Consultancy Report number 5, by Ophélie Ratel – October 2020

Analysis of secondary and primary exploited forest plots

Hypothesis

What are the effects of initial biomass on biomass productivity at the plot scale?

The higher the initial biomass, the lower the productivity.

What is the impact of landscape structure and dynamic to the recovery of forest productivity over time?

Landscape structure (configuration and composition) and dynamic -> impact on the regeneration potential of forests (proximity to crops, fragmentation dynamic, etc.).

The more complex the structure of the landscape, the more negatively it influences the rate of forest regeneration.

The increase in the number of interfaces between forests and intensive agriculture reduces the regeneration potential of forests.

The higher the dynamics of landscape fragmentation, the lower the potential for forest regeneration.

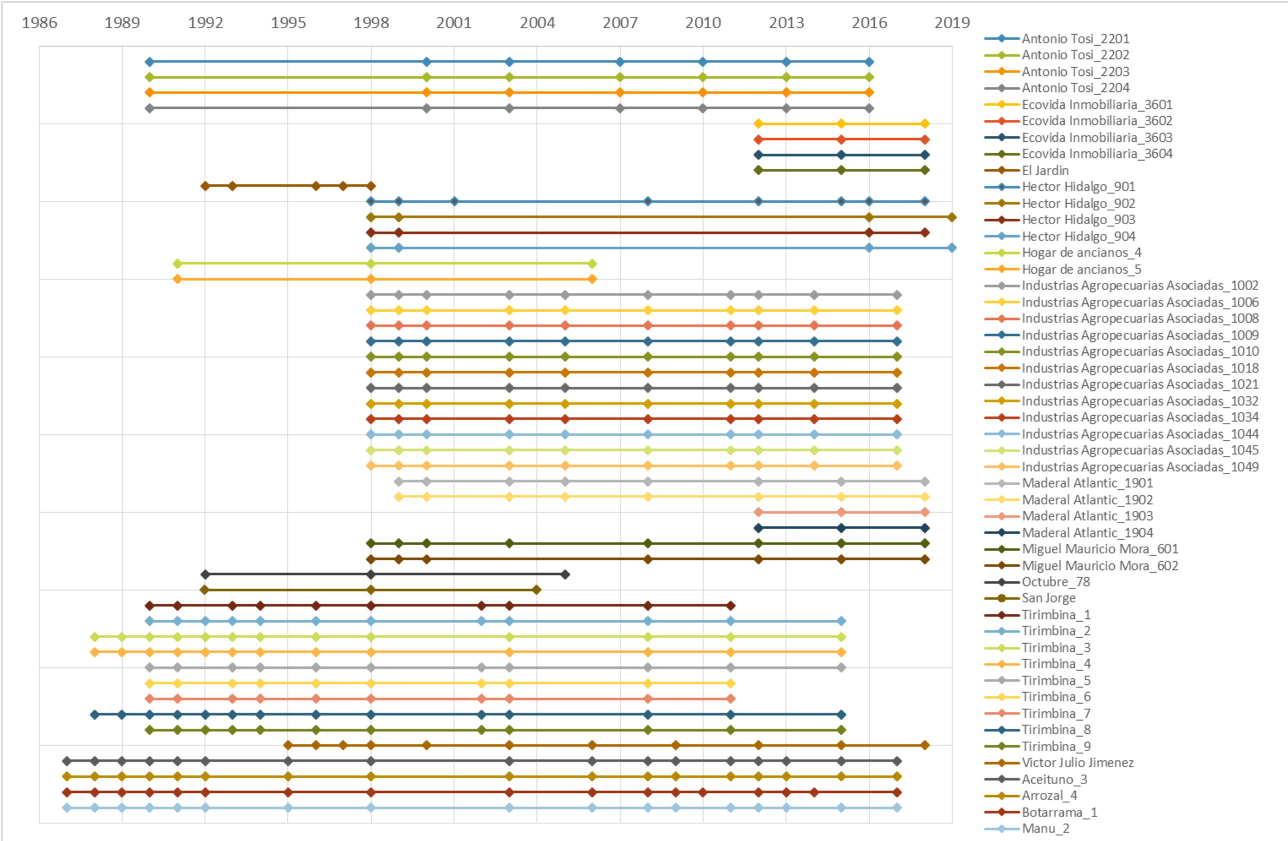
Data

For this analysis we have 49 plots from three institutions: CATIE, CODEFORSA and FUNDECOR. Four plots are within secondary forests, all the others are within primary exploited forest.

