

Supplementary Information for "Linking land-use and land-cover transitions to their ecological impact in the Amazon"

Cássio Alencar Nunes, Erika Berenguer, Filipe França, Joice Ferreira, Alexander C. Lees, Julio Louzada, Emma J. Sayer, Ricardo Solar, Charlotte Smith, Luiz E. O. C. Aragão, Danielle de Lima Braga, Plinio Camargo, Carlos Cerri, Raimundo Cosme, Mariana Durigan, Nárgila Moura, Victor Hugo Fonseca Oliveira, Carla Ribas, Fernando Vaz-de-Mello, Ima Vieira, Ronald Zanetti, Jos Barlow

Cássio Alencar Nunes Email: cassioalencarnunes@gmail.com

This PDF file includes:

Supplementary Methods Figures S1 to S12 Tables S1 to S7 SI References

Supplementary Methods

Detailed information on above-ground biomass and biodiversity sampling

Information on biomass estimation

To estimate tree biomass, species-specific wood density was obtained from the Global Wood Density Database (1). If a species-specific wood density was not available, the average of the genus or family was used. For all trees and saplings, we used the allometric equation of Chave et al (2) to calculate biomass based on diameter at breast height (DBH) and wood density, except for the genus *Cecropia* for which we used the Nelson allometric equation (3). We used the Gerwing & Farias equation (4) to estimate the biomass of lianas.

Biomass estimation of litter components (fine wood debris and leaf litter) was made by weighing the samplings after drying to constant weight. For coarse woody debris, each piece of dead wood was classified in terms of its decomposition state, and biomass was calculated by multiplying its volume by its density. More details on biomass estimation can be found in refs (5, 6).

Detailed information of animal sampling

Birds: Sampling was conducted at three points along each transect separated by 150 m (0, 150 and 300 m). At each of these three points, 15-minute point count samples were collected between 15 minutes before dawn and 9:30 am. These surveys were repeated at each point in the transect, but in reverse order to remove temporal bias. Solid-state sound recorders were used to record each point count survey to facilitate identification of any species not identified in the field. See ref (7) and ref (8) for details on the bird survey and a list of voucher species collected in Santarém and Paragominas, respectively.

Ants: We used passive (non-baited) pitfall traps in the Santarém region, whereas in Paragominas, the traps were baited with sardine and honey. Traps were plastic containers (12 cm height and 8 cm diameter) half-filled with a solution of water, salt (2%) and detergent (5%), and were left in the field for 48 h. In Paragominas, six traps were installed per transect, separated by 50 m (between 25 and 275 m in each transect). In Santarém, 10 traps were installed along each transect separated by 10 m (between 50 – 90 and 200 – 240 m in each transect). All collected specimens were kept in ethanol and then were taken to the Universidade Federal de Lavras and Universidade Federal de Viçosa – Minas Gerais – Brazil, where they were identified to the lowest taxonomic level. Final identification was checked by an expert taxonomist (Dr. Rodrigo Feitosa at the Museu de Zoologia de São Paulo). For a full list of voucher species collected in Paragominas see ref (9).

Dung Beetles: We used baited pitfall traps, which were 1-litre plastic containers (4 cm diameter and 9 cm height) with a bait container filled with 50 g of dung (80% pig and 20% human). All traps were halffilled with a solution of water, salt (2%) and detergent (5%). Dung beetles were sampled at three points along each transect separated by 150 m (0, 150 and 300 m). In each of these three sampling points, 3 pitfalls were installed in the corners of a triangle (3-m per side) and were left in the field for 48 h. The individuals were kept in ethanol for preservation and taken to the Universidade Federal de Lavras and Universidade Federal de Viçosa – Minas Gerais – Brazil for identification. Final identification was checked by an expert taxonomist (Dr. Fernando Zagury Vaz-de-Mello at the Universidade Federal do Mato Grosso).

