

Un modèle markovien de transition agraire

Fabien Campillo¹ — Angelo Raherinirina¹ — Rivo Rakotozafy²

¹ Projet MERE, INRIA/INRA, Montpellier, France
Fabien.Campillo@inria.fr

² Université de Fianarantsoa, Madagascar
rrakotozafy@uni-fianar.mg

.....
RÉSUMÉ. Nous présentons un modèle de Markov de transition agraire à 4 états. L'application de ce genre de modèles aux problèmes de successions agraires n'est pas nouveau, mais les outils relativement récents de l'inférence bayésienne numérique permettent de tester des lois a priori plus générales. En plus de l'estimateur du maximum de vraisemblance, nous considérons la loi a priori de Jeffreys (non informative) et calculons l'estimateur bayésien associé à l'aide d'une approximation de Metropolis-Hastings. Nous étudions les capacités et les limites de cette approche.

ABSTRACT. We present a Markov model for agricultural successions with 4 states. The application of Markov models to agricultural succession problems is not new, but relatively new tools of numerical Bayesian inference allow us to test general prior laws. In addition to the maximum likelihood estimate, we consider the Jeffreys prior (non-informative), and calculate the associated Bayesian estimator with a Metropolis-Hastings approximation procedure. We study the capabilities and limitations of this approach.

MOTS-CLÉS : Modèle de Markov, Monte Carlo par chaîne de Markov (MCCM), a priori de Jeffreys

KEYWORDS : Markov model, Markov chain Monte Carlo (MCMC), Jeffreys prior

1. Introduction

Les chaînes de Markov sont depuis longtemps utilisées afin de modéliser des successions agraires ou écologiques (cf. Waggoner [5] et Usher [4]). Inférer la dynamique considérée revient à estimer les paramètres de la matrice de transition.

Dans cet article nous nous plaçons dans le cadre d'une chaîne à 4 états. Nous présentons le problème d'inférence correspondant. Pour cela nous reprenons les travaux de Assoudou et Essebbar [1]. Nous proposons un calcul plus simple de la loi a priori de Jeffreys, que nous échantillonnons à l'aide d'une procédure de Metropolis–Hastings. Nous présentons enfin quelques essais numériques.

Les données étudiées nous ont été proposées par Dominique Hervé (IRD) et correspondent à des parcelles situées sur les versants et les bas fonds de la commune d'Ambendrana, à la lisière du corridor forestier de Ranomafana–Andringitra. La compréhension des dynamiques agricoles jouxtant ce corridor forestier est de la plus grande importance : ce corridor est le dernier vestige d'une vaste forêt qui recouvrait une grande partie de la province de Fianarantsoa à Madagascar. Cette forêt s'est peu à peu réduite à une bande de 200 km de long et d'au plus 20 km de large entre les aires protégées de Ranomafana et Andringitra. Les données utilisées dans cet article sont disponibles dans [3].

2. Le modèle

On dispose de l'historique de successions d'occupations du sol de $P = 43$ parcelles sur une période de $N = 22$ ans :

$$i_{1:N}^{1:P} = \begin{pmatrix} i_1^1 & \cdots & i_N^1 \\ i_1^2 & \cdots & i_N^2 \\ \vdots & & \vdots \\ i_1^P & \cdots & i_N^P \end{pmatrix} \quad (1)$$

où i_n^p désigne l'état de la parcelle p à l'année n , i_n^p peut prendre 4 valeurs : F (forêt), C (culture), J (jachère), B (bois). L'espace d'état est noté $I \stackrel{\text{def}}{=} \{F, C, J, B\}$.

Chaque parcelle part de l'état initial "forêt" (naturelle), elle peut alors être exploitée en "culture" ou laissée en "jachère", elle peut ensuite basculer de "culture" à "jachère" et inversement, l'état "bois" correspond à l'utilisation finale de la parcelle (état absorbant) qui est utilisée pour l'exploitation du bois (cf. Figure 1). Les autres transitions sont impossibles (i.e. correspondent à des probabilités de transition nulles) ; il est important de noter que ces transitions "impossibles" n'apparaissent pas dans les données (1) que nous avons eu à traiter.

