Special issue CARI'12

Markov analysis of land use dynamics

A Case Study in Madagascar

Fabien Campillo[1] — Dominique Hervé[2] — Angelo Raherinirina[3] — Rivo Rakotozafy[3]

[1] Inria, Modemic Team Projet (Inria/Inra), 2, place Pierre Viala, 34060 Montpellier, France
Fabien.Campillo@inria.fr

[2] IRD, MEM (IRD/University of Fianarantsoa),
UMR220 GRED (IRD/Montpellier3), Montpellier, France
dominique.herve@ird.fr

[3] University of Fianarantsoa, BP 1264, Andrainjato, 301 Fianarantsoa, Madagascar
angelo_raherinirina@yahoo.fr
rrakotozafy@uni-fianar.mg

ABSTRACT. We present a Markov model of a land-use dynamic along a forest corridor of Madagascar. A first approach by the maximum likelihood approach leads to a model with an absorbing state. We study the quasi-stationary distribution law of the model and the law of the hitting time of the absorbing state. According to experts, a transition not present in the data must be added to the model: this is not possible by the maximum likelihood method and we make of the Bayesian approach. We use a Markov chain Monte Carlo method to infer the transition matrix which in this case admits an invariant distribution law. Finally we analyze the two identified dynamics.

RÉSUMÉ. Nous présentons un modèle de Markov d'une dynamique d'utilisation des sols le long d'un corridor forestier de Madagascar. Une première approche par maximum de vraisemblance conduit à un modèle avec un état absorbant. Nous étudions la loi de probabilité quasi-stationnaire du modèle et la loi du temps d'atteinte de l'état absorbant. Selon les experts, une transition qui n'est pas présente dans les données doit néanmoins être ajoutée au modèle: ceci n'est pas possible par la méthode du maximum de vraisemblance et nous devons faire appel à une approche bayésienne. Nous faisons appel à une technique d'approximation de Monte Carlo par chaîne de Markov pour identifier la matrice de transition qui dans ce cas admet une loi de probabilité invariante. Enfin nous analysons les deux dynamiques ainsi identifiés.

KEYWORDS : Bayesian inference; Jeffreys prior; Land use dynamics; Markov model; Markov chain Monte Carlo; Quasi stationary distribution law.

MOTS-CLÉS : Inférence bayésienne; loi a priori de Jeffreys; dynamique d'usage des sols; modèles de Markov; Monte Carlo par Chaînes de Markov; loi quasi-stationnaire.

1. Introduction

Population pressure is one of the major causes of deforestation in tropical countries. In the region of Fianarantsoa (Madagascar), two national parks Ranomafana and Andringitra are connected by a forest corridor, which is of critical importance in maintaining the regional biodiversity [1, 2, 5], see Figure 1. The need for cultivated land pushes people to encroach on the corridor to look for lowlands to be converted into paddy fields, and then to clear slope forested parcels for cultivation. Once lowlands are all converted in paddy fields, the dynamic of slash and burn cultivation is clearly opposed to the dynamic of forest conservation and regeneration. To reconcile forest conservation with agricultural production, it is important to understand and model the dynamic of post-forest land use of these parcels.

This specific region of Madagascar has already motivated studies in terms of modeling. Notably, [1, 2] propose a hierarchical Bayesian model that relates demographic data with satellite images in order to analyze the links between population pressure and deforestation at the regional scale.

In contrast to previous models, we will work at the smaller scale of the township with a land-use data set of 42 plots over 22 years; and our model will not capture spatial dependencies. Indeed, the first treatment performed on such a dataset is to build the empirical transition matrix, which is equivalent to assume first that the dynamics of the plots are independent and second that the plot follows an homogeneous Markov dynamic of order 1. We will adopt this simple Markov framework and in this preliminary work we aim at inferring the time scale at which the dynamics of the agricultural producers stabilizes.

We will use a data set developed by IRD, at the western edge of the corridor, consisting of yearly transitions of 42 parcels initially in forest, these parcels have been cleared since. This set of data is presented in Table 1. Each parcel can take three possible states: annual crop (`A`), fallow (`F`), perennial crop (`P`).

The use of Markovian approaches to model land-use transitions and vegetation successions is widespread [19, 20, 22]. The success of these approaches is explained by the fact that agro-ecological dynamics may be represented as discrete-time successions on a finite number of states, each one with its sojourn time. Both agronomists and ecologists, in dialog, actually fail in predicting the future succession of these states, knowing the previous land-use history. They ask the mathematicians for detailing the characteristics of these dynamics and defining how to pilot them. The construction, manipulation and simulation of such models are fairly easy. The transition probabilities of the Markovian model are estimated from observed data. The classical reference [3] proposes the maximum likelihood method to estimate the transition probabilities of a Markov chain. An alternative is to consider Bayesian estimators [18, 13]. In this paper we explore and test several modeling tools, Markov chain, Bayesian estimation and MCMC procedure in order to find a better fit with the data. These results can then be utilized by agro-ecologists to model land-use dynamics at a parcel scale.

In Section 2, a first model is directly derived from the data with the help of the empirical transition matrix which also corresponds to the maximum likelihood estimate. This first model is not entirely satisfactory as one of the states is absorbing and does not correspond to current knowledge gained by agronomists in the field. Hence, in Section 3, a

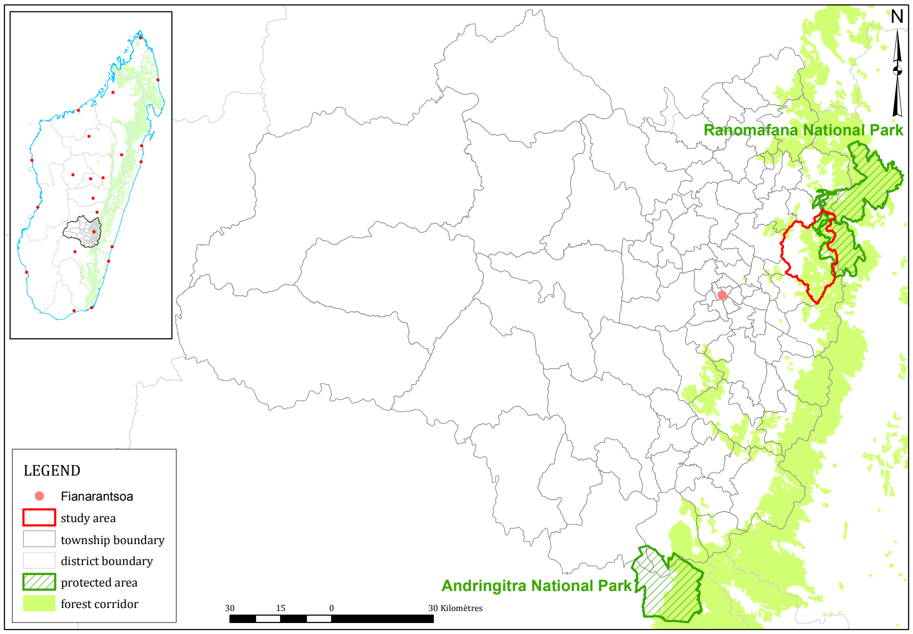


Figure 1. *The study area bordering the forest corridor between Ranomafana and Andringitra.*

second model is proposed in a Bayesian context and we use a Monte Carlo Markov Chain (MCMC) method to infer the transition matrix. This transition matrix appears to be more realistic. The model is evaluated in Section 4 and discussed in Section 5. Conclusion and perspectives are drawn in Section 6.

2. A first model derived from the maximum likelihood estimate

Consider the observation series of Table 1 of the annual states corresponding to 42 parcels over 22 years, from 1985 to 2006. The states are: annual crop ($\mathtt{A}$), fallow ($\mathtt{F}$), perennial crop ($\mathtt{P}$) and natural forest ($\mathtt{f}$). The transition from forest, $\mathtt{f}$ state, to annual crop, $\mathtt{A}$ state, is the slash and burn first forest clearing. The $\mathtt{f}$ state is systematically followed by the $\mathtt{A}$ state and none of the parcel comes back to the $\mathtt{f}$ state. So we can omit the $\mathtt{f}$ state from the series of observations and we assume that $\mathtt{A}$ is the initial state of each parcel. Note that it is possible because the dynamics of parcels are assumed independent.

The observation series will be denoted $(e^{(p)}_{1:N_p})_{p=1:42}$ where $e^{(p)}_n$ belongs to the state space:

$$E \stackrel{\text{def}}{=} \{\mathtt{A}, \mathtt{F}, \mathtt{P}\}$$

and N_p is the length of the observation series of the parcel p. Here the notation $n = n_1 : n_2$ stands for $n = n_1, \ldots, n_2$ for $n_1 \leq n_2$.