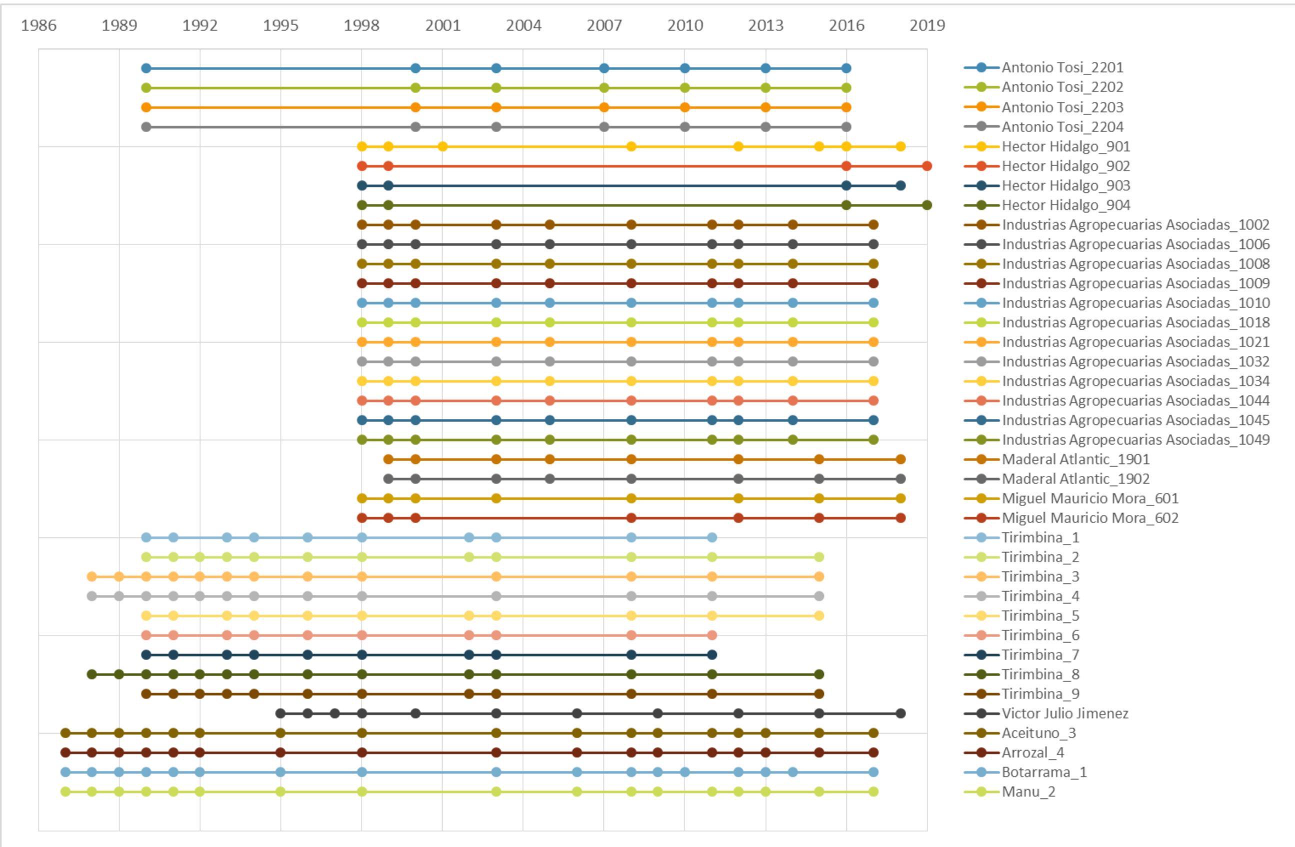
Institution	Forest type	Treatment	Plots number	Area (ha)
CATIE	Primary exploited forest	Exploited with silvicultural	9	1
	Secondary forest	Harvested and abandoned	4	1,16 ; 1,44 ; 1,6
CODEFORSA	Primary exploited forest	Exploited	5	1
FUNDECOR	Primary exploited forest	Exploited	30 1	0,3 1

- Number of sampled plots for each year between 1987 and 2019 :

This graph shows how many plots have been sampled every year. Each colour represents 1 plot. Samplings are represented by points. There are therefore 16 sites for 49 plots, half of the sites (8) consisting of only one plot.



Given the distribution of the data over time, only the plots sampled the most times over the longest period of time will be retained in the model. This represents 38 plots over the period 1990-2018. All these plots are not sampled at the same years but we could test two time periods, 1990-2010 and 1998-2018 for example.