Nous faisons deux hypothèses :

- (i) chaque parcelle est une réalisation indépendante d'un même processus ;
- (ii) ce processus est markovien (homogène).

Sous ces hypothèses (1) est la réalisation d'un processus :

$$X_{1:N}^{1:P} = \begin{pmatrix} X_1^1 & \cdots & X_N^1 \\ X_1^2 & \cdots & X_N^2 \\ \vdots & & \vdots \\ X_1^P & \cdots & X_N^P \end{pmatrix} \quad (2)$$

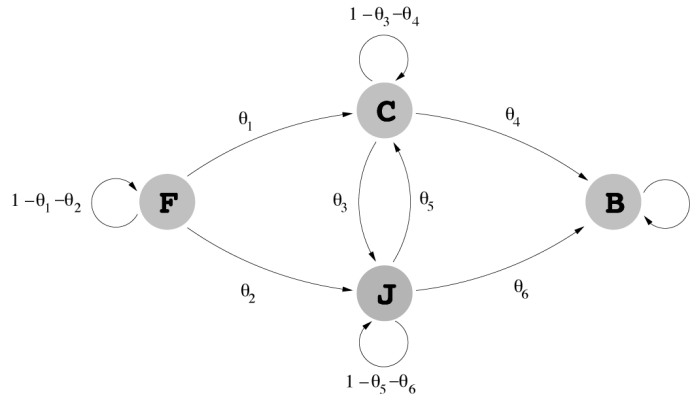


Figure 1. Le modèle de chaîne de Markov à 4 états : F (forêt), C (culture), J (jachère), B (bois). Le paramètre de la loi de transition est $\theta = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)$. F est l'état initial, B est un état absorbant.

où les lignes $X_{1:N}^p$, pour $p = 1, \dots, P$, sont indépendantes et sont chacune une chaîne de Markov de matrice de transition :

$$Q = \begin{pmatrix} 1 - \theta_1 - \theta_2 & \theta_1 & \theta_2 & 0 \\ 0 & 1 - \theta_3 - \theta_4 & \theta_3 & \theta_4 \\ 0 & \theta_5 & 1 - \theta_5 - \theta_6 & \theta_6 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (3)$$

Le paramètre inconnu est donc $\theta = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6) \in \Theta$ avec :

$$\Theta \stackrel{\text{déf}}{=} \{ \theta \in [0, 1]^6 ; \theta_1 + \theta_2 \leq 1, \theta_3 + \theta_4 \leq 1, \theta_5 + \theta_6 \leq 1 \}.$$

L'état initial de chaque parcelle est F, la loi initiale de chaque ligne $X_{1:N}^p$ est donc $\delta_F = [1 \ 0 \ 0 \ 0]$ (i.e. $\delta_{Fi} = 1$ si $i = F$ et 0 sinon). Les hypothèses (i) et (ii) impliquent que la loi de $X_{1:N}^{1:P}$ est :

$$\mathbb{P}_\theta(X_{1:N}^{1:P} = i_{1:N}^{1:P}) = \prod_{p=1}^P \mathbb{P}_\theta(X_{1:N}^p = i_{1:N}^p) = \prod_{p=1}^P \delta_{Fi_1^p} Q_{i_1^p i_2^p} \cdots Q_{i_{N-1}^p i_N^p}$$

pour tout $N > 1$, $P > 1$ et tout $i_n^p \in I$ avec $1 \leq n \leq N$ et $1 \leq p \leq P$.