Orchid bees: we collected male orchid bees using baited traps, made of 2-litre plastic bottles (0 cm diameter and 35 cm height). Traps were installed at four sampling points along each transect in the Paragominas region, and were separated by 50 m (50, 100, 150 and 200 m). In each bottle, radial holes were made at 20 cm height, where we inserted flower-like structures that were lined with coarse sand to give support to the bees. Inside the bottle, a stick with cotton ball was baited with eugenol, methyl salicylate, vanilla or eucalyptol (each trap received only one type of bait). These traps were tied to tree trunks 1.5 m above the ground and were left in the field for 48 h. Bees were kept frozen before sorting out and identification by an expert taxonomist (Dr. André Nemésio at EMBRAPA – Amazônia Oriental). For a full list of voucher species collected see ref (9).

Model validation analysis

Visual inspection of diagnostic plots revealed that some of our linear mixed-effect models (LMMs) presented residuals with deviations from a normal distribution, outlier influence or unequal variances (Fig. S9, S10 and S11). Most of these violations were relatively small and a result of low variance in certain response variables for pastures and mechanized agriculture land-use and land-cover (LULC) classes. Nonetheless, we ran a validation analysis to verify that the results of our LMMs were robust.

First, we ran non-parametric Quantile Generalized Additive Models (qGAM) using the function '*qgam*' from R package 'qgam' (10, 11). We used the standardized ecosystem variables as response variables and the LULC classes as the explanatory variable, together with mean elevation, mean slope and clay content as co-variates (i.e., the same model structure as our LMMs). We also included catchment ID and region as smoothed random factors in the qGAMs. The main difference between the qGAMs and LMMs was that we constructed the qGAM models based on the 0.5 quantile (i.e., the median), whereas the LMMs compared means. Second, we ran the simultaneous comparisons of median values in the different LULC class, using contrast matrices and accounting for inflation of type I error (in the same way as for LMMs). We used the functions '*glht*' and '*confint*' from 'multcomp' package (12) to obtain the standardized effect sizes and their confidence intervals for each LULC transition.

Finally, to validate our LMMs, we compared the effect sizes obtained using qGAMs with the effect sizes obtained using the LMM approach. We ran correlation tests using the base-R function 'cor.test'. The tests revealed that the effect sizes were highly correlated, with most of the *r* coefficients = 0.99 (Fig. S12). The only two models with lower *r* coefficients were for the soil carbon pool and soil Na, with $r = 0.88$ and $r = 0.90$, respectively. Crucially, these were non-significant models, so the weaker correlation does not affect the results. In terms of significance (p-value), of the 270 pairwise comparisons (18 variables x 15 transitions), 258 (95.5%) remained the same side of the significance threshold (0.05). Of the transitions that changed their significance (Table S7), most (3.7%, $n = 10$) revealed transitions that became significant using the $qGAM$, and only 0.75% (n = 2) became nonsignificant with the gGAM. There was no clear pattern in these changes in significance, and they occurred in both models that met all assumptions of goof fit and in models that presented deviations from normal distribution and unequal variances. Thus, we confirmed that the LMMs with residuals deviating slightly from a normal distribution and unequal variances between LULC were indeed robust, strengthening our results and conclusions.

Fig. S1. Study sites and design. **(a)** Map of the study sites in the Santarém (STM) and Paragominas (PGM) regions of the eastern Amazon, in the Brazilian state of Pará. **(b)** Diagram of plots, subplots, and sampling design points to collect ecological variables representing biodiversity, carbon pools and soil properties at the study sites, where CWD = coarse woody debris and FWD = fine woody debris.

Fig. S2. Land use and cover transition (LULCT) rates in the Brazilian Amazon. Mean annual LULCT rates from (a) 2006-2019, based on land-use change maps (13) and (b) 2006-2014, based on forest degradation maps (14). Primary forests have never been clear cut, and secondary forests are regenerating forests. Young secondary forests are < 20 years old and old secondary forests are ≥ 20 years old. Agriculture includes perennial and temporary crops. The width of the arrows is proportional to the mean annual rate of the transition.

Fig. S3. Impacts of 18 land-use and land-cover transitions on the species richness of seven biodiversity groups in the Brazilian Amazon. Values and size of symbols represent the species richness averages based on 21 undisturbed primary forests (UF), 68 logged primary forests (LF), 65 logged-and-burned primary forests (LBF), 72 pastures (PA), 26 mechanized agriculture fields (MA), 33 young secondary forests (SFy, < 20 years old) and 25 old secondary forests (SFo, ≥ 20 years old). Arrows indicate the transitions and their effect, where grey indicates no effect, green is a significant increase and blue a significant decrease.

