Inferring land use dynamics by semi-Markov model

Angelo Raherinirina, Dominique Hervé, Fabien Campillo

Inferring land use dynamics

by semi-Markov model

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

[1] University of Fianarantsoa, BP 1264, Andrainjato, 301 Fianarantsoa, Madagascar
angelo_raherinirina@yahoo.fr

[2] UMR220 GRED (IRD/Montpellier3), Montpellier, France

[3] Inria, Modemic Team Projet (Inria/Inra), 2, place Pierre Viala, 34060 Montpellier, France

ABSTRACT. We propose land use dynamics models corresponding to parcels located on the edge of the forest corridor, Madagascar. We use semi-Markov chain to infer the land-use dynamics. In addition to the empirical and maximum likelihood methods, we estimate the semi-Markov kernel by a Bayesian approach. In the latter case, we use Jeffreys' non-informative prior and we approximate the posterior distribution by Monte Carlo Markov Chain (MCMC) approximation. These three estimation methods lead to three different models, two are absorbing and one is regular. We study the asymptotic behavior of these models. We have determined the time scales of the considered land-use dynamics.

RÉSUMÉ. On considère des modèles de dynamique d'usages des sols correspondant aux utilisations de parcelles situées sur le bord du couloir forestier reliant les deux parcs nationaux de Ranomafana et d'Andringitra à Madagascar. Nous utilisons des modèles semi-markoviens pour inférer la dynamique d'usage des sols. En plus de la méthode empirique et de la méthode du maximum de vraisemblance, nous estimons le noyau semi-markovien par une méthode bayésienne. Dans ce dernier cas, la loi a priori utilisée est la loi non-informative de Jeffreys et l'approximation de la loi a posteriori associée se fait avec la méthode de Monte Carlo par chaîne de Markov (MCMC). Ces trois techniques conduisent à trois modèles différents dont deux absorbants et un régulier. Nous étudions les comportements asymptotiques de ces trois modèles. En termes applicatifs, nous avons pu déterminer l'échelle de temps de la dynamique dúsage des sols considérée.

KEYWORDS : semi-Markov models, maximum likelihood, Bayesian statistic, land use dynamics

MOTS-CLÉS : modèles semi-markoviens, maximum de vraisemblance, statistique bayésienne, dynamique d'usage des sols

1. Introduction

We observe land use dynamics by the succession of parcels occupations. The state of a given parcel corresponds to its use in a given time. Inferring such process returns to identify the law of the successions of the state during the observation period. Markov models are among the most requested tools in this kind of situation. In [3], we used Markov chain to modelize land use dynamics corresponding to the parcels on the edge of the forest corridor-Ranomafana-Andringitra, Madagascar.

Compared to [3], the dataset used in this paper is larger and has more states (see Figure 1). It contains seven states: " natural forest" ($\mathtt{Fn}$)", " annual crop" ($\mathtt{Ca}$), " fallow" ($\mathtt{Fa}$) , " perennial crop" ($\mathtt{Cp}$), " grass" ($\mathtt{Gr}$), "paddy field" ($\mathtt{Pf}$) and " secondary forest" ($\mathtt{Fs}$). The two states: " natural forest" and " paddy field" have not the same dynamics as the others and we have decided to omit them in our models. They remains five states:

$$E = \{\mathtt{Ca}, \mathtt{Fa}, \mathtt{Gr}, \mathtt{Fs}, \mathtt{Cp}\}.$$

In [4], we propose a parametric bootstrap method to test the fit between a Markov chain model with the dataset. This test concerns the distributions of the sojourn time on each state which should be necessarily a geometric distribution. According to this statistical tests: the distribution of the sojourn time in each state are not all geometric, and the distribution of the sojourn time on the state "fallow" depends on the next state visited. Hence, the classical Markov model is not suitable in this situation and the use of semi-Markov models is natural[2].

The semi-Markov model is a generalization of Markov chain model by removing the constraints of the distribution of the sojourn time. It allows more flexibility in the distribution in order to provide models closer to the reality.

We denote $\left(e^{(p)}_{n=1:N_p}\right)_{p=1:P}$, $e^{(p)}_n \in E$ the observations of the P parcels (Figure 1), where N_p is the length of the observation. Let $(t^{(p)}_k)_{k\in\mathbb{N}}$ the kth jump time for the parcel p and $(s^{(p)}_k)_{k\in\mathbb{N}}$ the sequence of sojourn time:

$$t^{(p)}_k = \sum_{l=0}^{k} s^{(p)}_l.$$

Finally, we denote

$$M^{(p)} = \max\{k \in \mathbb{N}, t^{(p)}_k \leq N_p\}$$

the number of jumps at time N_p and $u_{N_p} = N_p - t_{N_p}$ the censured sojourn time on the final state $e^{(p)}_{M^{(p)}}$.