Nous supposons à partir de maintenant que $i_1^p = F$ pour tout $p = 1, \dots, P$. Donc d'après (3), la fonction de log-vraisemblance s'écrit :

$$l(\theta) = \log L(\theta) \stackrel{\text{déf}}{=} \log \mathbb{P}_\theta(X_{1:N}^{1:P} = i_{1:N}^{1:P}) = \sum_{p=1}^P [\log Q_{i_1^p i_2^p} + \cdots + \log Q_{i_{N-1}^p i_N^p}]$$

On montre aisément que :

$$\begin{aligned} l(\theta) = & n_{FF} \log(1 - \theta_1 - \theta_2) + n_{FC} \log(\theta_1) + n_{FJ} \log(\theta_2) \\ & + n_{CC} \log(1 - \theta_3 - \theta_4) + n_{CJ} \log(\theta_3) + n_{CB} \log(\theta_4) \\ & + n_{JJ} \log(1 - \theta_5 - \theta_6) + n_{JC} \log(\theta_5) + n_{JB} \log(\theta_6) \end{aligned} \quad (4)$$

où n_{ij} désigne le nombre de transitions de i vers j dans les séquences $X_{1:N}^1$ à $X_{1:N}^P$, i.e. :

$$n_{ij} \stackrel{\text{def}}{=} \sum_{p=1}^P \sum_{n=2}^N 1_{\{X_{n-1}^p=i\}} 1_{\{X_n^p=j\}}, \quad \forall i, j \in I. \quad (5)$$

On calcule les estimateurs du maximum de vraisemblance des éléments de Q en annulant les dérivées partielles de la fonction de vraisemblance (4). On vérifie aisément que :

$$\begin{aligned} \hat{\theta}_1 &= \frac{n_{\text{FC}}}{n_{\text{FF}} + n_{\text{FC}} + n_{\text{FJ}}}, & \hat{\theta}_2 &= \frac{n_{\text{FJ}}}{n_{\text{FF}} + n_{\text{FC}} + n_{\text{FJ}}}, \\ \hat{\theta}_3 &= \frac{n_{\text{CJ}}}{n_{\text{CC}} + n_{\text{CJ}} + n_{\text{CB}}}, & \hat{\theta}_4 &= \frac{n_{\text{CB}}}{n_{\text{CC}} + n_{\text{CJ}} + n_{\text{CB}}}, \\ \hat{\theta}_5 &= \frac{n_{\text{JC}}}{n_{\text{JC}} + n_{\text{JJ}} + n_{\text{JB}}}, & \hat{\theta}_6 &= \frac{n_{\text{JB}}}{n_{\text{JC}} + n_{\text{JJ}} + n_{\text{JB}}}. \end{aligned}$$

3. Estimateur bayésien

Une alternative à l'estimateur du maximum de vraisemblance sont les estimateurs bayésiens. Ils s'appuient sur la formule de Bayes :

$$\tilde{\pi}(\theta) \propto L(\theta) \pi(\theta). \quad (6)$$

L'estimateur bayésien de θ est alors calculé de la façon suivante :

$$\tilde{\theta} \stackrel{\text{def}}{=} \int \theta \tilde{\pi}(\theta) d\theta. \quad (7)$$

Le choix de la loi a priori $\pi(\theta)$ est un point crucial de l'inférence bayésienne. Nous pouvons faire appel à une loi initiale uniforme ou beta, dans ce cas le calcul de $\tilde{\theta}$ est explicite. Nous avons toutefois obtenu de meilleurs résultats avec la loi initiale de Jeffreys : dans ce cas le calcul (7) n'est pas explicite et nous avons dû faire appel à une technique d'approximation de type MCMC. C'est cette approche que nous détaillons maintenant.

3.1. Calcul de la loi de Jeffreys

La loi de Jeffreys est définie par :

$$\pi(\theta) \propto \sqrt{\det \mathcal{I}_N(\theta)}$$

où $\mathcal{I}(\theta)$ est la matrice d'information de Fisher :

$$\mathcal{I}(\theta) \stackrel{\text{def}}{=} \left[\mathbb{E}_{\theta} \left(- \frac{\partial^2 l(\theta)}{\partial \theta_k \partial \theta_{\ell}} \right) \right]_{1 \leq k, \ell \leq 6}.$$

