HAL
open science

Modélisation markovienne des dynamiques d'usage des sols. Cas des parcelles situées sur le bord du corridor forestier Ranomafana-Andringitra

Angelo Raherinirina

To cite this version:

Angelo Raherinirina. Modélisation markovienne des dynamiques d'usage des sols. Cas des parcelles situées sur le bord du corridor forestier Ranomafana-Andringitra. Modélisation et simulation. Université de Fianarantsoa, 2013. Français. NNT : . tel-00870679

HAL Id: tel-00870679

https://theses.hal.science/tel-00870679v1

Submitted on 7 Oct 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université de Fianarantsoa

École Nationale d'Informatique

Thèse de doctorat

Spécialité : Mathématiques Appliquées

présentée par

Angelo Fulgence Raherinirina

Modélisation markovienne des dynamiques d'usage des sols

Cas des parcelles situées sur le bord du corridor forestier Ranomafana–Andringitra

---

Soutenue le 02 Août 2013 devant le jury composé de :

Manantsoa Victor	Président
Solo Randriamahaleo	Examinateur
Venot Ratiarson	Examinateur
Josvah Paul Razafimandimby	Rapporteur (interne)
Nicolas Picard	Rapporteur (externe)
André Totohasina	Rapporteur (externe)
Fabien Campillo	Directeur de thèse
Rivo Rakotozafy	Co-directeur de thèse

Remerciements

Cette thèse n'aurait pu être mené à bien sans l'aide précieuse de plusieurs personnes qui ont contribué de près ou de loin à l'aboutissement de ce travail de recherche, je tiens à les remercier maintenant.

En premier lieu, je tiens à remercier mon directeur de thèse, Fabien Campillo. Son expertise, son sens de perfectionnement et sa disponibilité m'ont permis d'avoir un encadrement de qualité. Fabien n'a jamais hésité de prendre les temps de m'expliquer, discuter, échanger et partager ses grandes expertises dans la recherche. Grâce à lui, j'ai pu terminer avec succès mes trois années de thèse. Je ne le remercierai jamais assez pour tout ce qu'il a fait pour moi. Je tiens à lui témoigner ma profonde gratitude.

Ces remerciements sont à partager avec Rivo Rakotozafy, mon co-directeur de thèse, pour son encadrement et ses conseils pendant ces trois années de thèse. Il a toujours su se rendre disponible pour me guider et m'éclaircir sur notre problème de recherche. Je lui dois beaucoup.

J'exprime aussi mes sincères gratitude à l'égard de Dominique Hervé, chercheur à l'IRD pour son implication dans mon travail de thèse. Ils m'a permis d'accéder à des problèmes de dynamiques agraires, et ce, depuis mon stage de DEA. Qu'il trouve ici l'expression de toute ma reconnaissance.

Je souhaite également vivement remercier Nicolas Picard et André Totohasina qui m'ont fait le grand plaisir d'être les rapporteurs de mon travail de thèse. Vos remarques et appréciations m'ont permis d'améliorer ce manuscrit.

Je suis très sensible à l'honneur que m'a fait Manantsoa Victor qui a accepté de présider le jury de ma soutenance. Je le remercie vivement pour sa disponibilité malgré ses différentes responsabilités.

Je remercie également Josvah Paul Razafimandimby, Venot Ratiarson et Solo Randriamahaleo qui ont accepté de faire partie des membres de jury de ma soutenance de thèse. Ses critiques constructives m'ont particulièrement bénéfiques dans l'aboutissement de ce travail.

Mes remerciements vont également aux différentes personnalités de l'université de Fianarantsoa, qui ont participé au bon déroulement de mes recherches. Je tiens à remercier plus particulièrement, Fontaine Rafamatanantsoa, Directeur de l'École Nationale d'Informatique, Clarisse Rasoamampionona, Directeur de l'École Normale Supérieure de Fianarantsoa et Monique Rasoazananiera, Présidente de l'Université de Fianarantsoa.