Fig. S4. Impacts of 18 land-use and land-cover transitions on the forest species richness of seven biodiversity groups in the Brazilian Amazon. Values and size of symbols represent the species richness averages based on 21 undisturbed primary forests (UF), 68 logged primary forests (LF), 65 logged-and-burned primary forests (LBF), 72 pastures (PA), 26 mechanized agriculture fields (MA), 33 young secondary forests (SFy, < 20 years old) and 25 old secondary forests (SFo, ≥ 20 years old). Arrows indicate the transitions and their effect, where grey indicates no effect, green is a significant increase and blue a significant decrease.

Fig. S5. Impacts of 18 land-use and land-cover transitions on four ecosystem carbon pools in the Brazilian Amazon. Values and size of symbols represent the averages of carbon pools in Mg ha⁻¹ based on 21 undisturbed primary forests (UF), 68 logged primary forests (LF), 65 logged and burned primary forests (LBF), 72 pastures (PA), 26 mechanized agriculture fields (MA), 33 young secondary forests (SFy, < 20 years old) and 25 old secondary forests (SFo, ≥ 20 years old). Arrows indicate the transitions and their effect, where grey indicates no effect, green is a significant increase and blue a significant decrease.

Fig. S6. Impacts of 18 land-use and land-cover transitions on seven soil properties at 0-10 cm depth in the Brazilian Amazon. Values represent the averages based on 21 undisturbed primary forests (UF), 68 logged primary forests (LF), 65 logged-and-burned primary forests (LBF), 72 pastures (PA), 26 mechanized agriculture fields (MA), 33 young secondary forests (SFy, < 20 years old) and 25 old secondary forests (SFo, ≥ 20 years old). Arrows indicate the transitions and their effect, where grey indicates no effect, green is a significant increase and blue a significant decrease. The size of symbols indicates nutrient concentrations or pH, where AI: Aluminium (mmolc dm⁻³), Ca+Mg: Calcium + Magnesium (mmolc dm⁻³), K: Potassium (mg dm⁻³), N: Nitrogen (%), Na: Sodium (mg dm⁻³), P: Phosphorus (mg dm-3).

Fig. S7. Annual rates (km² yr⁻¹) of land-use and land-cover transitions. Data for the Brazilian Amazon from 1985 to 2019, based on the MapBiomas collection 5 dataset (13).

Fig. S8. Annual rates (km² yr⁻¹) of land-use and land-cover transitions related to primary forest degradation and deforestation. Data for the Brazilian Amazon from 1992 to 2018, based on (14). The rates are the mean values for the intervals shown on the X axis.

Fig. S9. Model diagnostics for variables of the biodiversity ecosystem component. We checked all models for both overdispersion and homoscedasticity using the package DHARMa in R (10, 15). When variance was not homogeneous, we checked the boxplots of residuals per land-use and land-cover class (LULC) to decide about the adequacy of the models. Biodiversity models included species richness as response variables. The number of replicates are 21 undisturbed primary forests (UF), 68 logged primary forests (LF), 65 logged-and-burned primary forests (LBF), 72 pastures (PA), 26 mechanized agriculture fields (MA), 33 young secondary forests $(SFy, < 20$ years old) and 25 old secondary forests $(SFo, \geq 20$ years old).

Fig. S10. Model diagnostics for variables of the carbon ecosystem component. We checked all models for both overdispersion and homoscedasticity using the package DHARMa in R (10, 15). When variance was not homogeneous, we checked the boxplots of residuals per land-use and land-cover class (LULC) to decide about the adequacy of the models. Carbon models included the values of carbon pools in Mg ha $^{-1}$ as response variables. The number of replicates are 21 undisturbed primary forests (UF), 68 logged primary forests (LF), 65 logged-and-burned primary forests (LBF), 72 pastures (PA), 26 mechanized agriculture fields (MA), 33 young secondary forests (SFy, < 20 years old) and 25 old secondary forests (SFo, ≥ 20 years old).