On vérifie aisément que la matrice d'information de Fisher est diagonale par blocs :

$$\mathcal{I}_N(\theta) = \begin{pmatrix} A_1 & 0 & 0 \\ 0 & A_3 & 0 \\ 0 & 0 & A_5 \end{pmatrix}$$

avec :

$$A_k \stackrel{\text{d\'ef}}{=} -\mathbb{E}_\theta \left[\begin{pmatrix} \frac{\partial^2 l(\theta)}{\partial^2 \theta_k} & \frac{\partial^2 l(\theta)}{\partial \theta_k \partial \theta_{k+1}} \\ \frac{\partial^2 l(\theta)}{\partial \theta_k \partial \theta_{k+1}} & \frac{\partial^2 l(\theta)}{\partial^2 \theta_{k+1}} \end{pmatrix} \right].$$

Ainsi :

$$\pi(\theta) \propto \sqrt{\det A_1 \times \det A_3 \times \det A_5}. \quad (8a)$$

D'après (4) :

$$\begin{aligned} \frac{\partial^2 l(\theta)}{\partial^2 \theta_1} &= -\left(\frac{n_{FF}}{(1-\theta_1-\theta_2)^2} + \frac{n_{FC}}{\theta_1^2} \right), & \frac{\partial^2 l(\theta)}{\partial \theta_1 \partial \theta_2} &= -\frac{n_{FF}}{(1-\theta_1-\theta_2)^2}, \\ \frac{\partial^2 l(\theta)}{\partial^2 \theta_2} &= -\left(\frac{n_{FF}}{(1-\theta_1-\theta_2)^2} + \frac{n_{FJ}}{\theta_2^2} \right), & \frac{\partial^2 l(\theta)}{\partial \theta_2 \partial \theta_1} &= -\frac{n_{FF}}{(1-\theta_1-\theta_2)^2}, \\ \frac{\partial^2 l(\theta)}{\partial^2 \theta_3} &= -\left(\frac{n_{CC}}{(1-\theta_3-\theta_4)^2} + \frac{n_{CJ}}{\theta_3^2} \right), & \frac{\partial^2 l(\theta)}{\partial \theta_3 \partial \theta_4} &= -\frac{n_{CC}}{(1-\theta_3-\theta_4)^2}, \\ \frac{\partial^2 l(\theta)}{\partial^2 \theta_4} &= -\left(\frac{n_{CC}}{(1-\theta_3-\theta_4)^2} + \frac{n_{CB}}{\theta_4^2} \right), & \frac{\partial^2 l(\theta)}{\partial \theta_4 \partial \theta_3} &= -\frac{n_{CC}}{(1-\theta_3-\theta_4)^2}, \\ \frac{\partial^2 l(\theta)}{\partial^2 \theta_5} &= -\left(\frac{n_{JJ}}{(1-\theta_5-\theta_6)^2} + \frac{n_{JC}}{\theta_5^2} \right), & \frac{\partial^2 l(\theta)}{\partial \theta_5 \partial \theta_6} &= -\frac{n_{JJ}}{(1-\theta_5-\theta_6)^2}, \\ \frac{\partial^2 l(\theta)}{\partial^2 \theta_6} &= -\left(\frac{n_{JJ}}{(1-\theta_5-\theta_6)^2} + \frac{n_{JB}}{\theta_6^2} \right), & \frac{\partial^2 l(\theta)}{\partial \theta_6 \partial \theta_5} &= -\frac{n_{JJ}}{(1-\theta_5-\theta_6)^2} \end{aligned}$$

(les autres termes \'etant nuls) et :

$$\det A_1 = \left(\frac{\mathbb{E}_\theta[n_{FF}]}{(1-\theta_1-\theta_2)^2} + \frac{\mathbb{E}_\theta[n_{FC}]}{\theta_1^2} \right) \left(\frac{\mathbb{E}_\theta[n_{FF}]}{(1-\theta_1-\theta_2)^2} + \frac{\mathbb{E}_\theta[n_{FJ}]}{\theta_2^2} \right) - \left(\frac{\mathbb{E}_\theta[n_{FF}]}{(1-\theta_1-\theta_2)^2} \right)^2, \quad (8b)$$

