Measuring carbon using random forests to better assess forest conservation policy

Ilan Havinga[1*], Lars Hein[1], Mauricio Vega-Araya[2], Antoine Languillaume[3]
[1]Environmental Systems Analysis Group, Wageningen University
[2]Instituto de Investigación y Servicios Forestales, Universidad Nacional, Costa Rica
[3]Biometris, Wageningen University
[*]*Email: ilan.havinga@wur.nl*

1. Introduction

- The **Pago de Servicios Ambientales** scheme in **Costa Rica** reimburses private landowners for **forest ecosystem services**
- **Carbon storage** is one ecosystem service targeted by the scheme
- The **additional amount** of ecosystem services generated as a result of such a scheme is a **key measure** of **effectiveness**
- However, **ecosystem services** are **rarely quantified** to measure the effectiveness of payments

2. Research questions

RQ1. What is the spatial distribution of carbon storage in Costa Rica?
RQ2. What is the additional amount of carbon storage generated by the Pago de Servicios Ambientales scheme?

3. Method

3.1. Spatial distribution of carbon

- Employed machine learning algorithm **random forests**
- Random forests uses **decision trees** and **bootstrap sampling** to predict a response variable using a set of predictor variables
- **155 forest inventory plots** used to train response (carbon)
- **11 remote sensing** and **climatic variables** chosen as predictors

Table 1. Carbon storage predictor variables.

Dataset	Predictor variable
MODIS Nadir BRDF-Adjusted Reflectance Daily	Band 1 - blue (0.459-0.479 μm)
	Band 3 - red (0.620-0.670 μm)
	Band 4 - near-infrared (0.841-0.876 μm)
	Band 5 - short-wave infrared (1.230-1.250 μm)
	Band 6 - short-wave infrared (1.628-1.654 μm)
	Band 7 - short-wave infrared (2.105-2.155 μm)
	NDMI - normalised difference moisture index
WorldClim 2 climate surfaces	Bio7 - Temperature annual range (ºC)
	Bio12 - Annual precipitation (mm)
	Bio17 - Precipitation of the driest quarter (mm)
Digital elevation model	Elevation (m)

3.2. Measuring additional carbon stored

- **Two linear models** generated with PSA, non-PSA and national parks as categorical (policy) variable
- **400 points** randomly sampled in each area
- Added **covariates** slope, distance to road and nearest population
- **Tukey test** performed to measure differences in adjusted means of categorical policy variable

$$EQ1.\ \ y_{ij} = \mu + \beta_{policy}\ P_i + \beta_{slope}\ S_{ij} + \beta_{road}\ R_{ij} + \beta_{pop}\ Po_{ij} + \varepsilon_{ij}$$

$$EQ2.\ \ y_{ijk} = \mu + \beta_{policy}\ P_i + \beta_{slope}\ S_{ij} + \beta_{road}\ R_{ij} + \beta_{pop}\ Po_{ij} + \beta_{year}\ Y_k + \varepsilon_{ijk}$$

- Model **EQ1.** considers carbon storage in a single year (2013)
- Model **EQ2.** considers carbon storage over time (2011 and 2013)

