measured according to the NFI field manual and integrated in the NFI database management system. However they should be filtered out from the national level statistics calculation chains and separate calculation chains should be designed specifically for these samples.

## R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

summary(cars)

## speed dist   
## Min. : 4.0 Min. : 2.00   
## 1st Qu.:12.0 1st Qu.: 26.00   
## Median :15.0 Median : 36.00   
## Mean :15.4 Mean : 42.98   
## 3rd Qu.:19.0 3rd Qu.: 56.00   
## Max. :25.0 Max. :120.00

## Including Plots

You can also embed plots, for example:



Note that the echo = FALSE parameter was added to the code chunk to prevent printing of the R code that generated the plot.