We make the following hypothesis:

parcel number \ year	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
42	f	A	A	A	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P
41	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	A	A	A	A	A	A
40	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	A	A	A	A	A	A
39	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	A	F	A	A	A	F
38	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	A	F	F	A	F	A
37	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	A	A	A	A	A	A
36	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	A	A	A	A	A	F
35	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	A	A	A	A	F	A
34	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	A	A	A	F	A	A
33	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	A	A	A	A	F	A
32	f	f	f	A	F	A	A	F	F	F	A	A	F	F	F	A	A	F	F	F	A	A
31	f	f	f	A	A	F	A	A	F	F	A	A	A	F	A	A	A	F	F	A	A	F
30	f	f	f	A	A	F	F	A	F	F	F	F	A	A	F	F	F	F	F	F	A	A
29	f	f	f	A	A	A	F	F	A	A	A	F	F	A	F	F	F	A	A	A	A	A
28	f	f	f	A	A	A	A	A	F	F	A	A	A	A	A	F	F	A	A	F	F	F
27	f	f	f	A	A	A	A	A	A	F	F	F	A	A	A	A	A	F	F	A	A	A
26	f	f	f	A	A	A	F	F	F	F	F	F	F	F	F	F	F	A	A	F	A	A
25	f	f	f	A	A	A	A	A	A	F	F	F	F	F	F	F	F	A	A	F	F	A
24	f	f	f	A	A	A	A	A	A	F	F	F	F	F	F	A	A	F	F	F	A	F
23	f	f	f	f	f	f	f	f	f	f	f	A	A	A	A	A	F	F	A	A	A	F
22	f	f	f	f	f	f	f	f	f	f	f	A	A	A	A	A	F	A	A	A	F	A
21	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	F	A	A	A	A	F
20	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	A	A	F	F	F	F
19	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	F	F	F	F	F	F
18	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	A	A	A	F	F	F
17	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	F	F	F	F	A	F
16	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	F	F	F	F	A	A
15	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	F	F	F	F	A	A
14	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	A	A	A	A	F	F
13	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	A	A	A	A	F	F
12	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	A	F	F	F	F	F
11	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	A	A	A	A	A	F	A
10	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	A	A	F	A	A	F	A
9	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	A	A	A	A	A	F	F
8	f	f	f	f	f	f	f	f	f	f	f	f	f	A	A	A	F	F	A	A	A	A
7	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	A	A	A	P	P	P	P
6	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	A	A	A	A	A	F	A
5	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	A	A	A	A	A	A	A
4	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	A	A	A	A	A
3	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	A	A	A	A	A
2	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	A	A	A	P	P
1	f	f	f	f	f	f	f	f	f	f	f	f	f	f	f	A	A	F	F	F	A	A

Table 1. *Annual states corresponding to 42 parcels and 22 years. These data were collected by D. Hervé between the years 1985 and 2006. The parcels are located on the slopes and lowlands on the edge of the forest corridor of Ranomafana-Andringitra, Madagascar, see Figure 1. The states are: annual crop* (`A`)*, fallow* (`F`)*, perennial crop* (`P`) *and natural forest* (`f`)*. The* `f` *state, which will be omitted, is systemically followed by the* `A` *state that will be considered as the first observation. Hence the observation series* $(e^{(p)}_{n=1:N_p})_{p=1:42}$ *will be the sequence of states* $\{\mathtt{A},\mathtt{F},\mathtt{P}\}$ *and* N_p *will be the length of the observation series of the parcel* p*, e.g.* $N_1 = 7$*,* $N_2 = 6$ *etc.*

(H_1) *The dynamics of the parcels are independent and identical, they are Markovian and time-homogeneous.*

This means that $(e^{(p)}_{1:N_p})_{p=1:42}$, are 42 independent realizations of a same process $(X_n)_{n\geq 0}$ and that this process is Markovian and time-homogeneous.

This assumption is of course simplistic as the dynamics of a given parcel may depend on: farmer decisions; exposition, slope and distance from the forest; neighboring parcels. Assumption (H_1) is rather restrictive and unrealistic, nevertheless in the present context it permits us to infer some interesting results.

This hypothesis leads to a model $X = (X^{(p)}_{1:N_p})_{p=1:P}$, where $(X^{(p)}_n)_{n=1:N_p}$ are P independent Markov chains, with initial law δ_A and transition matrix Q of size 3×3:

$$\begin{aligned} \mathbb{P}(X^{(p)}_{1:N_p} = e^{(p)}_{1:N_p}\,,\ \forall p = 1, \ldots, P) &= \prod_{p=1}^{P} \mathbb{P}(X^{(p)}_{1:N_p} = e_{1:N_p}) \\ &= \prod_{p=1}^{P} \delta_{\mathtt{A}}\,(e^{(p)}_0)\; Q(e^{(p)}_0, e^{(p)}_1) \cdots Q(e^{(p)}_{N-2}, e^{(p)}_{N-1}) \end{aligned} \tag{1}$$

for all $e^{(p)}_n \in E$, where $Q = [Q(e, e')]_{e,e' \in \{\mathtt{A},\mathtt{F},\mathtt{P}\}}$.

2.1. Maximum likelihood estimate

We derive the maximum likelihood estimate of the transition matrix:

$$Q = [Q(i,j)]_{i,j\in E} = \begin{pmatrix} 1-\theta_1-\theta_2 & \theta_1 & \theta_2 \\ \theta_3 & 1-\theta_3-\theta_4 & \theta_4 \\ \theta_5 & \theta_6 & 1-\theta_5-\theta_6 \end{pmatrix} \tag{2}$$

where $\theta = (\theta_i)_{1\leq i\leq 6} \in \Theta$ with $\Theta \stackrel{\text{def}}{=} \big\{\theta \in [0,1]^6\,;\ \theta_1+\theta_2 \leq 1,\, \theta_3+\theta_4 \leq 1,\, \theta_5+\theta_6 \leq 1\big\}$. Let $\mathbb{P}_\theta$ denote the probability under which the Markov chain admit Q with parameter θ as a transition matrix. We applied the classical results of [3] to compute the MLE of the matrix Q. We consider the likelihood function $L(\theta|(e^{(p)}_{1:N_p})_{p=1:P})$ of θ given the data $(e^{(p)}_{1:N_p})_{p=1:P}$, for the sake of simplicity we denote it $L_1(\theta)$, hence:

$$\begin{aligned} L_1(\theta) &\stackrel{\text{def}}{=} \mathbb{P}_\theta(X^{(p)}_{1:N_p} = e^{(p)}_{1:N_p},\, \forall p = 1, \ldots, P) \\ &= \prod_{p=1}^{P} \delta_{\mathtt{A}}\,(e^{(p)}_1)\; Q(e^{(p)}_1, e^{(p)}_2) \cdots Q(e^{(p)}_{N_p-1}, e^{(p)}_{N_p}) \end{aligned}$$

for any $(e^{(p)}_{1:N_p})_{p=1:P} \in \mathcal{E} = \prod_{p=1}^{P} E^{N_p}$. Let $n^{(p)}_{ee'}$ be the number of transitions from state e to state e' for a parcel p in X:

$$n^{(p)}_{ee'} \stackrel{\text{def}}{=} \sum_{n=1}^{N-1} 1_{\{X^{(p)}_{n-1}=e\}}\, 1_{\{X^{(p)}_n=e'\}} \qquad \forall e, e' \in E \tag{3}$$

and $n_{ee'}$ be the total of number of transitions from state e to state e' ($e, e' \in E$):

$$n_{ee'} \stackrel{\text{def}}{=} \sum_{p=1}^{P} n_{ee'}^{(p)} \,. \tag{4}$$

According to (2) and (4):

$$\begin{aligned}
L_1(\theta) = \prod_{p=1}^{P} \big\{ & Q(1,1)^{n_{\mathtt{AA}}^{(p)}}\, Q(1,2)^{n_{\mathtt{AF}}^{(p)}}\, Q(1,3)^{n_{\mathtt{AP}}^{(p)}}\, Q(2,1)^{n_{\mathtt{FA}}^{(p)}}\, Q(2,2)^{n_{\mathtt{FF}}^{(p)}}\, Q(2,3)^{n_{\mathtt{FP}}^{(p)}} \\
& \qquad Q(3,1)^{n_{\mathtt{PA}}^{(p)}}\, Q(3,2)^{n_{\mathtt{PF}}^{(p)}}\, Q(3,3)^{n_{\mathtt{PP}}^{(p)}} \big\} \\
= \prod_{p=1}^{P} \big\{ & (1-\theta_1-\theta_2)^{n_{\mathtt{AA}}^{(p)}}\, \theta_1^{n_{\mathtt{AF}}^{(p)}}\, \theta_2^{n_{\mathtt{AP}}^{(p)}}\, \theta_3^{n_{\mathtt{FA}}^{(p)}}\, (1-\theta_3-\theta_4)^{n_{\mathtt{FF}}^{(p)}}\, \theta_4^{n_{\mathtt{FP}}^{(p)}} \\
& \qquad \theta_5^{n_{\mathtt{PA}}^{(p)}}\, \theta_6^{n_{\mathtt{PF}}^{(p)}}\, (1-\theta_5-\theta_6)^{n_{\mathtt{PP}}^{(p)}} \big\} \\
= (1-\theta_1 & -\theta_2)^{n_{\mathtt{AA}}}\, \theta_1^{n_{\mathtt{AF}}}\, \theta_2^{n_{\mathtt{AP}}}\, \theta_3^{n_{\mathtt{FA}}}\, (1-\theta_3-\theta_4)^{n_{\mathtt{FF}}}\, \theta_4^{n_{\mathtt{FP}}}\, \theta_5^{n_{\mathtt{PA}}}\, \theta_6^{n_{\mathtt{PF}}}\, (1-\theta_5-\theta_6)^{n_{\mathtt{PP}}}
\end{aligned}$$

so that the log-likelihood function is:

$$\begin{aligned}
l_1(\theta) = n_{\mathtt{AA}} \log(1-\theta_1-\theta_2) &+ n_{\mathtt{AF}} \log(\theta_1) + n_{\mathtt{AP}} \log(\theta_2) \\
&+ n_{\mathtt{FA}} \log(\theta_3) + n_{\mathtt{FF}} \log(1-\theta_3-\theta_4) + n_{\mathtt{FP}} \log(\theta_4) \\
&+ n_{\mathtt{PA}} \log(\theta_5) + n_{\mathtt{PF}} \log(\theta_6) + n_{\mathtt{PP}} \log(1-\theta_5-\theta_6) \,.
\end{aligned} \tag{5}$$