4. Results

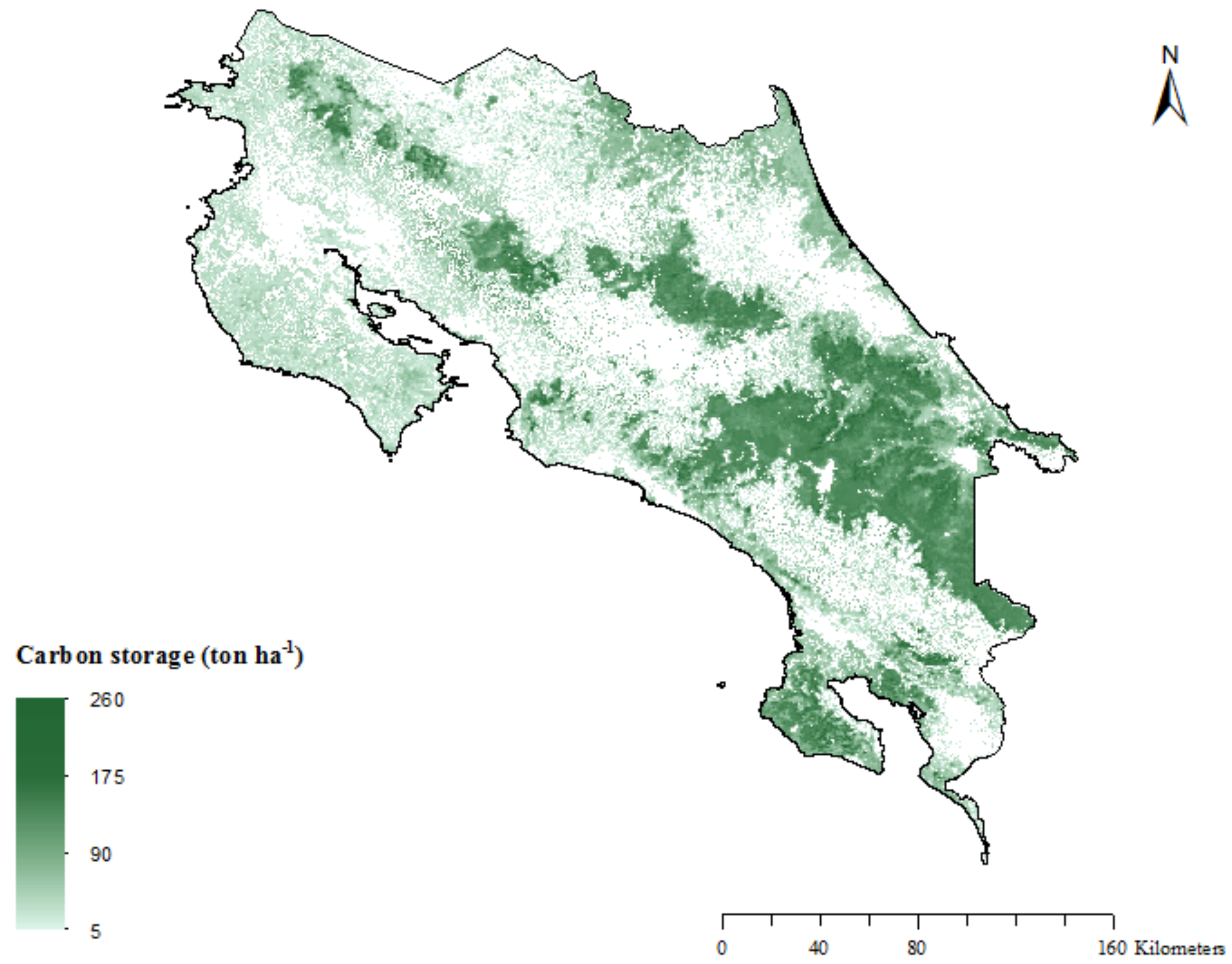
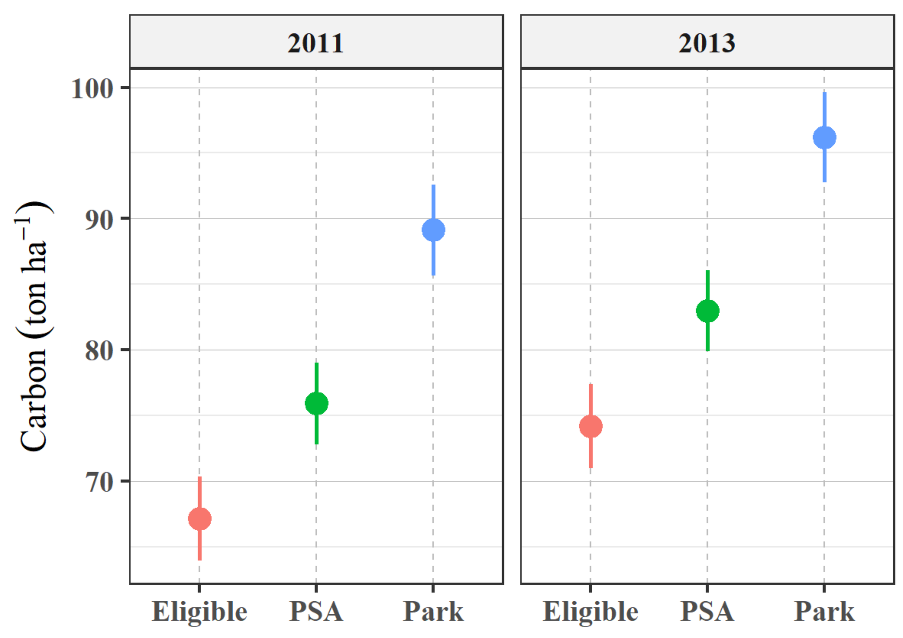


Figure 1. Estimated carbon storage in Costa Rica (ton ha^{-1}). The random forest model had an R^2 of 0.63.

Table 2. Carbon storage.

Land cover	Min (ton ha^{-1})	Median (ton ha^{-1})	Mean (ton ha^{-1})	Max (ton ha^{-1})	Total area (m ha)	Total carbon (Mt)
Primary forest	6.0	78.9	87.4	262.3	2.2	189.3
Secondary forest	6.1	42.2	49.0	253.6	0.6	28.8
Tree plantations	5.6	36.4	38.5	238.9	0.2	6.4
Palm swamp forest	6.4	55.6	59.2	188.2	0.1	8.0
Mangrove forest	9.0	55.3	57.5	201.3	0.05	2.8
Overall	5.6	57.0	75.8	262.3	3.1	235.3



Contrast	Estimate	SE	p-value
PSA – Eligible	8.76	1.91	<.0001
PSA – Park	-13.23	2.23	<.00001
Eligible – Park	−21.99	2.35	<.00001
2013 – 2011	7.06	1.54	<.00001

Figure 2. Adjusted means and contrasts for carbon storage per policy type for EQ2. The points represent the mean carbon stored (ton ha^{-1}), adjusted for the covariates, while the bars represent the 95% confidence intervals. The model had an R^2 of 0.42.

5. Discussion

- No significant difference found between carbon stored in PSA and non-PSA areas in 2013 alone (**EQ1.**)
- However, PSA areas enrolled in both 2011 and 2013 (**EQ2.**) stored an additional 9 tonC ha-1: suggests larger **long-term effect**
- Model struggled to predict **high biomass regions** but results in line with **official statistics** (Ecosytem accounts: 243 MtonC)
- Spatial data assits prioritisation of **high-value areas** and **compliance** but **direct compensation less feasible** due to model uncertainties

6. Conclusion

- Spatially quantifying ecosystem services to examine a payments for ecosystem service scheme is **feasible**
- This enables a **direct quantification** of the additional services delivered and can assist administrators improve policy effectiveness

Source: *Havinga, I., Hein, L., Vega-Araya, M., Languillaume, A., 2020. Spatial quantification to examine the effectiveness of payments for ecosystem services: A case study of Costa Rica's Pago de Servicios Ambientales. Ecol. Indic. 108, 105766.*