Fig. S11. Model diagnostics for variables of the soil ecosystem component. We checked all models for both overdispersion and homoscedasticity using the package DHARMa in R (10, 15). When variance was not homogeneous, we checked the boxplots of residuals per land-use and land-cover class (LULC) to decide about the adequacy of the models. Soil models included the values of nutrient concentrations as response variables; AI: Aluminium (mmolc dm⁻³), Ca+Mg: Calcium + Magnesium (mmolc dm⁻³), K: Potassium (mg dm⁻³), N: Nitrogen (%), Na: Sodium (mg dm⁻³), P: Phosphorus (mg dm⁻³). The number of replicates are 21 undisturbed primary forests (UF), 68 logged primary forests (LF), 65 logged-and-burned primary forests (LBF), 72 pastures (PA), 26 mechanized agriculture fields (MA), 33 young secondary forests (SFy, $<$ 20 years old) and 25 old secondary forests (SFo, \geq 20 years old).

Fig. S12. Results of model validation analysis showing the relationship between the standardized effect sizes for each LULC transition calculated using linear mixed-effect models (LMM) on the X axis and the standardized effect sizes obtained using quantile generalized additive models (qGAM) on the Y axis. Circle colors show the changes in significance when comparing the two approaches (LMM with qGAM). The correlation coefficients (*r*) are shown for each response variable.

Table S1. Number of sampling sites of each land-use and land-cover type in the two sampling regions in the Brazilian Amazon, State of Pará.

Table S2. Mean annual rates (km² yr⁻¹) of land-use and land-cover transitions (LULCT) in the Brazilian Amazon and their standard deviation (Sd). The rates were calculated and estimated (*) based on land-use change maps (2006-2019; (13)) and on forest degradation maps (2006-2014; (14) .

Table S3. Mean values of the 18 ecological variables sampled in 310 transects over two regions in the Brazilian Amazon, State of Pará. The values are shown for each land-use or land-cover type. Units for biodiversity variables is species richness, for carbon pools is Mg ha⁻¹, and for soil variables are AI and Ca+Mg = mmolc dm⁻³, K, Na and P = mg dm⁻³, and N is %. UF: Undisturbed primary forest, LF: Logged primary forest, LBF: Logged-and-burned primary forest, SFo: Old secondary forest, SFy: Young secondary forest, PA: Pasture, MA: Mechanized agriculture.

Ecosystem component	Variable	UF	LF	LBF	PA	МA	SFy	SFo
Biodiversity	Bird	51.24	50.35	47.12	27.96	11.35	38.67	39.84
	Ant	24.48	24.57	26.83	19.56	13.62	23.90	23.32
	Dung beetle	24.57	19.49	17.75	7.65	7.50	15.03	18.76
	Orchid bee	7.88	6.76	6.52	4.29	2.60	9.47	5.80
	Liana	13.57	14.03	11.02	0.04	0.00	6.03	8.96
	Small tree	57.81	64.51	53.63	0.96	0.00	30.61	47.48
	Large trees	65.00	61.79	48.49	0.19	0.00	17.88	39.60
Carbon	Above-ground pool	196.15	140.36	105.00	0.48	0.00	29.64	70.83
	Litter pool	7.24	8.23	8.01	1.69	0.84	6.44	6.60
	Dead wood pool	20.44	28.45	25.02	1.62	0.00	8.26	10.73
	Soil pool	48.13	60.71	57.48	47.83	52.55	53.07	58.17
Soil	pH	4.04	4.30	4.53	5.08	4.88	4.56	4.23
	N	0.11	0.16	0.15	0.11	0.14	0.13	0.16
	P	3.06	3.70	3.64	4.50	5.58	4.22	3.86
	K	25.15	35.82	38.66	44.25	55.19	32.75	32.34
	Na	19.26	26.50	26.43	26.82	27.50	19.40	19.15
	CaMg	0.70	1.31	1.96	2.41	2.99	1.78	1.51
	Al	1.62	1.53	1.19	0.55	0.61	1.04	1.61

Table S4. Percentage changes in the mean values of 18 ecological variables in response to 18 land-use and land-cover transitions (LULCT) in the Brazilian Amazon. Positive values represent an increase in the variable, while negative values represent the percentage of decrease. An increase above 100% is represented as a multiplier and marked in bold. NA values are increases from 0, making it impossible to give percentages. Biodiversity is the species richness of the listed groups. UF: Undisturbed primary forest, LF: Logged primary forest, LBF: Logged-and-burned primary forest, SFo: Old secondary forest, SFy: Young secondary forest, PA: Pasture, MA: Mechanized agriculture.