The MLE $\hat\theta$ is solution of $\partial l_1(\theta)/\partial\theta|_{\theta=\hat\theta} = 0$, that leads to:

$$\begin{aligned}
\hat\theta_1 &= n_{\mathtt{AF}}/(n_{\mathtt{AA}}+n_{\mathtt{AF}}+n_{\mathtt{AP}})\,, & \hat\theta_2 &= n_{\mathtt{AP}}/(n_{\mathtt{AA}}+n_{\mathtt{AF}}+n_{\mathtt{AP}})\,, \\
\hat\theta_3 &= n_{\mathtt{FA}}/(n_{\mathtt{FA}}+n_{\mathtt{FF}}+n_{\mathtt{FP}})\,, & \hat\theta_4 &= n_{\mathtt{FP}}/(n_{\mathtt{FA}}+n_{\mathtt{FF}}+n_{\mathtt{FP}})\,, \\
\hat\theta_5 &= n_{\mathtt{PA}}/(n_{\mathtt{PA}}+n_{\mathtt{PF}}+n_{\mathtt{PP}})\,, & \hat\theta_6 &= n_{\mathtt{PF}}/(n_{\mathtt{PA}}+n_{\mathtt{PF}}+n_{\mathtt{PP}})\,.
\end{aligned}$$

The resulting empirical transition matrix is:

$$\hat Q \stackrel{\text{def}}{=} \begin{pmatrix} 0.7447699 & 0.2426778 & 0.0125523 \\ 0.3385827 & 0.6614173 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{6}$$

it leads to the model depicted in Figure 8 (left). The main implication is that the state P is absorbing and the other states are transient so that the limit distribution is $\delta_{\mathtt{P}}$. A first question is to describe the behavior of the process before it reaches P and then to analyze the time taken to reach the state P.

2.2. Quasi-stationary distribution

The answer to the first question is given by the so-called quasi-stationary distribution, see Appendix C. The quasi-stationary distribution is roughly the "limit" distribution that the system reaches before being absorbed by P. The quasi-stationary distribution σ_{qs} associated with the maximum likelihood estimator of Q is described in Appendix C. The result is:

$$\sigma_{\text{qs}} = (\sigma_{\text{qs}}(\mathtt{A}), \sigma_{\text{qs}}(\mathtt{F})) = (0.5794, 0.4206) \,. \tag{7}$$

Hence, conditionally on the fact that the process does not reach P, it will spend 58% of its time in the A state and 42% of its time in the F state.

2.3. Distribution of the time to reach P

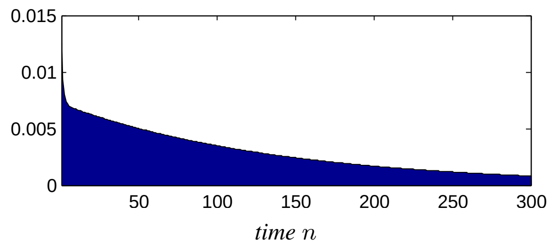


Figure 2. *Distribution of the time to reach state* P*. The mean time is 134 years and the standard deviation is 130 years, this standard deviation is rather high as the time to reach* P *is exponentially distributed.*

Define the first time to reach a given state e by:

$$\tau_e \stackrel{\text{def}}{=} \inf\{n \geq 1 \;:\; X_n = e\} \;\text{ if } \exists n \geq 1 \text{ such that } X_n = e,\; +\infty \text{ otherwise}\,. \tag{8}$$

We now compute the distribution law:

$$\mathbb{P}(\tau_{\mathtt{P}} = n | X_0 = \mathtt{A}) = \mathbb{P}(X_n = \mathtt{P},\, X_m \neq \mathtt{P},\, m = 1, \ldots, n-1 | X_0 = \mathtt{A})$$

of the first time to reach the absorbing state P. The derivation of this distribution is made in Appendix B and leads to the recurrence formula (24). The result is plotted in Figure 2. The mean time to reach P is 136 years with a standard deviation of 135 years, this rather high standard deviation corresponds to the fact that the time to reach P is exponentially distributed.

3. A second model

The model derived from the maximum likelihood estimate does not allow the transitions PA, FP and PF. The transitions FP and PF are not realistic in the geographical context of this study, but the transition PA can be observed on a time scale of several decades. So we suppose that:

(H_2) *The transitions* FP *and* PF *are not possible, all other transitions are possible.*

This hypothesis leads to the following transition matrix:

$$Q = \begin{pmatrix} 1-\theta_1-\theta_2 & \theta_1 & \theta_2 \\ \theta_3 & 1-\theta_3 & 0 \\ \theta_4 & 0 & 1-\theta_4 \end{pmatrix}, \tag{9}$$

with $\theta = (\theta_1, \theta_2, \theta_3, \theta_4) \in \Theta = \{\theta \in [0, 1]^4 ;\;\; \theta_1 + \theta_2 \leq 1\}$ that corresponds to the model depicted in Figure 3 where the probability of the transition PA will be small but

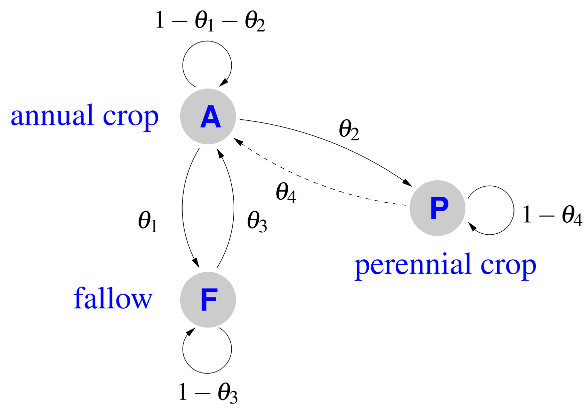


Figure 3. *Model 2. Three states regular Markov chain diagram: annual crop* (A)*, fallow* (F)*, perennial crop* (P)*;* A *is the initial state.*

positive. The maximum likelihood approach proposed in Section 2 will lead to the Model 1 of Figure 8 (left) where $Q(\mathtt{P},\mathtt{A}) = 0$. We therefore have to make use of Bayesian methods of estimation.

Given a prior distribution law $\pi(\theta)$ on the parameter θ, According to the Bayes rule, the posterior distribution law $\pi(\theta)$ on θ given the observations X is:

$$\pi_{\text{post}}\big(\theta\big|(e^{(p)}_{1:N_p})_{p=1:P}\big) \propto L\big(\theta\big|(e^{(p)}_{1:N_p})_{p=1:P}\big)\;\pi_{\text{prior}}(\theta) \tag{10}$$

where $L(\theta|(e^{(p)}_{1:N_p})_{p=1:P})$ is the likelihood function of the parameter θ given the observation $(e^{(p)}_{1:N_p})_{p=1:P}$. For the sake of simplicity, in the following $\pi_{\text{post}}(\theta|(e^{(p)}_{1:N_p})_{p=1:P})$ will be denoted $\pi_{\text{post}}(\theta)$, and $L(\theta|(e^{(p)}_{1:N_p})_{p=1:P})$ will be denoted $L_2(\theta)$ to emphasis that it is different from the previous likelihood function $L_1(\theta)$. Hence:

$$\begin{aligned}
L_2(\theta) &= \textstyle\prod_{p=1}^P \big\{Q(1,1)^{n^{(p)}_{\mathtt{AA}}}\,Q(1,2)^{n^{(p)}_{\mathtt{AF}}}\,Q(1,3)^{n^{(p)}_{\mathtt{AP}}}\,Q(2,1)^{n^{(p)}_{\mathtt{FA}}} \\
&\qquad\qquad Q(2,2)^{n^{(p)}_{\mathtt{FF}}}\,Q(3,1)^{n^{(p)}_{\mathtt{PA}}}\,Q(3,3)^{n^{(p)}_{\mathtt{PP}}}\big\} \\
&= \textstyle\prod_{p=1}^P \big\{(1-\theta_1-\theta_2)^{n^{(p)}_{\mathtt{AA}}}\,\theta_1^{n^{(p)}_{\mathtt{AF}}}\,\theta_2^{n^{(p)}_{\mathtt{AP}}}\,(1-\theta_3)^{n^{(p)}_{\mathtt{FF}}}\,\theta_3^{n^{(p)}_{\mathtt{FA}}}\,(1-\theta_4)^{n^{(p)}_{\mathtt{PP}}}\,\theta_4^{n^{(p)}_{\mathtt{PA}}}\big\} \\
&= (1-\theta_1-\theta_2)^{n_{\mathtt{AA}}}\,\theta_1^{n_{\mathtt{AF}}}\,\theta_2^{n_{\mathtt{AP}}}\,(1-\theta_3)^{n_{\mathtt{FF}}}\,\theta_3^{n_{\mathtt{FA}}}\,(1-\theta_4)^{n_{\mathtt{PP}}}\,\theta_4^{n_{\mathtt{PA}}},
\end{aligned}$$

and the log-likelihood function is:

$$\begin{aligned}
l_2(\theta) &= n_{\mathtt{AA}}\log(1-\theta_1-\theta_2) + n_{\mathtt{AF}}\log(\theta_1) + n_{\mathtt{AP}}\log(\theta_2) \\
&\qquad + n_{\mathtt{FF}}\log(1-\theta_3) + n_{\mathtt{FA}}\log(\theta_3) + n_{\mathtt{PP}}\log(1-\theta_4) + n_{\mathtt{PA}}\log(\theta_4).
\end{aligned} \tag{11}$$

