Accounting for spatial regional variability in modelling and forecasting deforestation – the fate of Madagascar's forests

Ghislain VIEILLEDENT[1,2,3,4,⋆] and Frédéric ACHARD[1]

[1] **European Commission** – JRC, Bio-economy Unit, I-21027 Ispra (VA), ITALY
[2] **CIRAD** – UPR Forêts et Sociétés, F-34398 Montpellier, FRANCE
[3] **CIRAD** – UPR AMAP, F-34398 Montpellier, FRANCE
[4] **Univ Montpellier** – AMAP, CIRAD, Montpellier, FRANCE

[⋆] **Corresponding author:** \E-mail: ghislain.vieilledent@cirad.fr \Phone: +33 4 67 61 49 09

Preambule

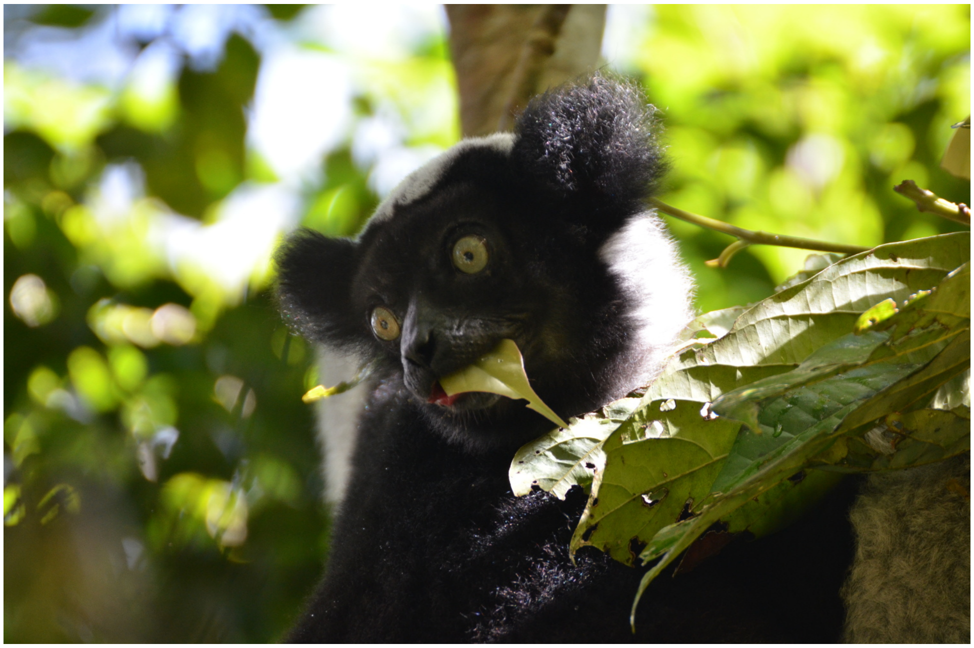

Figure 1: ***Indri indri* lemur speices in the Mantadia National Park in the eastern humid forest of Madagascar.** This species is highly threatened by habitat loss associated to deforestation.

Abstract

Deforestation models are useful tools in landscape ecology. They can be used to identify the main drivers of deforestation and estimate their relative contribution. When spatially explicit, models can also be used to predict the location of future deforestation. Deforestation forecasts can be used to estimate associated CO_2 emissions responsible of climate change, prioritize areas for conservation and identify refuge areas for biodiversity. Most of spatial deforestation models includes landscape variables such as the distance to forest edge, the distance to nearest road or the presence of protected areas. Such variables commonly explain a small part of the deforestation process and a large spatial variability remains unexplained by the model.

In the present study, we show how using an intrinsic conditional autoregressive (iCAR) model in a hierarchical Bayesian approach can help structure the residual spatial variability in the

deforestation process and obtain more realistic predictions of the deforestation (Dormann et al. 2007). We take Madagascar as a case study considering deforestation on the period 1990-2010 and forecasting deforestation on the period 2010-2050. We demonstrate that accounting for spatial autocorrelation increases the percentage of explained deviance of 21 points for the deforestation model in Madagascar. We also illustrate the use of the newly developed `forestatrisk` Python module to rapidly estimate the numerous parameters of a deforestation model including an iCAR process and efficiently forecast deforestation on large geographical areas at high spatial resolution.