Ce travail de thèse a été réalisé au sein de l'équipe-projet MODEMIC INRA-INRIA de l'UMR Mistea à Montpellier. Je remercie tous les membres de l'équipe pour leur accueils chaleureux. Je remercie plus particulièrement le directeur de l'Unité, Alain Rapaport, ainsi que les collègues doctorants qui ont échangé avec

moi pendant mes séjours de recherche au seins de l'équipe.

J'exprime aussi mes sincères remerciement à ceux qui m'ont aidé financièrement pour bien mener mon travail de recherche, en particulier le Service de Coopération et d'Action Culturelle de l'Ambassade de France à travers la bourse BGF, l'INRIA Sophia Antipolis et l'équipe STIC-Madagascar du LIRIMA. Leur contributions m'ont permis d'effectuer plusieurs séjours au sein de mon laboratoire d'accueil.

Un grand merci à mes parents, ma famille et tous mes amis pour leur soutien infaillible. Ils m'ont encouragé tout au long de mes études et je leur en suis très reconnaissant.

J'ai une pensée toute particulière pour ma fiancée, Laingoarivony Hasy, qui m'a soutenue durant toutes ces années de thèse. Je la remercie pour son amour et sa patience. Que notre grand jour arrive.

Je termine cette remerciement par louer l'Éternel, Dieu tout puissant pour les bénédictions qu'Il m'a accordé. Grâce à lui, j'ai eu la force et le courage nécessaire pour mener à terme ces trois années d'études. *À lui seul soit la gloire.*

Modélisation markovienne des dynamiques d'usage des sols : Cas de parcelles situées sur le bord du corridor forestier Ranomafana–Andringitra

Résumé : Nous proposons une démarche markovienne d'inférence et de modélisation de dynamiques agraires dans le cadre d'usage de parcelles situées en lisière du corridor forestier reliant les deux parcs nationaux de Ranomafana et d'Andringitra. La préservation de la forêt de la côte est de Madagascar est cruciale, il est donc pertinent de développer des outils permettant de mieux comprendre les dynamiques de déforestation puis de l'usage des parcelles et enfin de leur éventuel retour à l'état de forêt. Nous nous appuyons sur deux jeux de données de terrain établis par l'IRD. Dans ce genre d'étude, une étape préliminaire consiste à construire la matrice de transition empirique, cela s'apparente donc à une modélisation markovienne de la dynamique. Dans ce cadre nous considérons l'approche par maximum de vraisemblance et l'approche bayésienne. Cette dernière approche nous permet d'intégrer des informations non-présentes dans les données mais reconnues par les spécialistes, elle fait appel à des techniques d'approximation de Monte Carlo par chaînes de Markov (MCMC). Nous étudions les propriétés asymptotiques des modèles obtenus à l'aide de ces deux approches et notamment le temps de convergence vers la loi quasi-stationnaire dans le premier cas et vers la loi stationnaire dans le second. Nous testons différentes hypothèses portant sur les modèles. Cette approche markovienne n'est plus valide sur le deuxième jeu de données, plus étendu, où il a fallu faire appel à une approche semi-markovienne : les lois des temps de séjour dans un état donné ne sont plus nécessairement géométriques et peuvent dépendre de l'état suivant. À nouveau nous faisons appel aux approches par maximum de vraisemblance et bayésienne. Nous étudions le comportement asymptotique de chacun de ces modèles. En termes applicatifs, nous avons pu déterminer les échelles de temps de ces dynamiques.

Mots-clefs : Modèles markoviens, Modèles semi-markoviens, maximum de vraisemblance, statistique bayésienne, méthode de Monte Carlo par chaînes de Markov (MCMC), dynamiques agraires.