The Bayes estimator $\tilde\theta$ of the parameter θ is the mean of the a posteriori distribution:

$$\tilde\theta \stackrel{\text{def}}{=} \int_\Theta \theta\,\pi_{\text{post}}(\theta)\,\mathrm{d}\theta = \frac{\int_\Theta \theta\,L_2(\theta)\,\pi_{\text{prior}}(\theta)\,\mathrm{d}\theta}{\int_\Theta L_2(\theta)\,\pi_{\text{prior}}(\theta)\,\mathrm{d}\theta}\,. \tag{12}$$

3.1. Jeffreys prior

Numerical tests that will be performed in Section 3.2.1 suggest that the Jeffreys prior is well adapted to the present situation. This prior distribution (non-informative) is defined by [12]:

$$\pi_{\text{prior}}(\theta) \propto \sqrt{\det[\mathcal{I}(\theta)]} \tag{13}$$

where $\mathcal{I}(\theta)$ is the Fisher information matrix given by:

$$\mathcal{I}(\theta) \stackrel{\text{def}}{=} \left[\mathbb{E}_\theta\Big(-\frac{\partial^2 l_2(\theta)}{\partial\theta_k\,\partial\theta_l}\Big)\right]_{1\le k,l\le 4}$$

and $l_2(\theta)$ is the log-likelihood function (11). Hence:

$$\mathcal{I}(\theta) = \begin{pmatrix} A_{1,2} & 0 & 0 \\ 0 & a_3 & 0 \\ 0 & 0 & a_4 \end{pmatrix}$$

with

$$A_{k,\ell} \stackrel{\text{def}}{=} -\mathbb{E}_\theta \begin{pmatrix} \frac{\partial^2 l_2(\theta)}{\partial^2\theta_k} & \frac{\partial^2 l_2(\theta)}{\partial\theta_k\partial\theta_\ell} \\ \frac{\partial^2 l_2(\theta)}{\partial\theta_k\partial\theta_\ell} & \frac{\partial^2 l_2(\theta)}{\partial^2\theta_\ell} \end{pmatrix}, \quad a_k \stackrel{\text{def}}{=} -\mathbb{E}_\theta\left(\frac{\partial^2 l_2(\theta)}{\partial^2\theta_k}\right).$$

So $\det\mathcal{I}(\theta) = \det A_{1,2} \times a_3 \times a_4$ and

$$\pi_{\text{prior}}(\theta) \propto \sqrt{\det A_{1,2} \times a_3 \times a_4}\,. \tag{14a}$$

According to (11):

$$\begin{aligned}
\frac{\partial^2 l_2(\theta)}{\partial^2\theta_1} &= -\left(\frac{n_{\mathtt{AA}}}{(1-\theta_1-\theta_2)^2} + \frac{n_{\mathtt{AF}}}{\theta_1^2}\right), & \frac{\partial^2 l_2(\theta)}{\partial\theta_1\,\partial\theta_2} &= -\frac{n_{\mathtt{AA}}}{(1-\theta_1-\theta_2)^2}, \\
\frac{\partial^2 l_2(\theta)}{\partial^2\theta_2} &= -\left(\frac{n_{\mathtt{AA}}}{(1-\theta_1-\theta_2)^2} + \frac{n_{\mathtt{AP}}}{\theta_2^2}\right), & \frac{\partial^2 l_2(\theta)}{\partial\theta_2\,\partial\theta_1} &= -\frac{n_{\mathtt{AA}}}{(1-\theta_1-\theta_2)^2},, \\
\frac{\partial^2 l_2(\theta)}{\partial^2\theta_3} &= -\left(\frac{n_{\mathtt{FF}}}{(1-\theta_3)^2} + \frac{n_{\mathtt{FA}}}{\theta_3^2}\right), & \frac{\partial^2 l_2(\theta)}{\partial^2\theta_4} &= -\left(\frac{n_{\mathtt{PP}}}{(1-\theta_4)^2} + \frac{n_{\mathtt{PA}}}{\theta_4^2}\right)
\end{aligned}$$

and

$$\det A_{1,2} = \left(\tfrac{\mathbb{E}_\theta[n_{\mathtt{AA}}]}{(1-\theta_1-\theta_2)^2} + \tfrac{\mathbb{E}_\theta[n_{\mathtt{AF}}]}{\theta_1^2}\right)\left(\tfrac{\mathbb{E}_\theta[n_{\mathtt{AA}}]}{(1-\theta_1-\theta_2)^2} + \tfrac{\mathbb{E}_\theta[n_{\mathtt{AP}}]}{\theta_2^2}\right) - \left(\tfrac{\mathbb{E}_\theta[n_{\mathtt{AA}}]}{(1-\theta_1-\theta_2)^2}\right)^2, \tag{14b}$$

$$a_3 = \tfrac{\mathbb{E}_\theta[n_{\mathtt{FF}}]}{(1-\theta_3)^2} + \tfrac{\mathbb{E}_\theta[n_{\mathtt{FA}}]}{\theta_3^2}, \tag{14c}$$

$$a_4 = \tfrac{\mathbb{E}_\theta[n_{\mathtt{PP}}]}{(1-\theta_4)^2} + \tfrac{\mathbb{E}_\theta[n_{\mathtt{PA}}]}{\theta_4^2}. \tag{14d}$$

From (3) and (4):

$$\begin{aligned}\mathbb{E}_\theta[n_{ee'}] &= \sum_{p=1}^{P}\sum_{n=2}^{N_p}\mathbb{P}_\theta(X_n^{(p)}=e', X_{n-1}^{(p)}=e)\\ &= \sum_{p=1}^{P}\sum_{n=2}^{N_p}\mathbb{P}_\theta(X_n^{(p)}=e'|X_{n-1}^{(p)}=e)\,\mathbb{P}_\theta(X_{n-1}=e)\\ &= Q(e,e')\sum_{p=1}^{P}\sum_{n=2}^{N_p}\mathbb{P}_\theta(X_{n-1}^{(p)}=e)\\ &= Q(e,e')\sum_{p=1}^{P}\sum_{n=2}^{N_p}\big(\delta_{\mathtt{A}}Q^{n-1}\big)_e = Q(e,e')\sum_{p=1}^{P}\sum_{n=2}^{N_p}[Q^{n-1}](\mathtt{A},e) \qquad \text{(14e)}\end{aligned}$$

for all $e, e' \in E$. Note that [4] proposed a more complex method to compute the Jeffreys prior distribution.

3.2. MCMC method

Although the Jeffreys prior distribution is explicit, we cannot compute analytically the corresponding Bayes estimator (12). We propose to use a Monte Carlo Markov chain (MCMC) method, namely a Metropolis-Hastings algorithm with a Gaussian proposal distribution, see the corresponding algorithm in Appendix D.

3.2.1. Simulation study

We consider the simpler two states case $E = \{0, 1\}$. It has no connection with the Markov model considered in the present work but it allows us to easily compare the following different prior distributions: (*i*) the uniform distribution; (*ii*) the beta distribution of parameter $(\frac{1}{2}, \frac{1}{2})$; (*iii*) the non-informative Jeffreys distribution.

The Bayesian estimator is explicit for the two first priors, see Appendix A, the MCMC method is used in the last case. We use a Monte Carlo method to compare the maximum likelihood estimate and the Bayesian estimates: we compute the empirical distribution of the norm of the error between the real transition matrix and its estimation. More precisely, we sample 1000 independent values $\theta^{(\ell)}$ of the parameter according to a uniform distribution on $\Theta = [0, 1]^2$, $\ell = 1, \dots, 1000$, and we let:

$$Q_{\theta^\ell} = \begin{pmatrix} \theta_1^{(\ell)} & 1-\theta_1^{(\ell)} \\ 1-\theta_2^{(\ell)} & \theta_2^{(\ell)} \end{pmatrix}.$$

We simulated data $(e_{0:19}^{(p)})_{p=1:42}$ according to Q_{θ^ℓ}. Then we compute the MLE $\hat\theta^{(\ell)}$ and the three Bayes estimators $\tilde\theta_{\mathrm{U}}^{(\ell)}$, $\tilde\theta_{\mathrm{B}}^{(\ell)}$, $\tilde\theta_{\mathrm{J}}^{(\ell)}$ with the uniform, beta and Jeffreys priors respectively. Next we compute the errors:

$$\begin{aligned}\hat\varepsilon^\ell &= \|Q_{\theta^{(\ell)}} - Q_{\hat\theta^{(\ell)}}\|\,, & \tilde\varepsilon_{\mathrm{U}}^\ell &= \|Q_{\theta_{\mathrm{U}}^{(\ell)}} - Q_{\tilde\theta_{\mathrm{U}}^{(\ell)}}\|\,,\\ \tilde\varepsilon_{\mathrm{B}}^\ell &= \|Q_{\theta_{\mathrm{B}}^{(\ell)}} - Q_{\tilde\theta_{\mathrm{B}}^{(\ell)}}\|\,, & \tilde\varepsilon_{\mathrm{J}}^\ell &= \|Q_{\theta_{\mathrm{J}}^{(\ell)}} - Q_{\tilde\theta_{\mathrm{J}}^{(\ell)}}\|\end{aligned} \qquad (15)$$

for the two different norms, the Frobenius norm:

$$\|A\|_F^2 \stackrel{\text{def}}{=} \mathrm{trace}(A^*A) \qquad (16)$$

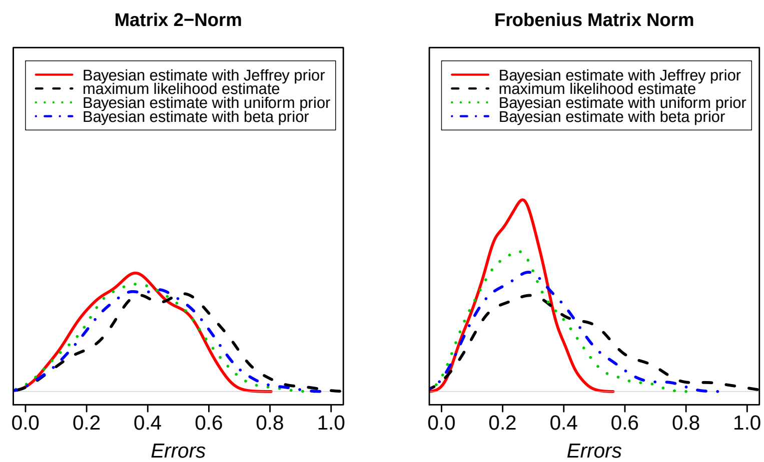


Figure 4. *Empirical PDF for the error terms* (15) *associated with the 2-norm* (16) *and the Frobenius norm* (17) *based on 1000 simulation of the parameter θ.*

and the 2-norm:

$$\|A\|_2 \stackrel{\text{def}}{=} \sqrt{\lambda_{\max}(A^*\,A)} \tag{17}$$

where $\lambda_{\max}(A^*\,A)$ is the largest eigenvalue of the matrix $A^*\,A$.