$$\det A_3 = \left(\frac{\mathbb{E}_\theta[n_{CC}]}{(1-\theta_3-\theta_4)^2} + \frac{\mathbb{E}_\theta[n_{CJ}]}{\theta_3^2} \right) \left(\frac{\mathbb{E}_\theta[n_{CC}]}{(1-\theta_3-\theta_4)^2} + \frac{\mathbb{E}_\theta[n_{CB}]}{\theta_4^2} \right) - \left(\frac{\mathbb{E}_\theta[n_{CC}]}{(1-\theta_3-\theta_4)^2} \right)^2, \quad (8c)$$

$$\det A_5 = \left(\frac{\mathbb{E}_\theta[n_{JJ}]}{(1-\theta_5-\theta_6)^2} + \frac{\mathbb{E}_\theta[n_{JC}]}{\theta_5^2} \right) \left(\frac{\mathbb{E}_\theta[n_{JJ}]}{(1-\theta_5-\theta_6)^2} + \frac{\mathbb{E}_\theta[n_{JB}]}{\theta_6^2} \right) - \left(\frac{\mathbb{E}_\theta[n_{JJ}]}{(1-\theta_5-\theta_6)^2} \right)^2. \quad (8d)$$

Enfin, d'après (5) :

$$\mathbb{E}_\theta[n_{ij}] = \sum_{p=1}^P \sum_{n=2}^N \mathbb{P}_\theta(X_n^p = j, X_{n-1}^p = i) = P Q_{ij} \sum_{n=2}^N [Q^{n-1}]_{Fi}. \quad (8e)$$

Le calcul de la loi a posteriori se fait donc à l'aide de la formule de Bayes (6), de l'expression de la log-vraisemblance (4), et de l'expression de la loi a priori (8). Ces calculs sont détaillés dans [3].

Dans [1], les auteurs proposent une version différente et un peu plus complexe de cette loi a priori. Ils se limitent de plus au cas à 2 états.

3.2. MCMC

La loi a posteriori $\tilde{\pi}(\theta)$ définie par (6), associée à la loi a priori de Jeffreys, n'est pas intégrable explicitement. Ainsi il n'est pas possible de calculer directement l'estimateur associé (7). Dans cette section, nous proposons une technique MCMC afin d'échantillonner selon $\tilde{\pi}(\theta)$ et de calculer $\tilde{\theta}$. Nous utilisons un échantillonneur indépendant : la loi de proposition est uniforme sur l'ensemble Θ . La procédure qui est découlée est décrite dans Algorithme 1.

```

choisir  $\theta^{(1)}$ 
 $\hat{\theta} \leftarrow \theta^{(1)}$ 
pour  $m = 2, 3, 4, \dots$  faire
   $\theta \sim U(\Theta)$  % loi uniforme sur  $\Theta$ 
   $u \sim U[0, 1]$ 
   $\alpha \leftarrow \tilde{\pi}(\theta^{\text{prop}}) / \tilde{\pi}(\theta^{(m-1)})$ 
  si  $u \leq \alpha$  alors
     $\theta^{(m)} \leftarrow \theta^{\text{prop}}$  % acceptance
  sinon
     $\theta^{(m)} \leftarrow \theta^{(m-1)}$  % rejet
  fin si
   $\hat{\theta} \leftarrow \frac{m-1}{m} \hat{\theta} + \frac{1}{m} \theta^{(m)}$ 
fin pour

```

Algorithme 1: Méthode MCMC avec tirage indépendant pour échantillonner selon la loi a posteriori $\tilde{\pi}(\theta)$ et calcul d'une approximation de $\hat{\theta}$ défini par (7).