Markov modeling of the dynamics of land uses: Case of plots on the edge of the Ranomafana-Andringitra forest corridor

Abstract: We consider a Markovian inference and modeling approach applied to land-use dynamics within the context of parcels of land on the edge of the forest corridor linking the two national parks of Ranomafana and Andringitra. Preservation of the forest on the east coast of Madagascar is crucial, so it is important to develop tools for inferring the dynamics of plots deforestation, of their use and of their possible return to the state forest. Our studies rely on two sets of field data established by IRD. The usual first step in such studies, which is to build the empirical transition matrix, is similar to a Markovian modeling of the dynamics. In this context, we consider the maximum likelihood approach and the Bayesian approach. This latter approach allows us to integrate information not present in the data but accredited by specialists, this approach uses Monte Carlo Markov Chain (MCMC) approximation techniques. We study the asymptotic properties of the models deduced from these two approaches, including the convergence time to the quasi-stationary distribution in the first case and to the stationary distribution in the second. We test various hypotheses on the models. The Markovian approach is no longer valid on the second larger data set where he had to use a semi-Markov models: the distributions of the sojourn times on given states are not necessarily geometric and may depend on the next state. Again we use maximum likelihood and Bayesian approaches. We study the asymptotic behavior of each of these models. In terms of applications, we have determined the time scales of these dynamics.

Keywords: Makov models, semi-Markov models, maximum likelihood, Bayesian statistics, Monte Carlo Markov chain (MCMC), land-use dynamics.

Notations

- $X_{1:n}$ désigne la suite $(X_1, \ldots, X_n)$.
- $X \sim Q$ signifie que la variable aléatoire X suit la loi Q.
- $X_{1:n} \overset{\text{iid}}{\sim}$ signifie que les variables X_i sont indépendantes et identiquement distribuées de loi Q.
- $1_A(X)$, fonction indicatrice de l'ensemble A

$$1_A(X) = \begin{cases} 1 & \text{si} \quad X \in A, \\ 0 & \text{sinon.} \end{cases}$$

- δ_{ij} est le symbole de Kronecker : $\delta_{ij} = 1$ si $i = j$ et 0 sinon.
- $f(x) \propto g(x)$ signifie que les fonctions $f(x)$ et $g(x)$ sont proportionnelles.

Table des figures

1.1 *Forêt dense et humide à l'est de Madagascar* 2

2.1 Zone d'étude : Village d'Ambendrana 6
2.2 Dynamique d'occupation des parcelles dans la village d'Ambendrana . 7
2.3 Jeu de données n° 2 : Usages des 131 parcelles observées pendant 50 ans (1956-2006) 10

3.1 Graphe d'une chaîne de Markov non irréductible 18
3.2 Graphe d'une chaîne de Markov à 3 classes 21

4.1 Modèle markovien d'usage des sols à 3 états 37
4.2 Modèle régulier à 3 états 40
4.3 Estimateur bayésien et EMV des paramètres de la matrice de transition du modèle markovien 44
4.4 Modèle à 2 états 45
4.5 Lois empiriques des erreurs d'estimations 48
4.6 Modèle markovien absorbant 49
4.7 Lois du temps d'absorption 49
4.8 Estimateur bayésien de la matrice de transition du modèle de Markov 51
4.9 Proportions d'usages des parcelles pour tous les états 54
4.10 Proportions des parcelles occupées par les états `Ca` `Jc` 55

5.1 Densités empiriques des temps de séjour 60
5.2 Résultats du test d'adéquation des temps de séjour 60
5.3 Test d'homogénéité en espace selon l'année de mise en culture 63
5.4 Test d'homogénéité en espace selon l'exposition 64
5.5 Test d'homogénéité en temps 66

6.1 Trajectoire d'une chaîne semi-markovienne 68

6.2 *Modèle semi-markovien à trois états.* 77
6.3 Loi de transition d'un modèle semi-markovien à 3 états 79
6.4 Convergence d'une loi de transition semi-markovienne vers son équilibre 80