Table S5. Complete list of model χ², degrees of freedom (Df), p-values and p-values for overdispersion and homoscedasticity tests. Small trees = DAP < 10 cm, large trees = $\text{DAP} \ge 10$ cm. LULC: land-use and land-cover, CLAY: clay content, ELEV_MEAN: mean elevation of transect (m.a.s.l), SLOPE_MEAN: mean slope of transect.

Table S6. Effect sizes, the lower and upper confidence interval (CI), the p-value and the standard error (Std error) of the effect size of all land-use and land-cover transitions (LULCT) on the different ecological variables of biodiversity, carbon, and soil ecosystem components. The column "Effect" shows whether the LULCT had a positive, negative or no effect on a given response variable. Small trees = DAP < 10 cm, large trees = \overline{D} AP ≥ 10 cm. UF: Undisturbed primary forest, LF: Logged primary forest, LBF: Logged-and-burned primary forest, SFo: Old secondary forest, SFy: Young secondary forest, PA: Pasture, MA: Mechanized agriculture.

SI References

- 1. A. Zanne, *et al.*, Data from: Towards a worldwide wood economics spectrum. *Dryad Digit. Repos.* (2009) https:/doi.org/https://doi.org/10.5061/dryad.234.
- 2. J. Chave, *et al.*, Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia* **145**, 87–99 (2005).
- 3. B. W. Nelson, *et al.*, Allometric regressions for improved estimate of secondary forest biomass in the central Amazon. *For. Ecol. Manage.* **117**, 149–167 (1999).
- 4. J. J. Gerwing, L. Farias, Integrating liana abundance and forest stature into an estimate of total aboveground biomass for an eastern Amazonian forest. *J. Trop. Ecol.* **16**, 327– 335 (2000).
- 5. T. A. Gardner, *et al.*, A social and ecological assessment of tropical land uses at multiple scales: the Sustainable Amazon Network. *Philos. Trans. R. Soc. B Biol. Sci.* **368**, 20120166 (2013).
- 6. E. Berenguer, *et al.*, A large-scale field assessment of carbon stocks in human-modified tropical forests. *Glob. Chang. Biol.* **20**, 3713–3726 (2014).
- 7. A. C. Lees, *et al.*, One hundred and thirty-five years of avifaunal surveys around Santarém, central Brazilian Amazon. *Rev. Bras. Ornitol.* **21**, 16–57 (2013).
- 8. A. C. Lees, *et al.*, Paragominas: A quantitative baseline inventory of an eastern Amazonian avifauna. *Rev. Bras. Ornitol.* **20**, 93–118 (2012).
- 9. R. R. de C. Solar, J. C. M. Chaul, M. Maués, J. H. Schoereder, A Quantitative Baseline of Ants and Orchid Bees in Human-Modified Amazonian Landscapes in Paragominas, PA, Brazil. *Sociobiology* **63**, 925 (2016).
- 10. R. C. Team, R: A language and environment for statistical computing (2021).
- 11. M. Fasiolo, S. N. Wood, M. Zaffran, R. Nedellec, Y. Goude, qgam : Bayesian Nonparametric Quantile Regression Modeling in R. *J. Stat. Softw.* **100**, 1–31 (2021).
- 12. T. Hothorn, F. Bretz, P. Westfall, Simultaneous inference in general parametric models. *Biometrical J.* **50**, 346–363 (2008).
- 13. C. M. Souza, *et al.*, Reconstructing Three Decades of Land Use and Land Cover Changes in Brazilian Biomes with Landsat Archive and Earth Engine. *Remote Sens.* **12**, 2735 (2020).
- 14. E. A. T. Matricardi, *et al.*, Long-term forest degradation surpasses deforestation in the Brazilian Amazon. *Science (80-.).* **369**, 1378–1382 (2020).
- 15. F. Hartig, DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models. R package version 0.4.5. (2022).