In Figure 4 we plotted the empirical distribution of the errors $\hat{\varepsilon}^\ell$ and $\tilde{\varepsilon}^\ell$, $\ell = 1, \dots, 1000$, for the two different norms. We see that the Bayes estimator with the Jeffreys prior give slightly better results for the 2-norm case and better results for the Frobenius norm. Hence the Jeffreys prior will be used with a MCMC method.

3.2.2. Application to the real data set

For the real data, the Bayes estimation of the transition matrix with the Jeffreys prior and the MCMC method is:

$$\tilde{Q} = \begin{pmatrix} 0.74216087 & 0.2410902 & 0.01674893 \\ 0.32459209 & 0.67540791 & 0 \\ 0.02460644 & 0 & 0.97539356 \end{pmatrix}. \tag{18}$$

As $[\tilde{Q}^2]_{ij} > 0$ for all i, j, $\tilde{Q}$ is a regular transition matrix (i.e. irreducible and aperiodic) and so there exists a unique invariant measure σ solution of:

$$\sigma = \sigma\,\tilde{Q} \tag{19}$$

indeed, according to the Perron–Frobenius theorem [16, p.3], the spectral radius of $\tilde{Q}$ is 1 and it is also an eigenvalue of $\tilde{Q}$ with multiplicity one, the corresponding eigenvector σ is strictly positive and could be chosen such that $\sum_{i\in E}\sigma(i) = 1$. After computation:

$$\sigma = (0.4127\,,\ 0.3065\,,\ 0.2808)\,. \tag{20}$$

We now study the rate of convergence toward the invariant measure σ following [10]. A first indication is given by the damping ratio $\alpha(\tilde{Q}) = \lambda_1/|\lambda_2|$, where $\lambda_1 = 1 > \lambda_2$ are

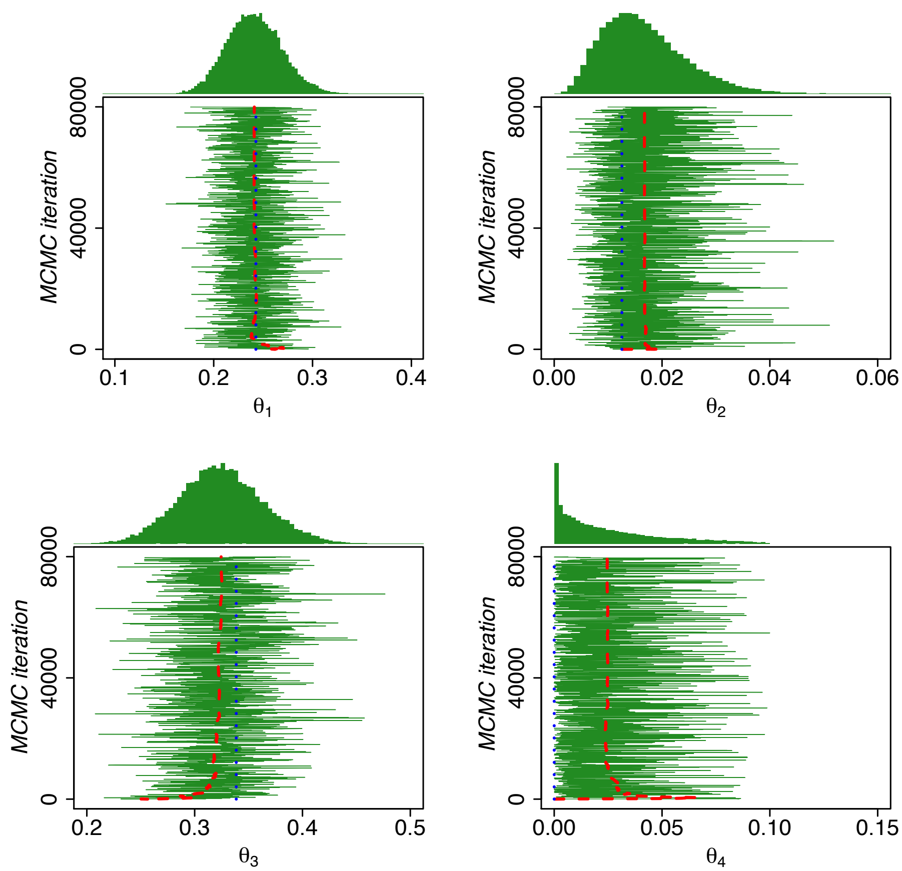


Figure 5. *Maximum likelihood and Bayes estimators for the real data set: posterior empirical distributions given by the MCMC iterations (—/green) and the associated mean (- - -/red) and the maximum likelihood estimates (· · · /blue).*

the two largest eigenvalues in modulus of $\tilde{Q}$, see [6] for details. The land-use dynamic converges to the equilibrium, in the long run, at least as fast as $\exp(-n\,\log\alpha(\tilde{Q})) = \exp(-0.036\,n)$ that is $3.6\%/\text{year}$.

An alternative to the damping ratio is the Dobrushin's coefficient defined as $\delta(\tilde{Q}) \stackrel{\text{def}}{=} \frac{1}{2}\,\max_{i,k\in E}\sum_{j\in\mathbb{E}}|P(i,j)-P(k,j)|$ for which we have the following inequality: $\|\mu_1\,\tilde{Q}^n - \mu_2\,\tilde{Q}^n\| \leq \|\mu_1-\mu_2\|\,\delta(\tilde{Q})^n$ for any n and any initial states μ_1 and μ_2, where $\|\cdot\|$ is the 1-norm. As this last inequality is true for all n, the Dobrushin's coefficient may be seen as a short-term lower bound for convergence rate, comparable to $\log\alpha(\tilde{Q})$. In our case $\delta(\tilde{Q}) = 0.975$ that can be used as $-\log(\delta(\tilde{Q})) = 0.025$ so $2.5\%/\text{year}$.

4. Model evaluation

4.1. Sojourn time test

In this section we test the fit between the data and the model. The sojourn time $S(e)$ of a given state $e \in E$ is the number of consecutive time periods the Markov chain X_n remains in this state:

$$S(e) \stackrel{\text{def}}{=} \inf\{n \in \mathbb{N};\ X_n \neq e\}$$

conditionally on $X_0 = e$. The distribution law of $S(e)$ is given by:

$$\begin{aligned}
\mathbb{P}(S(e) = n | X_0 = e) &= \mathbb{P}(X_1 = \cdots = X_n = e, X_{n+1} \neq e | X_0 = e) \\
&= \sum_{e' \neq e} \mathbb{P}(X_1 = \cdots = X_n = e, X_{n+1} = e' | X_0 = e) \\
&= \sum_{e' \neq e} \underbrace{Q(e,e) \cdots Q(e,e)}_{n\,-\,1 \text{ times}} Q(e,e') = (Q(e,e))^{n-1}\,(1 - Q(e,e))
\end{aligned}$$

for $n \geq 1$ and 0 for $n = 0$, that is a geometric distribution of parameter $1 - Q(e,e) = 1 - \mathbb{P}(X_{n+1} = e | X_n = e)$. Note that $\mathbb{E}(S(e)|X_0 = e) = 1/(1 - Q(e,e))$ and $\mathrm{var}(S(e)|X_0 = e) = Q(e,e)/(1 - Q(e,e))^2$.

As the transition PA is not present in the data set, we will consider only the sojourn time on the states A and F: we now test if the sojourn time on states A and F in the data set correspond to a geometric distribution of parameter $1 - Q(\mathtt{A},\mathtt{A})$ and $1 - Q(\mathtt{F},\mathtt{F})$ respectively.

Goodness-of-fit test

In order to test if the distribution of the sojourn time $S(e)$ on each state $e \in E$ of the data set $(e^{(p)}_{1:N_p})_{p=1:42}$ is geometric, we use a bootstrap technique for goodness-of-fit on empirical distribution function proposed in [9].