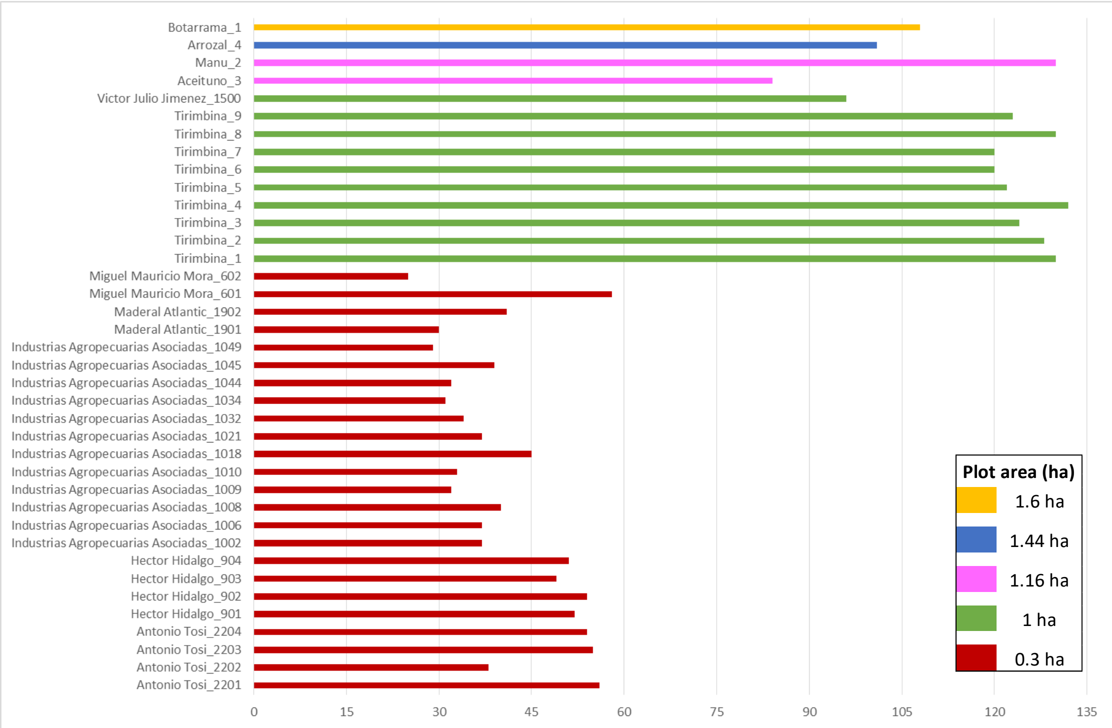


Synthesis of the plots implemented in the model

Id	Plots name	year n°1	final year	sampling number	sampling duration (years)
1	Aceituno*	1987	2017	17	30
6	Arrozal*	1987	2017	17	30
7	Botarrama*	1987	2017	17	30
35	Manu*	1987	2017	17	30
42	Tirimbina_3	1988	2015	13	27
43	Tirimbina_4	1988	2015	13	27
47	Tirimbina_8	1988	2015	13	27
41	Tirimbina_2	1990	2015	12	25
48	Tirimbina_9	1990	2015	12	25
44	Tirimbina_5	1990	2015	11	25
49	Victor Julio Jimenez	1995	2018	11	23
40	Tirimbina_1	1990	2011	10	21
45	Tirimbina_6	1990	2011	10	21
46	Tirimbina_7	1990	2011	10	21
19	Industrias Agro. Asos_1002	1998	2017	9	19
20	Industrias Agro. Asos_1006	1998	2017	9	19
21	Industrias Agro. Asos_1008	1998	2017	9	19
22	Industrias Agro. Asos_1009	1998	2017	9	19
23	Industrias Agro. Asos_1010	1998	2017	9	19
24	Industrias Agro. Asos_1018	1998	2017	9	19
25	Industrias Agro. Asos_1021	1998	2017	9	19
26	Industrias Agro. Asos_1032	1998	2017	9	19
27	Industrias Agro. Asos_1034	1998	2017	9	19
28	Industrias Agro. Asos_1044	1998	2017	9	19
29	Industrias Agro. Asos_1045	1998	2017	9	19
30	Industrias Agro. Asos_1049	1998	2017	9	19
36	Miguel Mauricio Mora_601	1998	2018	8	20
31	Maderal Atlantic_1901	1999	2018	8	19
32	Maderal Atlantic_1902	1999	2018	8	19
13	Hector Hidalgo_901	1998	2018	7	20
37	Miguel Mauricio Mora_602	1998	2018	7	20
2	Antonio Tosi_2201	1999	2016	7	17
3	Antonio Tosi_2202	1999	2016	7	17
4	Antonio Tosi_2203	1999	2016	7	17
5	Antonio Tosi_2204	1999	2016	7	17
14	Hector Hidalgo_902	1998	2019	4	21
16	Hector Hidalgo_904	1998	2019	4	21
15	Hector Hidalgo_903	1998	2018	4	20

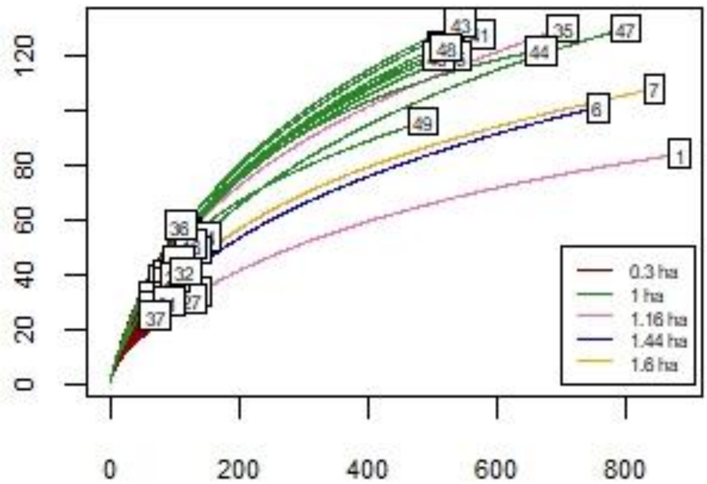
*Plots with a * are secondary forest plots.*

Specific richness

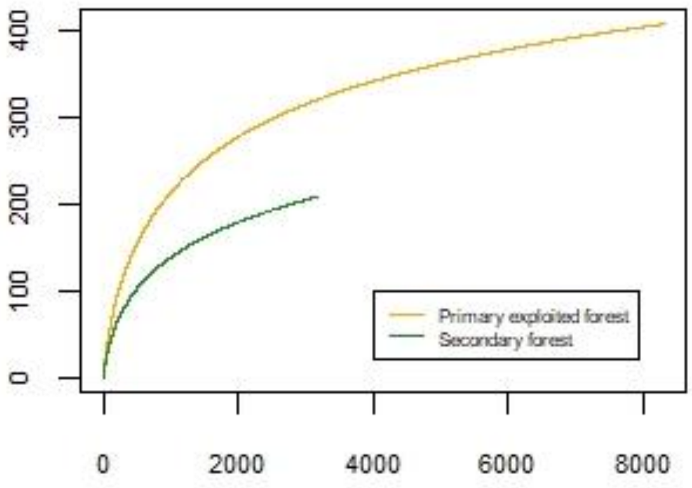


Specific richness per plot, according to plot area (ha). The first 4 lines correspond to secondary forest plots.