Discussions with specialists, along with the results of some statistical tests allowed us to make the following assumptions:

(h_1) *The parcels are independent and have the same dynamics.*

(h_2) *The process* $(Y_k)_{k\in\mathbb{N}}$ *which record the successive visited states is a Markov chain.*

(h_3) *The possible transitions are* CaFa, CaGr, CaCp, FaCa, FaFs, GrCa, GrFa, CpCa *and* FsCa. *Other transitions are unrealistic.*

(h_4) *The sojourn time of each state depends on the next visited state. The distributions of the sojourn time associated with the transitions* CaFa, CaGr , CaCp,*and* CaFa, *are geometric. The distributions of the sojourn time for other transitions is the Poisson distribution.*

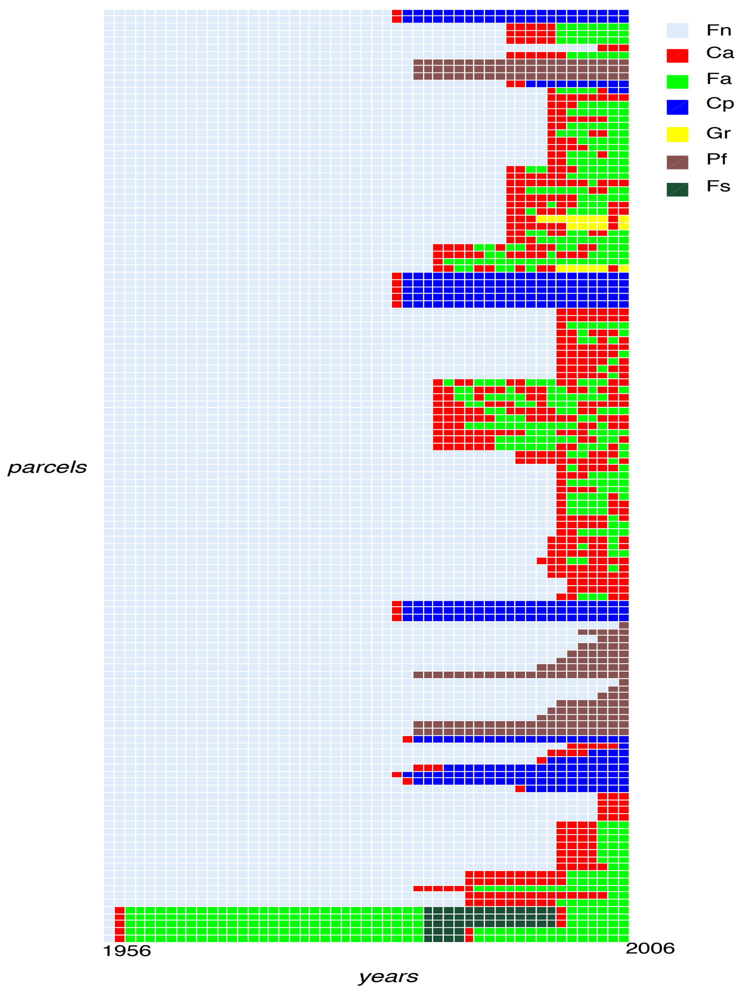

Figure 1. *Annual state of the 131 parcels on the western edge of the corridor Ranomafana-Andringitra for 50 years. We observed 7 states: natural forest* (`Fn`)*, annual crop* (`Ca`)*, fallow* (`Fa`)*, perennial crop* (`Cp`)*, grass* (`Gr`)*, paddy field* (`Pf`) *and secondary forest* (`Fs`).

In other words, we assume that the observations $(e^{(p)}_{n=1:N_P})_{p=1:P}$ are derived from a semi-Markov chain $(X_n)_{n\in\mathbb{N}}$.