Considering a sample $S_1, \ldots, S_k$ of size k from a discrete cumulative distribution function $(F(n))_{n \in \mathbb{N}}$, we aim to test the following hypothesis:

$$H_0\ :\ F \in \mathbb{F} \stackrel{\text{def}}{=} \{F_p : p \in \Theta\}\,.$$

In our case, F_p is a geometric cumulative distribution function (CDF) with parameter $p\ \in [0\ 1]$. Classically, we consider an estimator:

$$\hat{p} = T(S_{1:k})$$

of p and we compute the distance between the theoretical CDF $F_{\hat{p}}$ and the empirical CDF:

$$\hat{F}_{S_{1:k}}(n) \stackrel{\text{def}}{=} \frac{1}{k} \sum_{\ell=1}^{k} 1_{\{S_\ell \leq n\}}\,.$$

We use the Kolmogorov-Smirnov distance defined by:

$$K^* = K(S_{1:k}) \stackrel{\text{def}}{=} \sup_{n \in \mathbb{N}} \sqrt{k}\,|\hat{F}_{S_{1:k}}(n) - F_{T(S_{1:k})}(n)| \tag{21}$$

A (annual crop)	Sojourn time values	1	2	3	4	5	6	
	Number of occurrences	11	17	12	5	9	7	
F (fallow)	Sojourn time values	1	2	3	4	6	8	11
	Number of occurrences	16	12	7	4	2	1	1

Table 2. *Sojourn time values (year) on each state* A*,* F *in the data set.*

To establish whether K^* is significantly different from 0 or not, we simulate M samples of size k:

$$S_1^m, \dots, S_k^m \overset{\text{iid}}{\sim} F_{\hat{p}}, \qquad m = 1, \dots, M$$

and we let:

$$K^m \overset{\text{def}}{=} K(S_{1:k}^m)$$

where K is the function defined in (21).

The p-value associated to that test is:

$$\rho \overset{\text{def}}{=} \frac{1}{M} \sum_{m=1}^{M} 1_{\{K^m \geq K^*\}} \,.$$

If ρ is less than a given threshold α, corresponding to the probability chance of rejecting the null hypothesis H_0 when it is true, then H_0 is rejected.

Sojourn time goodness-of-fit test

In counting the sojourn time, states that appear at the end of the series of Table 1 are not treated (they are considered as censored data). Then the sojourn time values on each state A (annual crop), F (fallow) in the data set are given in Table 2.

To test the hypothesis H_0 we use the MLE for the parameter p of the geometric PDF:

$$\hat{p} \overset{\text{def}}{=} \frac{1}{1 + \frac{1}{k} \sum_{\ell=1}^{k} S_\ell} \,. \tag{22}$$

Indeed, the likelihood function is:

$$L(p) = (1-p)^{S_1}\, p \cdots (1-p)^{S_k}\, p = (1-p)^{\sum_{\ell=1}^{k} S_\ell}\, p^k$$

and $L'(p) = 0$ leads to $k - p\,(\sum_{\ell=1}^{k} S_\ell + k) = 0$ and (22).

The complete test procedure is given in Appendix E.

Results

In Figure 6 we plotted the empirical PDFs of sojourn time of states “Annual crop” and “Fallow”, associated with the data set of Table 1, and the geometric PDF corresponding to the parameter p estimated by (22). In Figure 7 we plotted the empirical PDF of K^m for the two states states. For the “Annual crop” state A, we have $K^* = 1.086$ and the associated p-value is 0.224. For the “Fallow” state F, we have $K^* = 1.104$ and the associated p-value is 0.255.

In conclusion the data appear to be consistent with the null hypothesis for the two states A and F.

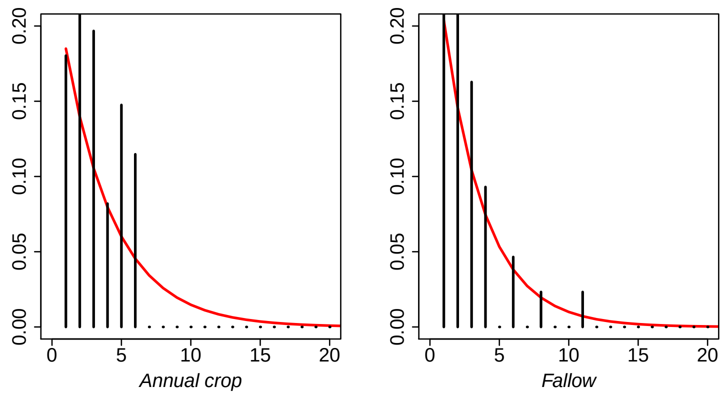


Figure 6. *Empirical PDFs of sojourn time of states* $(\mathtt{A}, \mathtt{F})$ *associated with the data set of Table 1; and the geometric PDF, represented as continuous read lines, corresponding to the parameter* p *estimated by* (22).

4.2. Homogeneity test

We can observe in the data set (Table 1) an early clearing period and a late clearing period, due to different authorized periods of slash and burn. The later clearing period, corresponding to plots cultivated from the 12th year, is limited by the year 2000 conservation measures. This explains why forest clearing is intensive just before conservation measures: after year 2001 ($n = 17$) no `fA` transition occurs.

We now want to question the homogeneity of the dynamics, i.e. the homogeneity of the Markov model, between these two time periods. The first group is $\mathcal{P}_1 = \{24, 25, \ldots, 32, 42\}$ and the second group is $\mathcal{P}_2 = \{1, \ldots, 42\} \setminus \mathcal{P}_1$. Let $\hat{Q}^{(g)}$ be the MLE of the transition matrix obtained from the data set $\mathcal{P}_g$, $g = 1, 2$. We want to test the null hypothesis:

$$H_0 : \hat{Q}^{(1)} = \hat{Q}^{(2)}$$

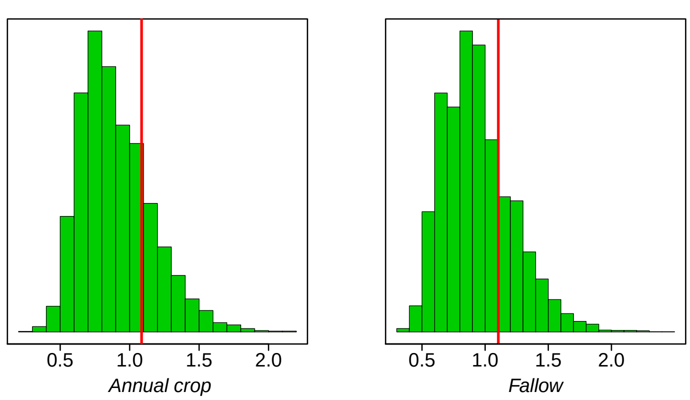
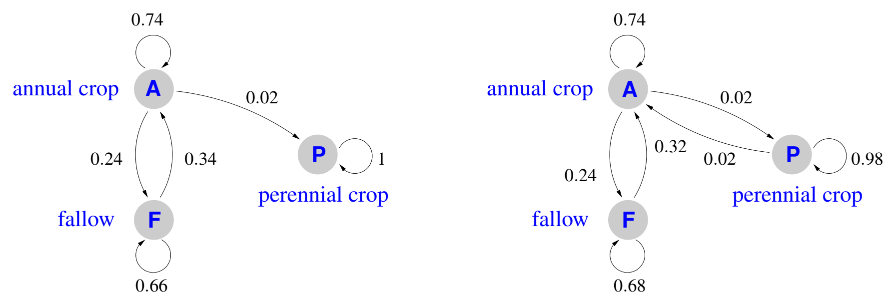


Figure 7. *Empirical PDF of* K^m *(sampled from* K^**) for the two states* $(\mathtt{A}, \mathtt{F})$ *et* K^* *value (vertical line).*



Model 1
Maximum likelihood estimation

Model 2
Bayesian estimation

Figure 8. *Comparing two models:* ***The Model 1*** *(left) is derived from the maximum likelihood estimate;* `A` *is the initial state and* `P` *is an absorbing state, the transitions* `PA`*,* `FP` *and* `PF` *do not appear in the data set of Table 1, so they are not integrated in the resulting model; the model admits a quasi-stationary distribution.* ***The Model 2*** *(right) is directly derived from a Bayesian analysis with a prior that admits the transition* `PA`*; this model is regular (i.e. it admits an invariant distribution).*

and are identical to $\hat{Q}$. The associated likelihood ratio statistics is [3]:

$$T = 2 \sum_{e \in E} \sum_{e' \in E} \sum_{g=1}^{2} n_{ee'}^{(g)} \log \left(\frac{\hat{Q}^{(g)}(e, e')}{\hat{Q}(e, e')} \right) \tag{23}$$

is asymptocally distributed as χ^2 with 6 degrees of freedom [3], and $n_{ee'}$ is the total of number of transitions from state e to state e' in the group $\mathcal{P}_g$. The estimated matrices are:

$$\hat{Q}^{(1)} = \begin{pmatrix} 0.68181818 & 0.30681818 & 0.01136364 \\ 0.31168831 & 0.68831169 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \hat{Q}^{(2)} = \begin{pmatrix} 0.78145695 & 0.20529801 & 0.01324504 \\ 0.38 & 0.62 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

We obtain $T = 3.687853$ that corresponds to a probability of 0.1344684. The null hypothesis cannot be rejected and the Markov model can be considered as homogenous over the two groups.

5. Discussion

We derived two Markovian models depicted in Figure 8. In Model 1, `P` is an absorbing state and the limit distribution will charge this state only. Model 2 is regular and admits a limit invariant measure which is strictly positive for the 3 states. The difference being that for the first model the probability of the transition `PA` is null and for the second model it is 0.02. Starting from the state `A`, the evolution of the proportions of parcels in state `A`, `F` and `P` for the two models is depicted in Figure 9.

These evolutions are almost identical over a 20 year time scale, and appear to be different only after a few decades. Over centuries they are radically different: for Model 1 almost all parcels are in the state `P`, Model 2 converges to an equilibrium after 100 years. In the first model we saw that the mean time to reach the state `P` is 136 years. In the second model, the equilibrium is reached after one century and it is 41% of parcels in state `A`, 31% in state `F`, 28% in state `P`.