Given the large difference in surface area between the plots, an attempt was made to estimate whether the sampling effort had been sufficient to enable the specific richness of the different plots to be compared with each other. For this purpose, rarefaction curves were created for each plot as a function of the number of individuals *(with R package VEGAN).*



This rarefaction curve shows specific richness (y axis) according to sample size (x axis) and plot area (see colours). At the bottom left are the plots with a surface area of 0.3 ha, with a low number of individuals sampled. The id number corresponds to the first column in the synthetic plots table.

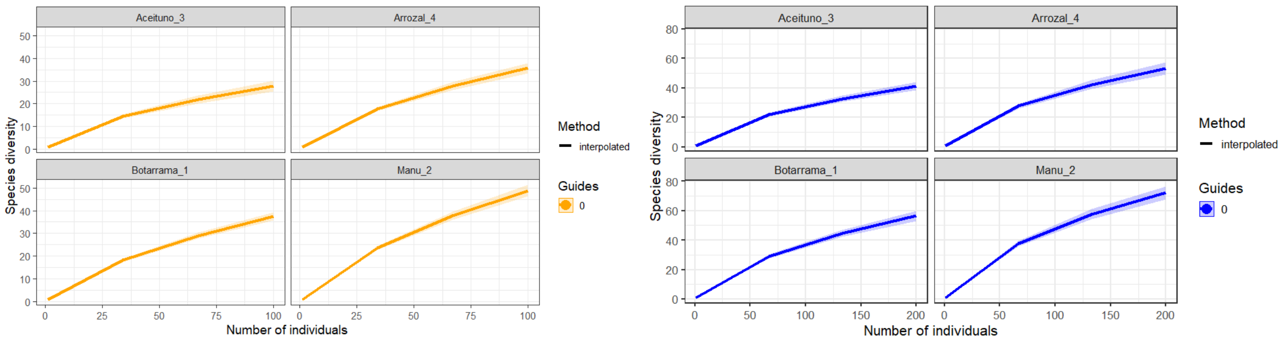


This rarefaction curve shows specific richness (y axis) according to sample size (x axis) and forest type (see colours). As there are only 4 plots of secondary forests, we see that the sampling effort is not optimal.

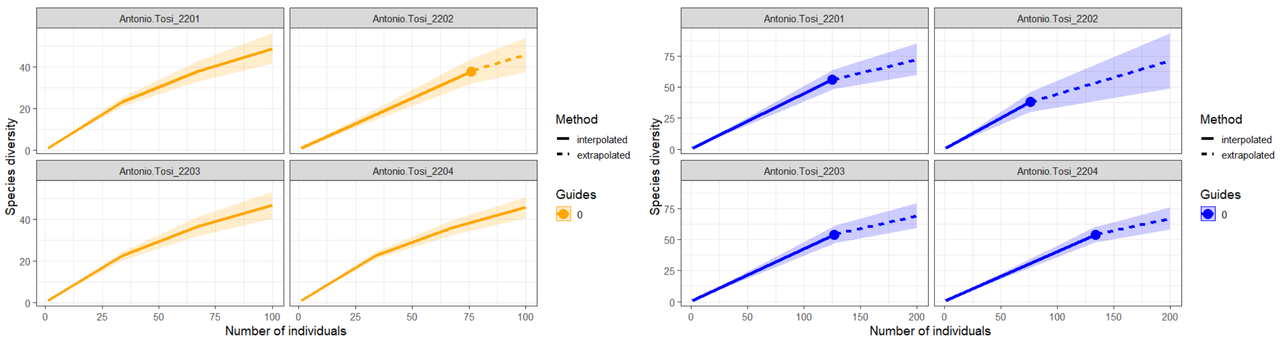
With these graphs it is clear that it is difficult to interpret the specific richness data since the majority of the plots have not reached their plateau and therefore are not representative of the real richness (not enough sampled individuals, and a too small plot area). For plots that have reached a plateau, we see that some sites are richer than others, this corresponds to secondary forests and larger plots.

Species richness interpolation/extrapolation

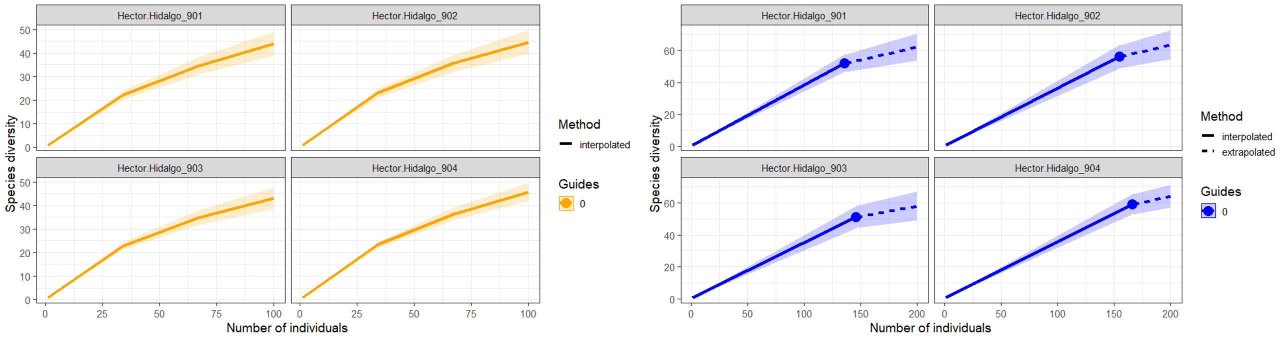
Using iNEXT R package, estimations of species richness have been calculated for each plot, using the observed sample of abundance to compute diversity estimates and the associated 95% (default) confidence intervals. Calculation was carried out for a number of individuals of 100 (orange curves) then 200 (blue curves).