Let $(Y_k)_{k\in\mathbb{N}}$ the embedded Markov chain, $(T_k)_{k\in\mathbb{N}}$ the successive jump time and $(S_k)_{k\in\mathbb{N}}$ the sojourn time. The transition matrix of $(Y_k)_{k\in\mathbb{N}}$ is

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

where

$$\theta = (\theta_i)_{1\leq i\leq 4} \in \Theta = \{\theta \in [0,1]^4; \theta_1+\theta_2 \leq 1\}$$

and the conditional sojourn time distribution is:

$$f(k) = \begin{pmatrix} 0 & \mathcal{G}(\gamma_1,k) & \mathcal{G}(\gamma_2,k) & 0 & \mathcal{G}(\gamma_3,k) \\ \mathcal{G}(\gamma_4,k) & 0 & 0 & \mathcal{P}(\lambda_1,k) & 0 \\ \mathcal{P}(\lambda_2,k) & \mathcal{P}(\lambda_3,k) & 0 & 0 & 0 \\ \mathcal{P}(\lambda_4,k) & 0 & 0 & 0 & 0 \\ \mathcal{P}(\lambda_5,k) & 0 & 0 & 0 & 0 \end{pmatrix} \tag{2}$$

where $\mathcal{G}(\gamma_i, .)$ is the geometric distribution with parameter $\gamma_i, \ , \ 1 \leq i \leq 4$ and $\mathcal{P}(\lambda_j, .)$ is the Poisson distribution with parameter $\lambda_j, \ 1 \leq j \leq 5$.

We get the semi-Markov kernel q by

$$\begin{aligned} \forall n \in \mathbb{N}, \ q_{ij}(n) &= \mathbb{P}(Y_k = j, \ S_k = n | Y_{k-1} = i) \\ &= Q_{ij} \, f_{ij}(n), \ \forall n > 0, \text{ and } i \neq j \end{aligned} \tag{3}$$

with $q_{ij}(0) = 0$, for all $i, j \in E$.

2. Absorbing models derived from empirical estimate and maximum likelihood estimate

2.1. Empirical estimates

The empirical estimates of the transition matrix of $(Y_k)_{k\in\mathbb{N}}$ and the conditional sojourn time distribution are:

$$\bar{Q}_{ij} = \frac{n_{ij}}{n_i}, \quad \bar{f}_{ij}(n) = \frac{n_{ij}(k)}{n_{ij}}, \ i, j \in E, \ k \in \mathbb{N},$$

where n_{ij} the number of transitions from i to j, $n_{ij}(k)$ the number of transitions from i to j, with sojourn time in state i equal to k, and n_i the number of visit to state i. If $n_i = 0$, we set $\bar{Q}_{ij} = 0$ and if $n_{ij} = 0$, we set $\bar{f}_{ij}(k) = 0$, for all $i, j \in E$.

2.2. Maximum likelihood estimates

Maximum likelihood and empirical methods coincide in the estimation of the transition matrix of the embedded Markov chain:

$$\begin{aligned} \hat{\theta}_1 &= \frac{n_{\texttt{CaFa}}}{n_{\texttt{CaFa}} + n_{\texttt{CaGr}} + n_{\texttt{CaCp}}}, & \hat{\theta}_2 &= \frac{n_{\texttt{CaGr}}}{n_{\texttt{CaFa}} + n_{\texttt{CaGr}} + n_{\texttt{CaCp}}}, \\ \hat{\theta}_3 &= \frac{n_{\texttt{FaCa}}}{n_{\texttt{FaCa}} + n_{\texttt{FaFs}}}, & \hat{\theta}_4 &= \frac{n_{\texttt{GrCa}}}{n_{\texttt{GrCa}} + n_{\texttt{GrFa}}}. \end{aligned}$$

They correspond to a matrix of an absorbing Markov chain:

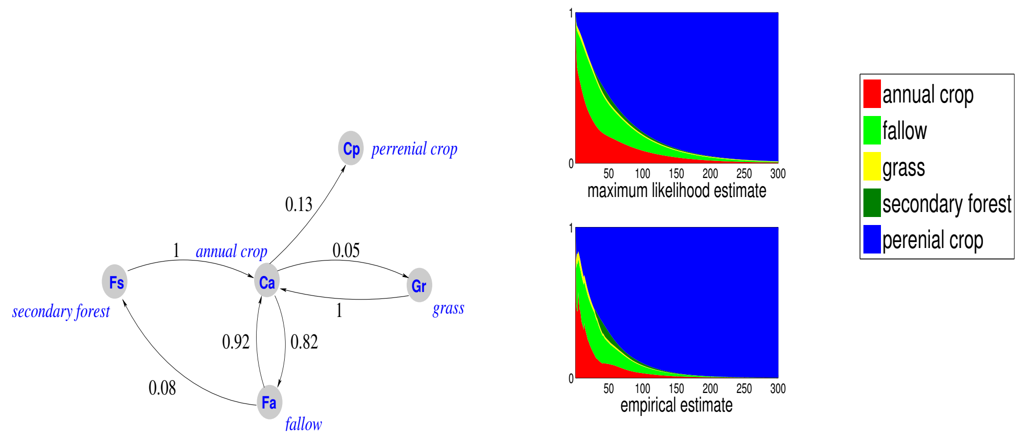

Figure 2. **Left:** *Embedded Markov chain of the absorbing semi-Markov deduced from empirical and maximum likelihood methods.* **Right:** *Distributions of the absorption times deduced from the empirical and maximum likelihood estimates.*

The maximum likelihood estimates(MLEs) of the parameters of the sojourn times distribution f are:

$$\hat{\gamma_1} = \frac{\sum_{k=1}^{M} n_{\texttt{CaFa}}(k)}{\sum_{k=1}^{M} k\, n_{\texttt{CaFa}}(k)}, \quad \hat{\gamma_2} = \frac{\sum_{k=1}^{M} n_{\texttt{CaGr}}(k)}{\sum_{k=1}^{M} k\, n_{\texttt{CaCp}}(k)}, \quad \hat{\gamma_3} = \frac{\sum_{k=1}^{M} n_{\texttt{CaCp}}(k)}{\sum_{k=1}^{M} k\, n_{\texttt{CaCp}}(k)},$$

$$\hat{\gamma_4} = \frac{\sum_{k=1}^{M} n_{\texttt{FaCa}}(k)}{\sum_{k=1}^{M} k\, n_{\texttt{FaCa}}(k)}, \quad \hat{\lambda}_1 = \frac{\sum_{k=1}^{M} k\, n_{\texttt{FaFs}}(k)}{\sum_{k=1}^{M} n_{\texttt{FaFs}}(k)}, \quad \hat{\lambda}_2 = \frac{\sum_{k=1}^{M} k\, n_{\texttt{GrCa}}(k)}{\sum_{k=1}^{M} n_{\texttt{HeCa}}(k)},$$

$$\hat{\lambda}_3 = \frac{\sum_{k=1}^{M} k\, n_{\texttt{GrFa}}(k)}{\sum_{k=1}^{M} n_{\texttt{GrFa}}(k)}, \quad \hat{\lambda}_4 = \frac{\sum_{k=1}^{M} k\, n_{\texttt{FsCa}}(k)}{\sum_{k=1}^{M} n_{\texttt{FsCa}}(k)}, \quad \hat{\lambda}_5 = \frac{\sum_{k=1}^{M} k\, n_{\texttt{CpCa}}(k)}{\sum_{k=1}^{M} n_{\texttt{CpCa}}(k)},$$

with $M = max(M^{(p)}, p = 1 \dots P)$.

These two estimations of f lead to two absorbing semi-Markov models with semi-Markov kernel estimates:

$$\bar{q}_{ij}(n) = \bar{Q}_{ij}\ \bar{f}_{ij}(n) \quad \forall i, j \in E, \forall n \in \mathbb{N},$$

$$\hat{q}_{ij}(n) = \hat{Q}_{ij}\ \hat{f}_{ij}(n) \quad \forall i, j \in E, \forall n \in \mathbb{N}.$$

2.3. Absorbing time

We focus on the time taken by the system before reaching the absorbing state "perennial crop" Cp (see Figure 2). Let $T_i^{\texttt{Cp}}$ the hitting time of Cp starting from a state $i \in E \setminus \{\texttt{Cp}\}$,ie.

$$T_i^{\texttt{Cp}} \stackrel{\text{def}}{=} T_m, \text{ with } m = \min\{l \in \mathbb{N},\ Y_l = \texttt{Cp},\ Y_0 = i\}.$$

We note $g_i^{\texttt{Cp}}(n) \stackrel{\text{def}}{=} \mathbb{P}_i(T_i^{\texttt{Cp}} = n), \quad n \in \mathbb{N}$ the absorption probability at time n starting from a state i with $g_i^{\texttt{Cp}}(0) = 0$, $\forall\, i \in E \setminus \{\texttt{Cp}\}$.

The mean $m_{i\mathtt{Cp}}$ of the absorption time satisfies

$$m_{i\mathtt{Cp}} = \nu_i + \sum_{k \neq \mathtt{Cp}} Q_{ik}\, m_{k\mathtt{Cp}} \tag{4}$$

where ν_i is the mean sojourn time of each state $i \in E \setminus \{\mathtt{Cp}\}$.