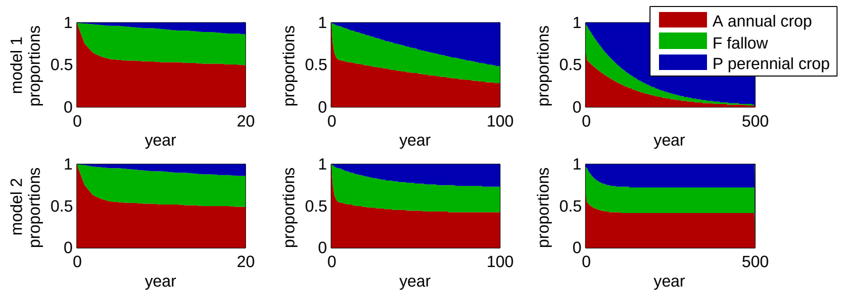


Figure 9. *Evolution of the proportions of parcels in state* A, F *and* P *during 20, 100 and 500 years according to the Models 1 and 2.*

Model 1 presents a quasistationary behavior. Indeed if we consider the evolution of the relative proportion of parcels in state A and in state F, Model 1 and Model 2 present quasi identical profiles (see Figure 10).

Hence after less than 5 years, the relative proportion of parcels in state A and in state F are close to 58% and 42%. This proportion may be interpreted as a state of agricultural intensification without external fertilizers. In order to test the effect of intensification (more frequent A and less frequent F) we would require further data.

6. Conclusion and perspectives

We proposed Markovian models of land use dynamics for parcels near the forest corridor of Ranomafana and Andringitra national parks in Madagascar. The first model is derived directly from the data set from an empirical procedure that corresponds to the maximum likelihood estimator. The resulting model features an absorbing state P and the limit distribution of the model only charges that state. We studied the quasistationary distribution of the model and the time to reach state P. The PA transition, even if it does not appear in the data set, is realistic in the time scale considered here and we need to use a Bayesian approach so that it can appear in the model. Among the priors, that of Jeffreys gives the best results. As it was not possible to explicitly calculate the associated estimator, a MCMC technique was used. The resulting Markov model is regular and admits a

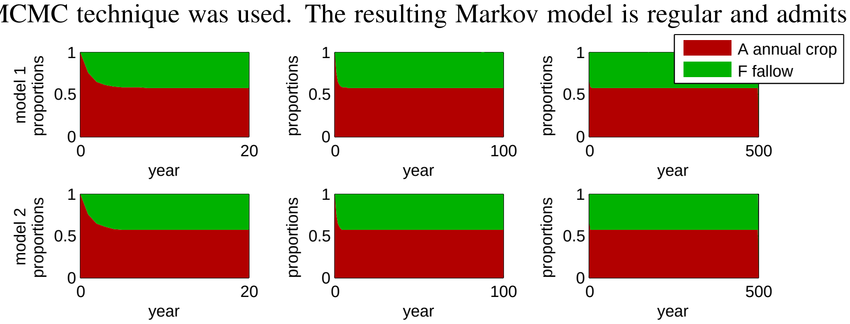


Figure 10. *Evolution of the proportions of parcels in state* A *and* F *during 20, 100 and 500 years according to the Models 1 and 2.*

unique invariant measure that charges all the states. We studied the speed of convergence to this limit distribution.

We assessed the adequacy of the model to real data. We focused on the sojourn times: we tested if the empirical sojourn times correspond to a geometric distribution. We used a parametric bootstrap goodness-of-fit on empirical distribution. We also tested the spatial homogeneity of the model.

A new database is currently being developed by the IRD. It will cover a longer period of time and a greater number of parcels, it will also allow us to consider a more detailed state space comprising more than three states.

Part of the complexity of these agro-ecological temporal data comes from the fact that some transitions are "natural", due to ecological dynamics, while others come from human decisions (annual cropping, crop abandonment, planting perennial crops, etc.). It should also be interesting to study the dynamics of parcels conditionally on the dynamics of the neighboring parcels. This model could be more realistic but would require a larger number of unknown parameters that the present data set will not permit to infer and a more in depth study of the farmers' practices will hence be necessary.

In the case of retrospective studies and forest-agriculture transitions, we lack a long term series of plot use histories, especially in tropical southern countries. The effect of a limited database, either in size (number of plots) or in time (period), is poorly known. In this context, the Bayesian approach is of interest in developing further conclusions based on complementary knowledge from agronomists, ecologists and geographers.

Acknowledgements

This works was partially supported by the LIRIMA/INRIA program and by the Cooperation and Cultural Action Office (SCAC) of the France Embassy in Madagascar.

Appendices

A. Explicit Bayes estimators for the two states case

Let $(X_n)_{0\leq n\leq N}$ be a Markov chain with two states $\{0,1\}$ and transition matrix

$$Q \stackrel{\text{def}}{=} \left(\begin{smallmatrix} p & 1-p \\ 1-q & q \end{smallmatrix}\right) .$$

We suppose that the initial law is the invariant distribution $\mu = (q/p+q, p/p+q)$, that is the solution of $\mu Q = \mu$. The unknown parameter is $\theta = (p,q) \in [0,1]^2$ and the associated likelihood function is

$$L_N(\theta) \stackrel{\text{def}}{=} \mathbb{P}_\theta(X_{0:N} = x_{0:N}) = (1-p)^{n_{00}}\, p^{n_{01}}\, q^{n_{10}}\, (1-q)^{n_{11}}$$

where $n_{ij} \stackrel{\text{def}}{=} n_{ij}(x_{0:N}) = \sum_{n=0}^{N-1} 1_{\{X_n=i\}}\, 1_{\{X_{n+1}=j\}}$ is the number of transition $i \to j$ in $x_{0:N}$.

We consider the following priori distributions: the uniform distribution π^{U} on $[0,1]^2$ and the beta distribution π^{B} with parameters (a,b), that is

$$\pi^{\text{B}}(\theta) = \frac{1}{\beta(a,b)}\, \theta^{a-1}\, (1-\theta)^{b-1}$$

where $\beta(a,b)$ is the beta function:

$$\beta(a,b) \stackrel{\text{def}}{=} \int_0^1 x^{a-1}\,(1-x)^{b-1}\,\mathrm{d}x = \frac{\Gamma(a)\,\Gamma(b)}{\Gamma(a+b)}$$

with $\Gamma(z) \stackrel{\text{def}}{=} \int_0^{+\infty} t^{z-1}\,e^{-t}\,\mathrm{d}t$. Here we will choose $a=b=1/2$, note that $\Gamma(1)=1$ and $\Gamma(\frac{1}{2})=\sqrt{\pi}$. For these two priors we can explicitly compute the posterior distriution and the associated Bayes estimators. Indeed the posterior distribution π_{post} is given by the Bayes formula: $\pi_{\text{post}}(\theta) \propto L_N(\theta)\,\pi(\theta)$, that is:

$$\begin{aligned}
&\pi^{\text{U}}_{\text{post}}(\theta) \propto L_N(\theta) = (1-p)^{n_{00}}\,p^{n_{01}}\,q^{n_{10}}\,(1-q)^{n_{11}}\,,\\
&\pi^{\text{B}}_{\text{post}}(\theta) \propto L_N(\theta)\,\pi^{\text{B}}(\theta) = (1-p)^{n_{00}-\frac{1}{2}}\,p^{n_{01}-\frac{1}{2}}\,q^{n_{10}-\frac{1}{2}}\,(1-q)^{n_{11}-\frac{1}{2}}
\end{aligned}$$

and the corresponding Bayes estimator are:

$$\tilde{\theta}^{\text{U}} = \int_{[0,1]^2} \theta\,\pi^{\text{U}}_{\text{post}}(\theta)\,\mathrm{d}\theta\,, \qquad \tilde{\theta}^{\text{B}} = \int_{[0,1]^2} \theta\,\pi^{\text{B}}_{\text{post}}(\theta)\,\mathrm{d}\theta\,.$$

We can easily check that the estimators of p and q for the uniform prior:

$$\tilde{p}^{\text{U}} = \frac{n_{01}+1}{n_{01}+n_{00}+2}\,, \qquad \tilde{q}^{\text{U}} = \frac{n_{10}+1}{n_{10}+n_{11}+2}$$

and for the beta prior:

$$\tilde{p}^{\text{B}} = \frac{n_{01}+\frac{1}{2}}{n_{01}+n_{00}+1}\,, \qquad \tilde{q}^{\text{B}} = \frac{n_{10}+\frac{1}{2}}{n_{10}+n_{11}+1}\,.$$

Note that in this case the MLEs are:

$$\hat{p}^{\text{MLE}} = \frac{n_{01}}{n_{00}+n_{01}}\,, \qquad \hat{q}^{\text{EMV}} = \frac{n_{10}}{n_{11}+n_{10}}\,.$$

B. Distribution law of the time to reach a given state

Let X_n be an homogeneous Markov chain with finite state space E and transition matrix Q. We aim to get an explicit expression of the distribution law:

$$f_{ee'}(n) \stackrel{\text{def}}{=} \mathbb{P}(\tau_{e'}=n|X_0=e)$$

of the first passage time $\tau_{e'}$ defined by (8). Note that for $n>1$:

$$\begin{aligned}
Q^n(e,e') &= \mathbb{P}(X_n=e'|X_0=e)\\
&= \mathbb{P}(X_n=e',\,\tau_{e'}=1|X_0=e) + \mathbb{P}(X_n=e',\,\tau_{e'}=2|X_0=e) + \cdots\\
&\quad \cdots + \mathbb{P}(X_n=e',\,\tau_{e'}=n-1|X_0=e) + \mathbb{P}(X_n=e',\,\tau_{e'}=n|X_0=e)\\
&= \mathbb{P}(\tau_{e'}=1|X_0=e)\,\mathbb{P}(X_n=e'|X_1=e')\\
&\quad + \mathbb{P}(\tau_{e'}=2|X_0=e)\,\mathbb{P}(X_n=e'|X_2=e') + \cdots\\
&\quad \cdots + \mathbb{P}(\tau_{e'}=n-1|X_0=e)\,\mathbb{P}(X_n=e'|X_{n-1}=e') + \mathbb{P}(\tau_{e'}=n|X_0=e)\\
&= f_{ee'}(1)\,Q^{n-1}(e',e') + f_{ee'}(2)\,Q^{n-2}(e',e') + \cdots\\
&\quad \cdots + f_{ee'}(n-1)\,Q^1(e',e') + f_{ee'}(n)
\end{aligned}$$

hence $f_{ee'}(n)$ could be computed recursively according to

$$f_{ee'}(n) = Q^n(e,e') - \sum_{k=1}^{n-1} f_{ee'}(k)\, Q^{n-k}(e',e') \tag{24}$$

with $f_{ee'}(1) = Q(e,e')$; for the mean we can use $\mathbb{E}(\tau_{ee''}|X_0 = e) = \sum_{n=1}^{\infty} n\, f_{ee'}(n)$.