7.1 Test d'adéquation des lois des temps de séjour à la loi géométrique . . 82
7.2 Test d'adéquation de la loi des temps de séjour en "jachère" 83
7.3 Modèle semi-markovien à 5 états 85
7.4 Chaîne de Markov incluse 85
7.5 Chaîne incluse des modèles semi-markoviens absorbants 100
7.6 Estimateur empirique et EMV des lois de temps de séjour 101
7.7 Proportions d'usages des parcelles calculées par l'EMV et l'estimateur empirique 102
7.8 Absorption du processus par l'état absorbant `Cp` 104
7.9 MCMC pour la matrice de transition de la chaîne incluse 106
7.10 MCMC pour les temps de séjour géométriques 107
7.11 MCMC pour les temps de séjour poissonniens 108
7.12 Lois de transitions semi-markovinne pendant les 50 années d'observations 109
7.13 *Loi limite du modèle semi-markovien régulier déduit de l'estimateur bayésien* 110
7.14 Lois limites des lois semi-markoviennes 111
7.15 Proportion à l'équilibre 112
7.16 Evolution des proportions d'usages des parcelles sans l'état absorbant 113
7.17 Proportion des parcelles en "jachère/non jachère" 114

Liste des algorithmes

3.1 *Algorithme de Metropolis.* 32
3.2 *Algorithme de Metropolis-Hastings* 33
4.1 Algorithme MCMC loi a posteriori pour le modèle markovien . . . 43
5.1 Test d'adéquation par bootstrap paramétrique 59
7.1 MCMC Metropolis-Hastings pour la matrice de transition de la chaîne incluse 96
7.2 MCMC pour les temps de séjour géométriques 98
7.3 MCMC pour les temps de séjour poissonniens 98

Glossaire

abattis brûlis pratique agraire qui consiste à défricher les champs (souvent forestière) par le feu, puis sont cultivés pendant une brève période pour être ensuite mis en jachère..

Culture annuelle Culture (plante) qui ne vive qu'une saison, qui naît, fructifie et mort dans la l'année (mais, manioc,...).

Culture pérenne culture dont la durée de vie peut s'étaler sur plusieurs années.

EMV Estimateur du maximum de vraisemblance.

Forêt naturelle forêt primaire, forêt vierge qui n'a jamais été exploité.

Forêt secondaire forêt qui a repoussé (plantée ou par régénération).

forêt tropicale forêt qui croît dans les regions tropicales et subtropicales : forêt avec des grands arbres, climat chaud et beaucoup de pluie.

Herbe Végétation principalement herbacée.

INRA Institut nationale de recherche en Agronomie.

INRIA Institut nationale de recherche en informatique et en automatique.

IRD Institut de recherche pour le développement .

Jachère Terre cultivable laissé en repos.

MCMC Méthode de monté carlo par chaîne de Markov.

MODEMIC Modélisation et optimisation des dynamiques des ecosyst emes microbiens.

parcelle portion de terrrain de même culture.

processus stochastique Une famille $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires. Il sert à decrire l'évolution d'une grandeur aléatoire en fonction du temps.

Rizière Parcelle reservée à la culture du riz .

Thèse préparée au sein
de l'Équipe-projet MODEMIC INRIA/INRA
UMR MISTEA
2 Place Pierre Viala
Montpellier 34000
France

Table des matières

Table des figures **viii**

Liste des algorithmes **x**

Table des matières **xv**

1 Introduction **1**

2 Problème de dynamique agraire **5**
2.1 Contexte 5
2.2 Les données 6
2.3 Démarche de modélisation 8

3 Généralités sur les chaînes de Markov **13**
3.1 Concepts de base 13
3.1.1 Définitions et propriétés 13
3.1.2 Propriétés de Markov 16
3.1.3 Classification des états 18
3.1.4 Période d'un état 21
3.2 Chaîne de Markov régulière 22
3.2.1 Définitions et propriétés 22
3.2.2 Loi stationnaire 23
3.3 Chaîne de Markov absorbante 28
3.3.1 Définitions et propriétés 28
3.3.2 Loi quasi-stationnaire 30
3.4 Méthodes MCMC 31
3.4.1 Principe de base 31
3.4.2 L'algorithme de Metropolis–Hastings 31

