**Ph.D. Thesis proposal**

Objectives:

* Create a gridded product based on a earth system model to simulate regeneration processes.
* Develop a large-scale monitoring tool for restoration sites derived from publicly available data.
* Improve algorithms that represent secondary forest regrowth.
* Predict suitability of land for specific restoration strategies across the landscape.

Roughly 70% of forested areas has been affected by forest management, clearing and abandonment (FAO,2010). Natural regeneration processes have an important role in the resilience of forest during mortality events. Regeneration processes such as reproductive allocation, seed dispersal, germination and seedling survival are influenced by environmental variables such as light availability, water, temperature and soil composition. This variables can have a deterministic effect in where do trees regrow and at what rate.

Non-governmental organizations dedicated to tropical forest restoration often struggle to meet the goals set by the funds that finance their operations. It is even more difficult to set out a monitoring strategy that allows them to report the total restored land at a low cost. Large scale monitoring of forest regrowth is possible by using remote sensing techniques such as drone imagery, LIDAR technology, and satellites. Some of the challenges

There are vegetation demographic models to represent vegetation dynamics.

Moving from the ground, to the drones, to the satellites.

**STRI- Auto Crown project**

There is available drone imagery for sites such as 50 ha plot in BCI and the Ava plot, which are processed into point clouds, digital elevation models and Orth mosaics by using photogrammetry. STRI has conducted monthly flights during several years, and now we have weekly flights. There is potential to conduct analysis on tree mortality, phenology and flowering. However, crown mapping has been done manually with ground truthing by experienced field technicians, which is laborious and time consuming. The rate at which we are collecting RGB imagery calls for automatization of crown delineation and the consolidation of the multiple datasets collected in this site. Artificial intelligence and neural networks have been used to map crowns in other sites, these algorithms require large training datasets to perform.

The Auto Crown project aims to consolidate the latest years datasets into timeseries that track most of the crowns in the 50 ha BCI plot. To achieve this goal, we are testing available deep-learning tree crown delineation tools such as *DeepForest* and *Detecttree2* in our datasets and comparing F1 scores. Detectree2 is a python package deriving from the detectron2 framework for object detection. They offer methods for transfer learning and pre trained models for tree crown detection on the Linux platform. Deep forest is another python package for training and prediction of individual tree crowns, which is training with data from the National Ecological Observatory network. Furthermore, we explore alternative neural network framework to derive an algorithm trained in tropical forest crown.

The algorithms mentioned are convolutions neural networks (CNN) and are difficult to develop due to the large datasets required to detect trees, and we are limited on time and resources to develop large training datasets from manual crown delineation. The idea of transfer learning is to initialize the training of a new neural network from a previously trained CNN.

How to split the data to avoid overfitting? The issue is that training the AI algorithm with crowns delineated from BCI 50 ha plot results in overfitting. Alternatives to this issue is to train the dataset with newly delineated crowns from the AVA plot which is relatively close, the flights images are at the same altitude and collected in the same day. Or training with the BCI 50 ha plot delineated crowns and predict the crowns in the AVA plot. Other issues arise from the quality of the training datasets, which need to be more extensively evaluated.

Scores to evaluate models. Intersection over union (IoU) is defined as the intersection area between the predicted and ground truth area. A threshold is then defined to distinguish between True positives, false positive, or false negatives. Precision is defined as the true positives divided by the sum of true positives and false positives. Recall is defined as true positives divided by the sum of the true positives and the false negatives. The F1 score is given by . The total crown area can be compared between ground delineation and AI delineation datasets with R-square and RMSE scores.

**STRI- Precipitation reanalysis datasets compared to ACP rain-gauge data.**