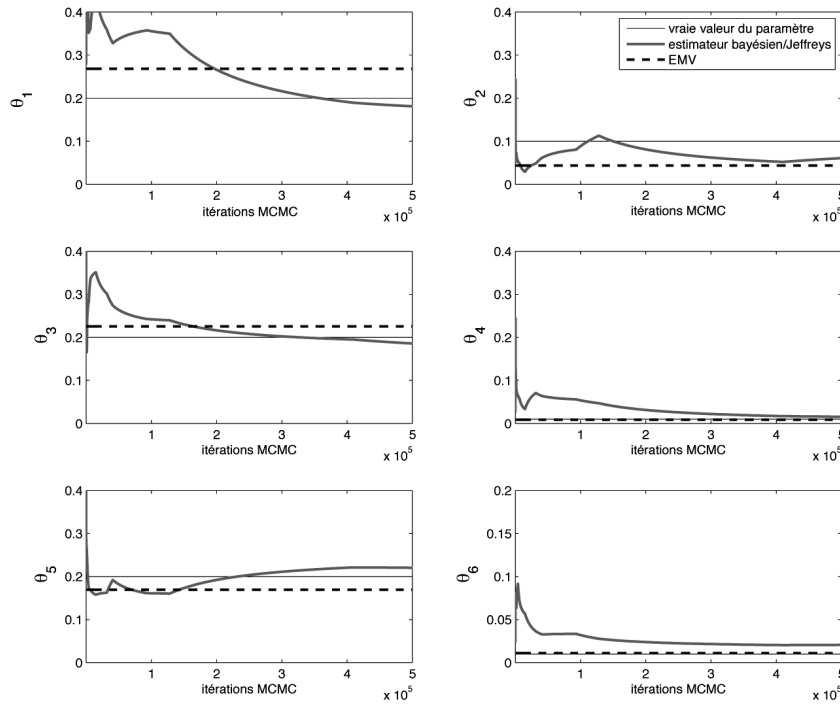


Figure 2. Données simulées : Comparaison entre la vraie valeur du paramètre, l'EMV et les itérations de l'algorithme MCMC pour le calcul de l'estimateur bayésien. Les erreurs quadratique sont de 0.051 pour l'estimateur bayésien et de 0.097 pour l'EMV.

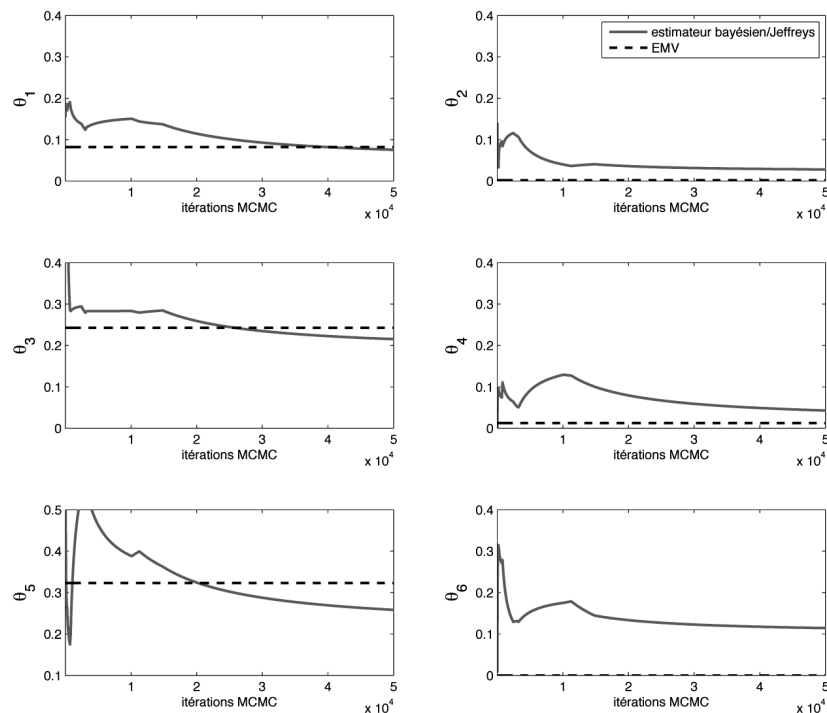


Figure 3. Données réelles : Comparaison entre l'EMV et les itérations de l'algorithme MCMC pour le calcul de l'estimateur bayésien.