By (4), we can write

$$m_{\mathtt{CaCp}} = \frac{\nu_{\mathtt{Ca}} + Q_{12}\nu_{\mathtt{Fa}} + Q_{13}\nu_{\mathtt{Gr}} + Q_{12}Q_{24}\nu_{\mathtt{Fs}}}{1 - (Q_{13}Q_{31} + Q_{12}Q_{21} + Q_{12}Q_{24}Q_{41})}.$$

Calculated with the empirical and maximum likelihood estimates, the means of the absorption time are:

$$\bar{m}_{\mathtt{CaCp}} = 39 \text{ years}$$

$$\hat{m}_{\mathtt{CaCp}} = 65 \text{ years.}$$

3. Regular model deduced from the Bayesian estimate

A Bayesian approach is proposed in [3] to estimate the transition matrix of a discrete time Markov chain . In this section,we use the same technique to estimate the transition matrix of the embedded Markov chain $(Y_k)_{k\in\mathbb{N}}$, then we compute the Bayesian estimates of the sojourn time distribution. In both cases, the calculations of the posterior distributions will be done by a Monte Carlo Markov Chain (MCMC) approximation.

3.1. Transition matrix of embedded Markov chain

We assume that the parameter θ of the transition matrix (1) is a random variable with probability density $\pi(\theta)$. Prior knowlegde on θ is summarized in a prior distribution $\pi_{\text{prior}}(\theta)$; we consider the non informative Jeffreys' prior

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

where $\mathcal{I}(\theta)$ is the matrix of Fisher Information.

We deduce the posterior distribution by the Bayes formula $\pi_{\text{post}}(\theta) \propto L_1(\theta)\ \pi_{\text{prior}}(\theta)$, where $L_1(\theta)$ is the likelihood function associated with the parameter θ.

And the Bayes estimates is

$$\tilde{\theta} = \int_{\Theta} \theta\, \pi_{\text{post}}(\theta)\, \mathrm{d}\theta = \frac{\int_{\Theta} \theta\, L_1(\theta)\, \pi_{\text{prior}}(\theta)\, \mathrm{d}\theta}{\int_{\Theta} L_1(\theta)\, \pi_{\text{prior}}(\theta)\, \mathrm{d}\theta}\,. \tag{5}$$

Computing (5) can not be explicit, we resolve it by using a MCMC algorithm (see [3]).The corresponding result is a regular matrix $\tilde{Q}$.

3.2. Distribution of the sojourn times

The Jeffreys priors of the parameters $\gamma_i,\ 1 \le i \le 4$ and $\lambda_j,\ 1 \le j \le 5$ are

$$\pi^{g}_{\text{prior}}(\gamma_i) \propto \sqrt{\frac{n}{\gamma_i^2}\ \left(\frac{1}{(1-\gamma_i)}\right)}, \qquad\qquad \pi^{p}_{\text{prior}}(\lambda_j) \propto \sqrt{\frac{n}{\lambda_j}}.$$

The posterior distributions and the Bayes estimates are:

$$\pi^g_{\text{post}}(\gamma_i) \propto L_g(\gamma_i)\,\pi^g_{\text{prior}}(\gamma_i), \qquad \pi^p_{\text{post}}(\lambda_j) \propto L_p(\lambda_j)\,\pi^p_{\text{prior}}(\lambda_j),$$
$$\tilde{\gamma_i} = \int \gamma_i\,\pi^g_{\text{post}}(\gamma_i)\,\mathrm{d}\gamma_i, \qquad \tilde{\lambda_j} = \int \lambda_j\,\pi^p_{\text{post}}(\lambda_j)\,\mathrm{d}\lambda_j.$$

3.3. Limit distribution of the semi-Markov model

Bayesian estimate of the transition matrix is a regular matrix. The embedded Markov chain associated is irreducible and admits an invariant distribution π. Further, the distribution of the sojourn times on each state have a finite mean. Thus, the semi-Markov chain $(X_n)_{n\in\mathbb{N}}$ has a limit distribution [2]:

$$\Pi_j = \frac{\pi_j\,\nu_j}{\sum_{i\in E}\pi_i\,\nu_i},\ j \in E \tag{6}$$

where π is the invariant distribution of the embedded Markov chain $(Y_k)_{k\in\mathbb{N}}$ and ν the means of the sojourn times in each state. We find:

$$\Pi = (0.1490,\ 0.2134,\ 0.0524,\ 0.0379,\ 0.5473). \tag{7}$$

This distribution represents the proportion of the parcel occupied by each state in equilibrium, see figure. That is, at equilibrium: 14.9% in state `Ca`, 21.34% in `Fa`, 5.24% in `Gr`, 3.79% in `Fs` and 54.73% in `Cp`; see figure 3.