C. Quasi-stationary distribution

We consider the probability to be in $e \in \{\mathtt{A},\mathtt{F}\}$ before reaching $\mathtt{P}$ and starting from $\mathtt{A}$, i.e. $\mu_n(e) = \mathbb{P}(X_n = e|\tau_{\mathtt{P}} > n,\, X_0 = \mathtt{A})$ where $\tau_{\mathtt{P}}$ is the first time to reach $\mathtt{P}$, see (8). When

$$\mu_n(e) \xrightarrow[n\to\infty]{} \sigma_{\text{qs}}(e)\,,\ e \in \{\mathtt{A},\mathtt{F}\}$$

the probability distribution $(\sigma_{\text{qs}}(e))_{e\in\{\mathtt{A},\mathtt{F}\}}$ is called quasi-stationary probability distribution. This problem was originally solved in [7] (see [21]): $\sigma_{\text{qs}} = [\sigma_{\text{qs}}(\mathtt{A}), \sigma_{\text{qs}}(\mathtt{F})]$ exists and is solution of

$$\sigma_{\text{qs}}\, Q_{\text{qs}} = \lambda\, \sigma_{\text{qs}}$$

with $\sigma_{\text{qs}}(e) \geq 0$ and $\sigma_{\text{qs}}(\mathtt{A}) + \sigma_{\text{qs}}(\mathtt{F}) = 1$, where Q_{qs} is the 2×2 submatrix defined by

$$Q = \left(\begin{array}{c|c} Q_{\text{qs}} & q \\ \hline 0 & 1 \end{array}\right),$$

and λ is the spectral radius of Q_{qs}.

D. Metropolis-Hastings algorithm

For the MCMC method of Section 3.2, we propose a Metropolis-Hastings algorithm with a Gaussian proposal distribution:

choose θ
$\bar\theta \leftarrow \theta$
for $k = 2, 3, 4, \dots$ **do**
  $\varepsilon \sim \mathcal{N}(0, \sigma^2)$
  $\theta^{\text{prop}} \leftarrow \theta + \varepsilon$
  $u \sim U[0,1]$
  $\alpha \leftarrow \min\left\{1, \dfrac{\pi_{\text{post}}(\theta^{\text{prop}})\, g(\theta - \theta^{\text{prop}})}{\pi_{\text{post}}(\theta)\, g(\theta^{\text{prop}} - \theta)}\right\}$ % g PDF of $\mathcal{N}(0,\sigma^2)$
  if $u \leq \alpha$ **then**
    $\theta \leftarrow \theta^{\text{prop}}$ % acceptation
  end if
  $\bar\theta \leftarrow \dfrac{k-1}{k}\,\bar\theta + \dfrac{1}{k}\,\theta$
end for

The target distribution is $\pi_{\text{post}}(\theta)$ defined by (10), the Gaussian proposal PDF (probability density function) is $g(\cdot - \theta)$ (g PDF of the $\mathcal{N}(0,\sigma^2)$ distribution) where θ is the current value of the parameter.

D. Parametric bootstrap for goodness-of-fit

The algorithm for the parametric bootstrap for goodness-of-fit test of Section 4.1 with the geometric distribution of parameter p is:

$\hat{p} \leftarrow T(S_{1:k})$
$\bar{n} \leftarrow \sup(S_1, \ldots, S_k)$
$K^* \leftarrow \sup\{\sqrt{k}|F^{S_{1:k}}(n) - F^{T(S_{1:k})}(n)|\,,\, 0 \le n \le \bar{n}\}$
for $m = 1, 2, \ldots, M$ **do**
  $S_1^m, \ldots, S_k^m \overset{\text{iid}}{\sim} F_{\hat{p}}$
  $\bar{n} \leftarrow \sup(S_1^m, \ldots, S_k^m)$
  $K^m \leftarrow \sup\{\sqrt{k}|F^{S_{1:k}^m}(n) - F^{T(S_{1:k}^m)}(n)|\,,\, 0 \le n \le \bar{n}\}$
  $\rho \leftarrow \frac{1}{M}\sum_{m=1}^{M} 1_{\{K^m \ge K^*\}}$
end for
if $\rho \le \alpha$ **then**
  accept H_0
else
  reject H_0
end if

A. References

[1] D. K. Agarwal, A. Gelfand, and J. Silander. Investigating tropical deforestation using two-stage spatially misaligned regression models. *Journal of Agricultural, Biological, and Environmental Statistics*, 7(3):420–439, 2002.

[2] D. K. Agarwal, J. A. Silander Jr., A. E. Gelfand, R. E. Deward, and J. G. Mickelson Jr. Tropical deforestation in Madagascar: analysis using hierarchical, spatially explicit, Bayesian regression models. *Ecological Modelling*, 185:105–131, 2005.

[3] T. W. Anderson and L. A. Goodman. Statistical inference about Markov chains. *Annals of Mathematical Statistics*, 28:89–109, 1957.

[4] S. Assoudou and B. Essebbar. A Bayesian model for binary Markov chains. *International Journal of Mathematics and Mathematical Sciences*, 8:421–429, 2004.

[5] S. M. Carrière, D. Hervé, F. Andriamahefazafy, and P. Méral. Corridors: Compulsory passages ? The Malagasy example. In Catherine Aubertin and Estienne Rodary, editors, *Protected Areas, Sustainable Land ?*, pages 53–69. IRD Editions - Ashgate, 2011.

[6] H. Caswell. *Matrix Population Models: Construction, Analysis, and Interpretation*. Sinauer Associates, second edition, 2001.

[7] J. N. Darroch and E. Seneta. On quasi-stationary distributions in absorbing discrete-time finite Markov chains. *Journal of Applied Probability*, 2(1):88–100, 1965.

[8] C. Genest and B. Rémillard. Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 44:1096–1127, 2008.

[9] N. Henze. Empirical-distribution-function goodness-of-fit tests for discrete models. *The Canadian Journal of Statistics / La Revue Canadienne de Statistique*, 24(1):81–93, 1996.

[10] M. F. Hill, J. D. Witman, and H. Caswell. Markov chain analysis of succession in a rocky subtidal community. *The American Naturalist*, 164:E46–E61, 2004.

[11] J. G. Kemeny and J. L. Snell. *Finite Markov Chains*. Springer, second edition, 1976.

[12] J.-M. Marin and C. P. Robert. *Bayesian Core: A Practical Approach to Computational Bayesian Statistics*. Springer-Verlag, 2007.

[13] M. R. Meshkani and L. Billard. Empirical Bayes estimators for a finite Markov chain. *Biometrika*, 79(1):185–193, 1992.

[14] C. R. Rakotoasimbahoaka, V. Ratiarson, B. O. Ramamonjisoa, and D. Hervé. Modélisation de la dynamique d'aménagement des bas-fonds rizicoles en forêt. In *Proceedings of the 10th International African Conference on Research in Computer Science and Applied mathematics (CARI'10), October 18–21, Yamoussoukro, Côte d'Ivoire*, pages 293–300, 2010.

[15] V. Ratiarson, D. Hervé, C. R. Rakotoasimbahoaka, and J.-P. Müller. Calibration et validation d'un modèle de dynamique d'occupation du sol postforestière à base d'automate temporisé à l'aide d'un modèle markovien. *Cahiers Agricultures*, 20(4):274–9, 2011.

[16] E. Seneta. *Nonnegative matrices and Markov chains*. Springer-Verlag, New York, second edition, 1981.

[17] W. Stute, W. Manteiga, and M. Quindimil. Bootstrap based goodness-of-fit-tests. *Metrika*, 40(1):243–256, December 1993.

[18] M. Sung, R. Soyer, and N. Nhan. Bayesian analysis of non-homogeneous Markov chains: Application to mental health data. *Statistics in Medecine*, 26:3000–3017, 2007.

[19] B. C. Tucker and M. Anand. The application of Markov models in recovery and restoration. *International Journal of Ecology and Environmental Sciences*, 30:131–140, 2004.

[20] M. B. Usher. Markovian approaches to ecological succession. *Journal of Animal Ecology*, 48(2):413–426, 1979.

[21] E. A. van Doorn and P. K. Pollett. Quasi-stationary distributions for reducible absorbing Markov chains in discrete time. *Markov Processes Relat. Fields*, 15:191–204, 2009.

[22] P. E. Waggoner and G. R. Stephens. Transition probabilities for a forest. *Nature*, 225:1160–1161, 1970.