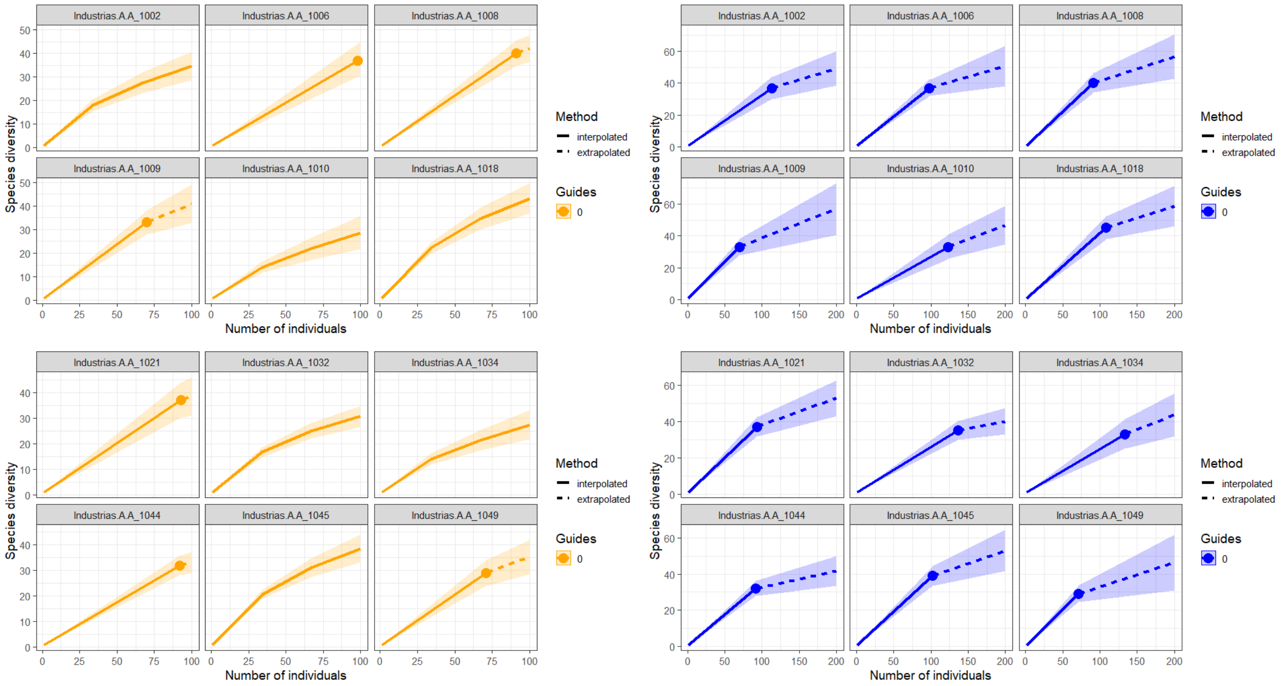
Interpolation curves (orange = 100 ind; blue = 200 ind) for the 4 secondary forest plots (Aceituno_3, Arrozal_4, Manu_2 & Botarrama_1).



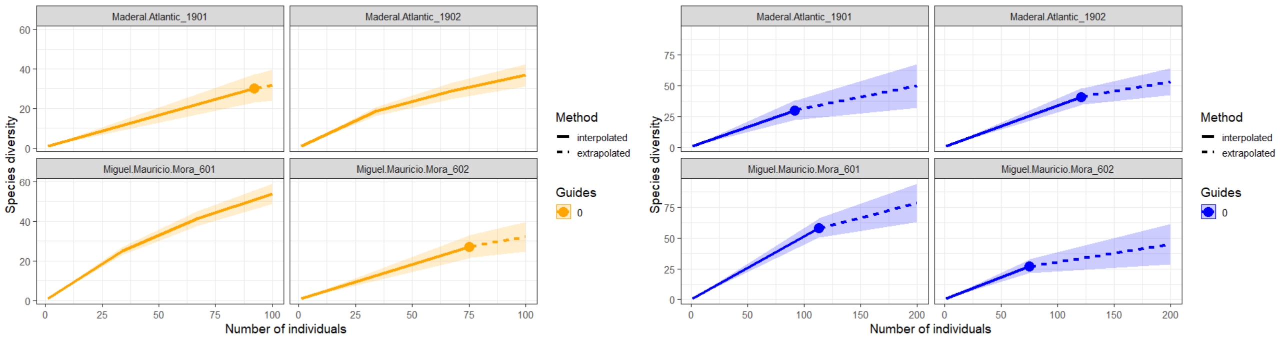
Interpolation/extrapolation curves (orange = 100 ind; blue = 200 ind) for the 4 primary exploited plots in Antonio Tosi site.