4. Tests numériques

Sur des exemples à 2 états (non présentés ici), l'estimateur bayésien donne de meilleurs résultats que l'EMV. Pour 4 états et dans le cas de données simulées selon le modèle présenté dans cet article, ces deux estimateurs donnent des résultats comparables (cf. Figure 2). Des essais systématiques sont en cours afin de comparer ces estimateurs. Dans le cas de données réelles, l'estimateur bayésien et l'EMV donnent des résultats différents (cf. Figure 3). Dans ces conditions, il semble nécessaire de faire évoluer le modèle considéré. Nous développons ce point dans la section suivante.

5. Perspectives

Dans ce travail nous avons considéré un modèle de successions agraires à 4 états. Nous avons fait deux hypothèses : (i) les parcelles ont les mêmes dynamiques et leurs dynamiques sont indépendantes entre elles ; (ii) ces dynamiques sont markoviennes. La matrice de transition comprend 6 paramètres que l'on a estimés à l'aide de deux estimateurs : l'estimateur du maximum de vraisemblance (EMV) et l'estimateur bayésien fondé sur la loi a priori de Jeffreys.

Lors d'essais numériques (non présentés ici) avec 2 paramètres correspondant à un modèle markovien à deux états, l'estimateur bayésien se comporte nettement mieux que

l'EMV. Dans le cas à 6 paramètres, les deux estimateurs semblent donner des résultats comparables. Des essais numériques plus systématiques sont en cours.

L'hypothèse (i) (parcelles indépendantes) est clairement violée. Il est en effet plus réaliste de considérer la dynamique d'une parcelle comme dépendante de l'état des parcelles voisines. Prendre en compte une dynamique spatio-temporelle est plus naturel mais nécessite beaucoup plus de paramètres.

L'hypothèse (ii) pourrait être assouplie. Au vu des données [3], l'état initial "forêt" devrait nécessiter un traitement particulier. Dans un modèle markovien, les temps de séjour en chacun des états suit une loi géométrique. Le temps de séjour dans cet état initial n'a pas de raison de suivre une loi géométrique : certaines parcelles plus récemment exploitées présentent de plus longues suites de F. Il serait donc nécessaire soit de commencer les historiques des parcelles au moment où elles quittent l'état F, soit de proposer une loi non géométrique pour le temps de séjour dans l'état F. Dans ce dernier cas on se placerait dans un cadre semi-markovien. Une étude est actuellement en cours.

Par ailleurs l'idée de faire appel à des modèles de Markov d'ordre 2 n'est pas réaliste dans notre cadre, elle augmenterait de façon significative le nombre de paramètres à estimer.

En conclusion il nous semble intéressant de développer un modèle dynamique de parcelle prenant en compte l'état des parcelles voisines. Afin de limiter le nombre de paramètres à estimer, cela devra s'appuyer sur une étude en amont portant sur les logiques d'exploitation des acteurs concernés. Dans le cas de périodes d'observations durant lesquelles des changements exogènes ont été constatés, comme des réformes agraires, il semble pertinent de faire appel à des modèles de Markov cachés.

Remerciements : Les auteurs souhaitent remercier Dominique Hervé (IRD) pour avoir attiré notre attention sur les problèmes de transitions agraires du corridor forestier de Ranomafana-Andringitra de la province de Fianarantsoa.

6. Bibliographie

- [1] Souad Assoudou and Belkheir Essebbbar. A Bayesian model for binary Markov chains. *International Journal of Mathematics and Mathematical Sciences*, 8 :421–429, 2004.
- [2] Patrick Billingsley. *Statistical inference for Markov processes*, Statistical Research Monographs, Vol. II. The University of Chicago Press, Chicago, Ill. 1961.
- [3] Fabien Campillo, Nicolas Desassis, Angelo Raheiririna, Rivo Rakotozafy. Markovian modeling of agricultural successions. En préparation.
- [4] M.D. Usher. Modelling ecological succession, with particular reference to Markovian models. *Plant Ecology*, 46 :11–18, 1981.
- [5] Paul E. Waggoner and George R. Stephens. Transition probabilities for a forest. *Nature*, 225 :1160–1161, 1970.