We advocate the use of such models to obtain more accurate land-use change predictions. Such an approach could be used to estimate better the impact of future deforestation in the global carbon cycle and define more efficient strategies for biodiversity conservation in tropical countries.

Introduction

Madagascar est une île de l'océan indien qui s'est détachée de l'Afrique puis de l'Inde il y a plusieurs dizaines de millions d'années. La biodiversité y a évolué de façon isolée. Ainsi, près de 90% des espèces n'existent que sur l'île. Cette biodiversité est fortement menacée par les changements climatiques et la déforestation. Le projet BioSceneMada, en s'appuyant sur des modèles climatiques, écologiques et paysagers, se propose d'établir des scénarios d'évolution de la biodiversité à Madagascar sous l'effet conjoint du changement climatique et de la déforestation. Les cartes issues du projet permettent d'identifier les zones à risque de perte de biodiversité et de déforestation ainsi que les zones refuges pour la biodiversité. Ces cartes peuvent être utilisées afin d'agir efficacement pour la conservation des forêts et de la biodiversité à Madagascar, notamment en priorisant les actions de conservation sur des zones cibles, tout en s'appuyant sur le réseau d'aires protégées actuel.

Figure 2: **Population de baobabs de l'espèce *Adansonia suarezensis* sur le site de la Montagne des Français au nord de Madagascar.** Cette espèce est fortement menacée par la perte d'habitat associée aux changements climatiques.

Materials and Methods

Data

We used historical deforestation maps for Madagascar at 30m resolution for three time-periods: 1990-2000, 2000-2010, and 2010-2017 (Vieilledent *et al.*, 2018). We tried to model the observed spatial deforestation process on the period 2000-2010 at the national level. Period 1990-2000 was used to compute the distance to past deforestation for each forest pixel in 2000. Period 2010-2017 was used to compare model forecasts with observations.

To explain the observed spatial deforestation on the period 2000-2010, we considered various spatial explicative variables describing: topography (altitude and slope), accessibility (distances to nearest road, town and river), forest landscape (distance to forest edge), deforestation history (distance to past deforestation) and land-tenure variables (protected area system). Characteristics of each variables are summarized in Tab. 1.

Altitude (in m) and slope (in degree) at 90m resolution were obtained from the SRTM Digital Elevation Database v4.1 (http://srtm.csi.cgiar.org/). Distances (in m) to nearest road, town and river at 150m resolution were derived from the OpenStreetMap (OSM) project for Madagascar (http://www.geofabrik.de/). To obtain the road network in Madagascar, we considered the categories "motorway", "trunk", "primary", "secondary" and "tertiary" for the "highway" key in OSM. To obtain the network of populated places in Madagascar (that we simply call "towns" in the present study), we considered the categories "city", "town" and "village" for the "place" key in OSM. To obtain the river network in Madagascar, we considered the categories "river" and "canal" for the "waterway" key in OSM. For a more detailed description of each category, see the OSM wiki page (https://wiki.openstreetmap.org/wiki/Tags). Distance to forest edge was computed at 30m resolution from the forest cover map in 2000. Distance to past deforestation was computed at 30m resolution from the 1990-2000 forest cover change map. For the protected area system, we used the 20/12/2010 version of the SAPM "Système des Aires Protégées à Madagascar" (http://rebioma.net/) and considered both Protected Areas (created before 2003) and New Protected Areas in the SAPM terminology. Polygons representing protected areas were rasterized at 30m resolution. In total, we obtained 8 spatial variables that could explained the location of the deforestation.