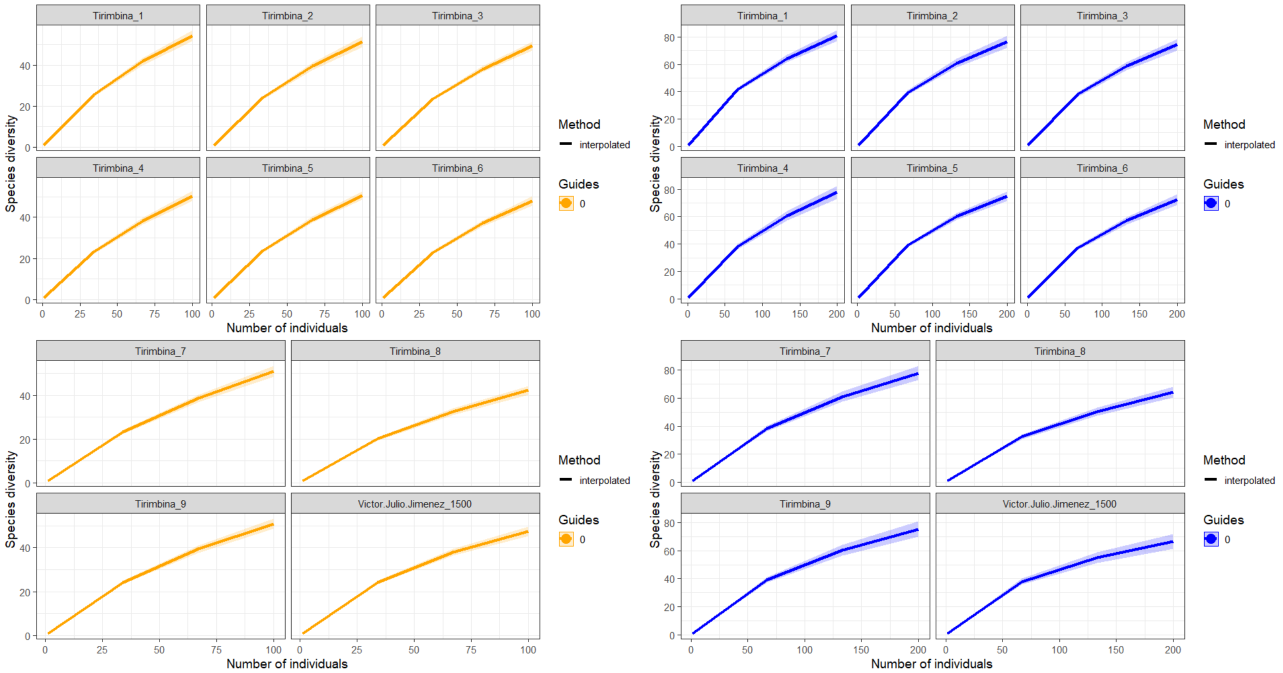
Interpolation/extrapolation curves (orange = 100 ind; blue = 200 ind) for the 4 primary exploited plots in Hector Hidalgo site.



Interpolation/extrapolation curves (orange = 100 ind; blue = 200 ind) for the 12 primary exploited plots in Industrias Agropecuarias Asociadas site.



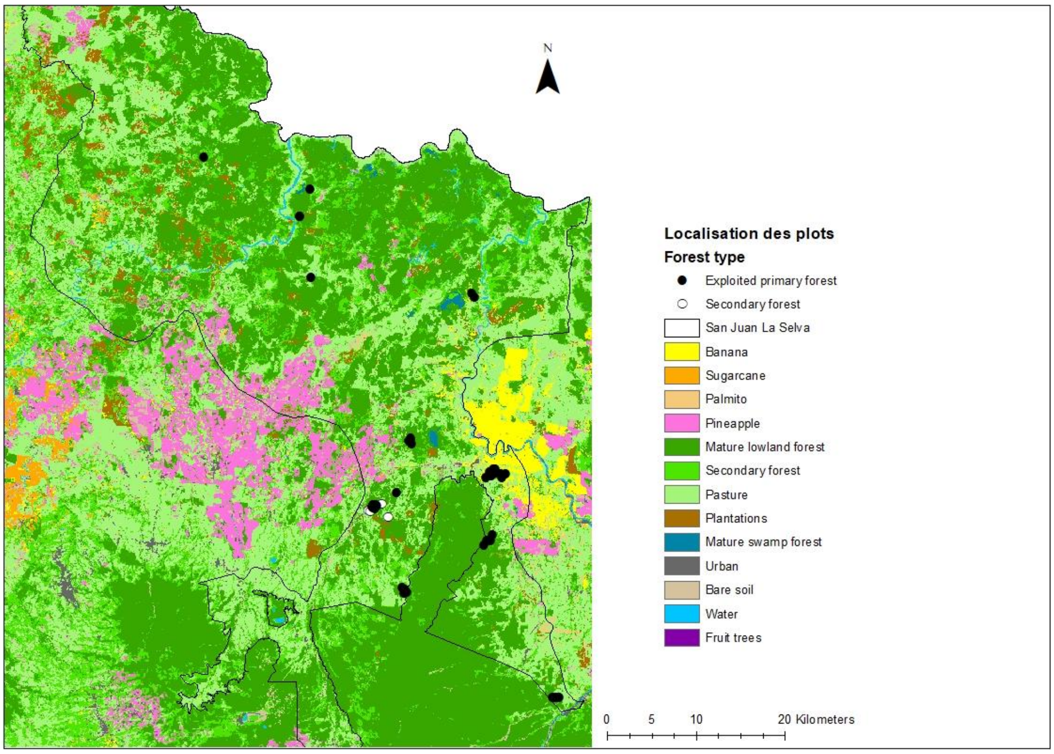
Interpolation/extrapolation curves (orange = 100 ind; blue = 200 ind) for the primary exploited plots in Maderal Atlantic and Miguel Mauricio Mora sites.