4. Conclusion

In this paper,we propose an inferring approach for land use dynamics by semi-Markov models. Assumptions about the independence of parcel, equality of the dynamics for all parcels and the homogeneity of the dynamics were posed. These assumptions are essentials in the construction of models. In our case, the inference of the dynamics returns to estimate the semi-Markov transition matrix which directs the successions of the different land uses: annual crops, fallow, grass, secondary forest and perennial crop. This matrix is obtained by the transition matrix of the embedded Markov chain Q and the conditional sojourn time distribution f.We estimate Q and f by empirical, maximum likelihood and Bayesian method.

The first two methods consist in using only the observed data, without taking into account any priori knowledge about the system. They lead to an absorbing model, that is: according to these models, in long term, the parcel use converges to an unique land use : perennial crop. We have calculated the mean of the time to reach this state: 39 years for the model deduced by empirical estimate and 65 years for the model deduced by the maximum likelihood estimate.

Bayesian approach is considered more credible by the specialists because it allows an integration of the available prior knowledge of the system. We use Jeffreys priors and the identifications of the parameters of Q and f are done with a Markov chain Monte Carlo approximation. This method leads to a regular model which has an invariant distribution. This invariant distribution ensures the stabilization of the proportions evolution of the land use in the long term. Making comparisons on the absorbing and regular models: we find that they are very close in the short term. The difference is clear after some decades.

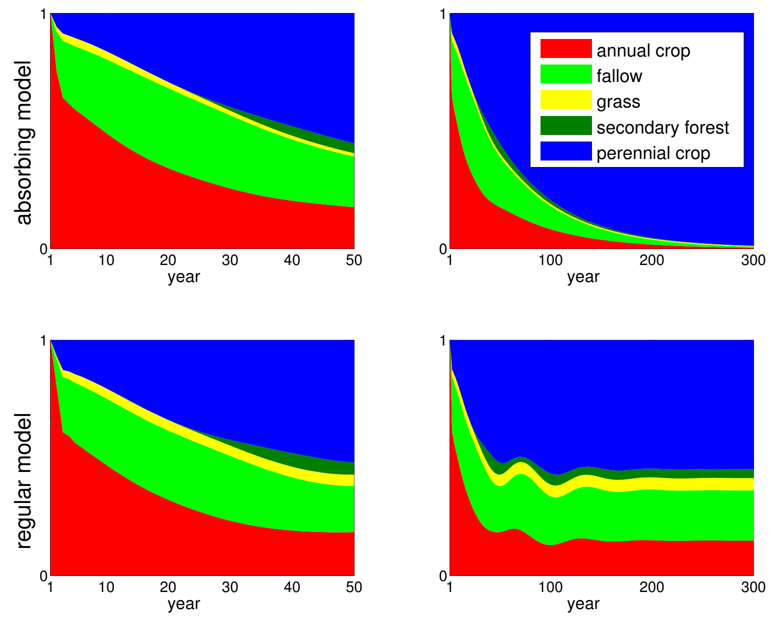

Figure 3. *Comparing proportions of parcel use according to the absorbing model (top) and the regular (bottom) models.*

We put in evidence in this work that Markov models are relevant tools to modelize land use dynamics. These models are relatively simple but allow for thorough analysis due to their mathematical properties. We can apply the tools developed in this article to various area related applications.

5. References

[1] 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.

[2] V. S. Barbu, J. Bulla, and A. Maruotti. Estimation of the stationary distribution of a semi-Markov chain. *Journal of Reliability and Statistical Studies*, 5:15–26, 2012.

[3] F. Campillo, Hervé D., A. Raherinirina, and R. Rakotozafy. Markov analysis of land use dynamics: A case study in Madagascar,. In *XI Colloque Africain sur la Recherche en Informatique en Mathématiques Appliqués (CARI)*, 2012.

[4] A. Raherinirina. *Modélisation markovienne des dynamiques d'usage des sols – Cas de parcelles situées sur le bord du corridor forestier Ranomafana-Andringitra*. Thèse soutenue à l'Université de Fianarantsoa en 2013.

[5] 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.