Models

We compared two deforestation models. The first model is a simple logistic regression model. We consider the random variable y_i which takes value 1 if the forest pixel i is deforested on the period 2000-2010 and 0 if it is not. We assumed that y_i follows a Bernoulli distribution of parameter θ_i. In our moded, θ_i represents the spatial probability of deforestation for pixel i. We assumed that θ_i is linked, through a logit function, to a linear combination of the explicative variables $X_i\beta$, where X_i is the vector of explicative variables for pixel i and β are the model's parameters to be estimated. The first model is summarized in Equation (1).

$$\begin{aligned} y_i &\sim \mathcal{B}ernoulli(\theta_i) \\ \text{logit}(\theta_i) &= X_i\beta \end{aligned} \tag{1}$$

The second model (Eq. (2)) includes a spatial random effect ρ_j for each spatial cell of a 10×10 km grid over Madagascar.

$$\begin{aligned} y_i &\sim \mathcal{B}ernoulli(\theta_i) \\ \text{logit}(\theta_i) &= X_i\beta + \rho_j \\ \rho_j &\sim \mathcal{N}ormal(\mu_j, V_\rho/n_j) \end{aligned} \tag{2}$$

We used a sample of 20,000 forest pixels in 2000 to fit the two models. The sample was stratified between 10,000 deforested pixels in 2000-2010 and 10,000 non-deforested pixels. A balanced data-set is preferable in our case as we are not interested in estimating the intensity of deforestation but only the relative effect of each variable on the spatial probability of deforestation (Dezécache *et al.*, 2017).

Parameter inference was done in a hierarchical Bayesian framework. Non-informative priors were used for all parameters: $\beta \sim Normal(\text{mean} = 0, \text{var} = 10^6)$ and $V_\rho \sim 1/Gamma(\text{shape} = 0.05, \text{rate} = 0.0005)$. We used 2000 iterations for the burnin phase.

We run a Markov Chain Monte Carlo (MCMC) of 5000 iterations, thining the chain at each 5 iterations. We obtained 1000 estimates for each parameter. MCMC convergence was visually checked looking a the trace and parameter posterior distributions. Function `model_binomial_iCAR()` from the `forestatrisk` Python package was used for parameter inference. This function calls an adaptive Metropolis-within-Gibbs algorithm (Rosenthal *et al.*, 2011) written in C for maximum computation speed.

Model comparison through cross-validation

Forecasting deforestation

Forecasts skill scores

Results

Discussion

Ackowledgements

Tables

Table 1: **Bases de données sur la biodiversité de Madagascar.** La base de données incluent des points de présence pour 4969 espèces réparties dans différents groupes taxonomiques. Ces espèces sont représentatives de la biodiversité à Madagascar.

	Espèces	Genres	Obs.
Plantes			
Arbres	531	283	40178
Palmiers	178	16	5105
Fougères	317	82	1664
Légumineuses	724	149	30305
Graminées	283	113	3469
Autres	1229	359	34265
Vertébrés			
Mammifères	189	69	28316
Lémuriens	64	15	3136
Oiseaux	285	172	60895
Reptiles	153	41	4938
Amphibiens	78	21	208
Invertébrés			
Escargots	537	113	1635
Fourmis	379	46	70012
Papillons	262	82	16396
Autres	355	203	6202
TOTAL=	4969	1749	303588

Figures

References

Dezécache, C., Salles, J.M., Vieilledent, G. & Hérault, B. (2017) Moving forward socio-economically focused models of deforestation. *Global Change Biology*, **23**, 3484–3500. URL https://doi.org/10.1111/gcb.13611

Rosenthal, J.S. *et al.* (2011) Optimal proposal distributions and adaptive MCMC. *Handbook of Markov Chain Monte Carlo*, **4**.

Vieilledent, G., Grinand, C., Rakotomalala, F.A., Ranaivosoa, R., Rakotoarijaona, J.R., Allnutt, T.F. & Achard, F. (2018) Combining global tree cover loss data with historical national forest cover maps to look at six decades of deforestation and forest fragmentation in Madagascar. *Biological Conservation*, **222**, 189–197. URL https://doi.org/10.1016/j.biocon.2018.04.008