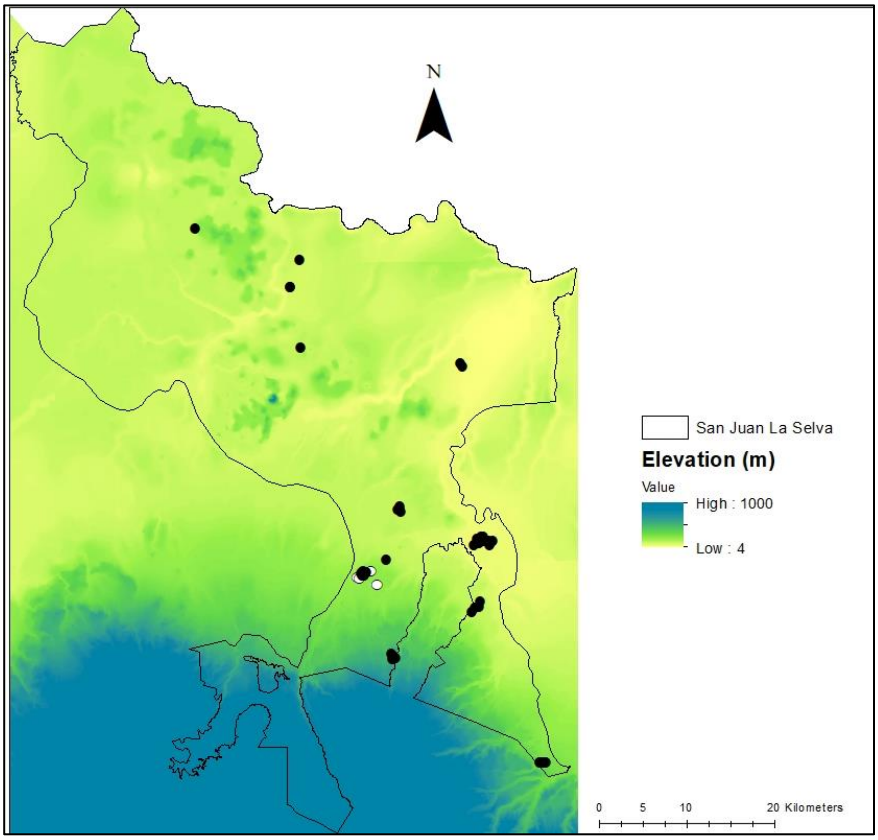
Interpolation/extrapolation curves (orange = 100 ind; blue = 200 ind) for the primary exploited plots in Tirimbina and Victor Julio Jimenez (bottom right) sites.

- Plots location in the study site :

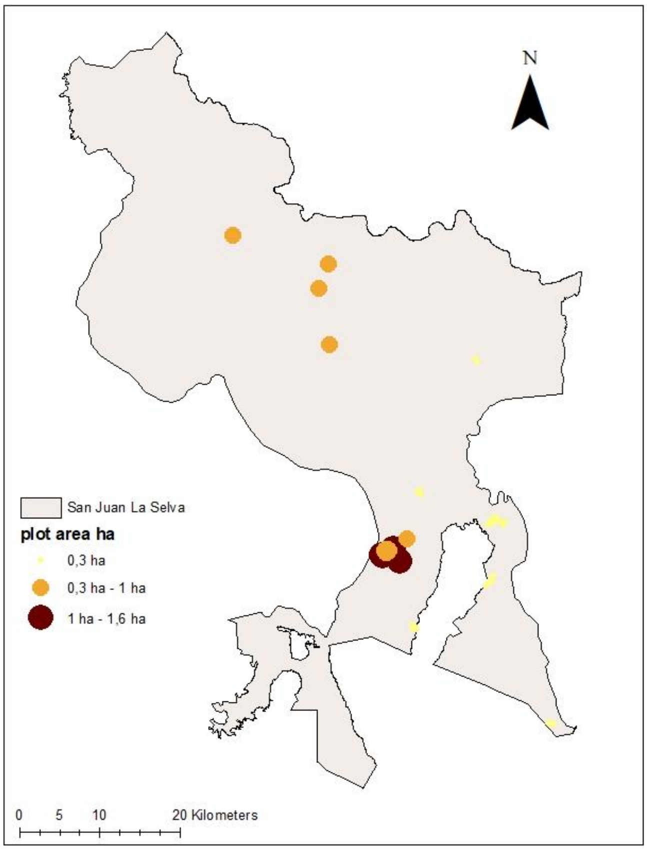
According to land-use classification



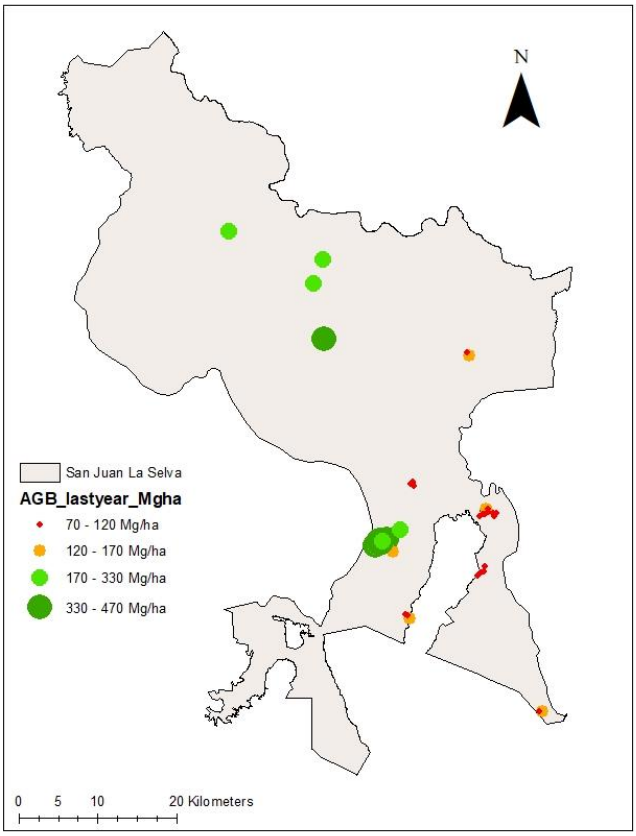
According to elevation (in meters)



- Plots are represented according to their (1) surface area and (2) biomass:



(1)



(2)

For each plot, biomass value (AGB) has been calculated from the last sampling year. It is the sum of biomass per plot and per hectare, expressed in Mg per ha.

NB: For the rest of the analysis, biomass rate will be assessed within the period of study chosen, by taking the delta biomass value between the first and last year of sampling, 1990-2010 for some of the plots and 1998-2018 for the others.

Model Framework

The idea would be to express the quantity of biomass recovered over time by including co-variables (landscape, climate, topography, soil, functional component) and developing our model in a Bayesian framework, which is particularly adapted when there is little data, thanks to the addition of information that could be described as "non-pure data" or "priors". These priors can be based on previous studies or expert knowledge: they are a way to make sure that our predictions are within the range of acceptable values given our prior knowledge on similar processes, and are thus especially important when data is scarce. Moreover, the Bayesian approach allows a rigorous estimation of parameters correlation and uncertainty.

Variable to explain : X = Quantity of biomass recovered over time (biomass production)

Explicative variables : Y = landscape metrics, climate, topography, soil, functional component, plot history (year of disturbance, land-use history on secondary forest plots, exploitation type in logged forest).

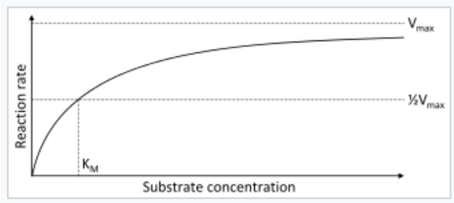
Proposed model:

We chose a model for biomass accumulation during forest succession with the following criteria:

- Simplicity (low number of parameters),
- Interpretability (parameters can be understood in terms of ecological quantities),
- Saturation (biomass cannot increase forever).

The Michaelis-Menten model (kinetic equation) fills the above-mentioned criteria. This model gives a saturation curve according to a parameter, here the maximum biomass value reachable in a forest (Y axis) over time (X axis).

$$v_i = \frac{v_{\max} \times [S]_0}{K_M + [S]_0}$$



With:

- Vmax = maximum biomass value = Vmax(t0) + covariable 1, 2 … n
- Km = growth rate to reach 50% of the maximum biomass = Km + covariable 1, 2 … n

[S]0 : here it would be the recovery time Then the objective would be to attribute to each parameter the co-variables likely to impact it. For example, we could say that the maximum biomass value can be impacted by soil characteristics, climate, topography or the functional component. While the growth rate may be influenced by landscape metrics or the plot history.

Therefore a theoretical expression of this model could be :

$$V\,\mathrm{max} = \frac{Vmax(t0) + \alpha 1.cov1 + \alpha 2.cov2 + \alpha 3.cov3}{Km + \alpha 4.cov4 + \alpha 5.cov5}$$

cov1 = soil characteristics

cov2 = climate

cov3 = functional component

cov4 = lanscape metrics

cov5 = plot history

A further question is whether we can use both types of forest (secondary or exploited) in the same modeling framework. Otherwise, there may be too little data to implement two different models.

We plan to implement this model in a Bayesian framework, using the following priors:

- **Vmax** : we expect the maximum AGB value of secondary forests to be similar to the AGB of old-growth forests. We therefore chose a prior that follows a normal distribution with a mean of 300 (in Mg/ha), and a standard deviation of 150. The high standard deviation reflects our uncertainty on the expected maximum AGB value.
- **Km:** we expect the time to reach Vmax/2 to be around 30 years for a secondary forest (Poorters et al., 2016). We chose a prior that follows a normal distribution with a mean of 30 (in years), and a standard deviation of 15.