4 Modèles markoviens **35**

4.1 Description du modèle 35
4.1.1 Les observations 35
4.1.2 Caractérisation du modèle 36
4.2 Inférence des modèles 38
4.2.1 Méthode du maximum de vraisemblance 38
4.2.2 Méthode bayésienne 40
4.3 Résultats et discussions 48
4.3.1 Modèle 1 : chaîne de Markov absorbante déduite de la méthode de vraisemblance 48
4.3.2 Modèle 2 : chaîne de Markov régulière déduite de l'approche bayésienne 51
4.3.3 Discussions 54

5 Evaluation des modèles markoviens 57
5.1 Test d'adéquation des lois de temps de séjour 57
5.1.1 Lois des temps de séjour 57
5.1.2 Test d'adéquation par bootstrap paramétrique 58
5.1.3 Résultats 59
5.2 Tests d'homogénéité 61
5.2.1 Test d'homogénéité en espace selon la date de mise en culture 61
5.2.2 Test d'homogénéité en espace selon l'exposition 62
5.2.3 Test d'homogénéité en temps 64

6 Généralités sur les chaînes semi-markoviennes 67
6.1 Définitions et propriétés 67
6.1.1 Chaîne semi-markovienne 68
6.1.2 Lois des temps de séjour 71
6.1.3 Loi de transition semi-markovienne 71
6.2 Comportement asymptotique 74
6.2.1 Classification des états 74
6.2.2 Loi stationnaire 75
6.3 Exemple d'application 76

7 Modèles semi-markoviens 81
7.1 Motivation 81
7.2 Description du modèle 83
7.2.1 Les observations 83
7.2.2 Hypothèses et présentation du modèle 84
7.3 Inférence des modèles 86
7.3.1 Méthode empirique 87

7.3.2 Méthode du maximum du vraisemblance 90
7.3.3 Méthode bayésienne 93
7.4 Résultats et discussions 99
7.4.1 Modèles absorbants 99
7.4.2 Modèle régulier 104
7.4.3 Discussion 109

Conclusions et perspectives **115**

Bibliographie **119**

Chapitre 1

Introduction

Madagascar est réputée pour sa grande forêt tropicale située le long de la côte est du pays (cf. Figure 1.1). Cette forêt, qui abrite une grande biodiversité essentiellement constituées d'espèces endémiques, est l'une des grandes richesses de l'île.

La population malgache est principalement constituées d'agriculteurs, pourtant le problème d'auto-suffisance alimentaire est encore une grande préoccupation pour le pays. Ce paradoxe s'explique par le fait que les paysans malgaches utilisent toujours des techniques traditionnelles basées sur les systèmes d'abattis brûlis. Ces techniques sont souvent peu productives et ont un impact néfaste sur l'environnement et les ressources naturelles. En particulier, elles ont largement contribué au phénomène de déforestation qui ravage la forêt de l'est du Pays. De nombreux travaux suivent et tentent d'expliquer ce phénomène [20, 5].

Les politiques agricoles doivent donc prendre en compte deux objectifs antagonistes : tendre vers une production suffisante tout en préservant les ressources naturelles et notamment les forêts primaires et la biodiversité associée. Il est donc particulièrement important d'analyser et comprendre les dynamiques agraires notamment celles en bordure du corridor forestier Ranomafana-Andringitra.

Cette thèse vise à proposer des modèles de dynamiques d'usages des sols. Ces modèles sont inférés en s'appuyant sur des mesures de terrains, et vont permettre de répondre aux questions suivantes :

(*i*) *Les dynamiques d'usages des sols se stabilisent-elles ?*

(*ii*) *Quelle est la natures des dynamiques stabilisées ?*

(*iii*) *Enfin, à quelle vitesse ces dynamiques se stabilisent-elles ?*

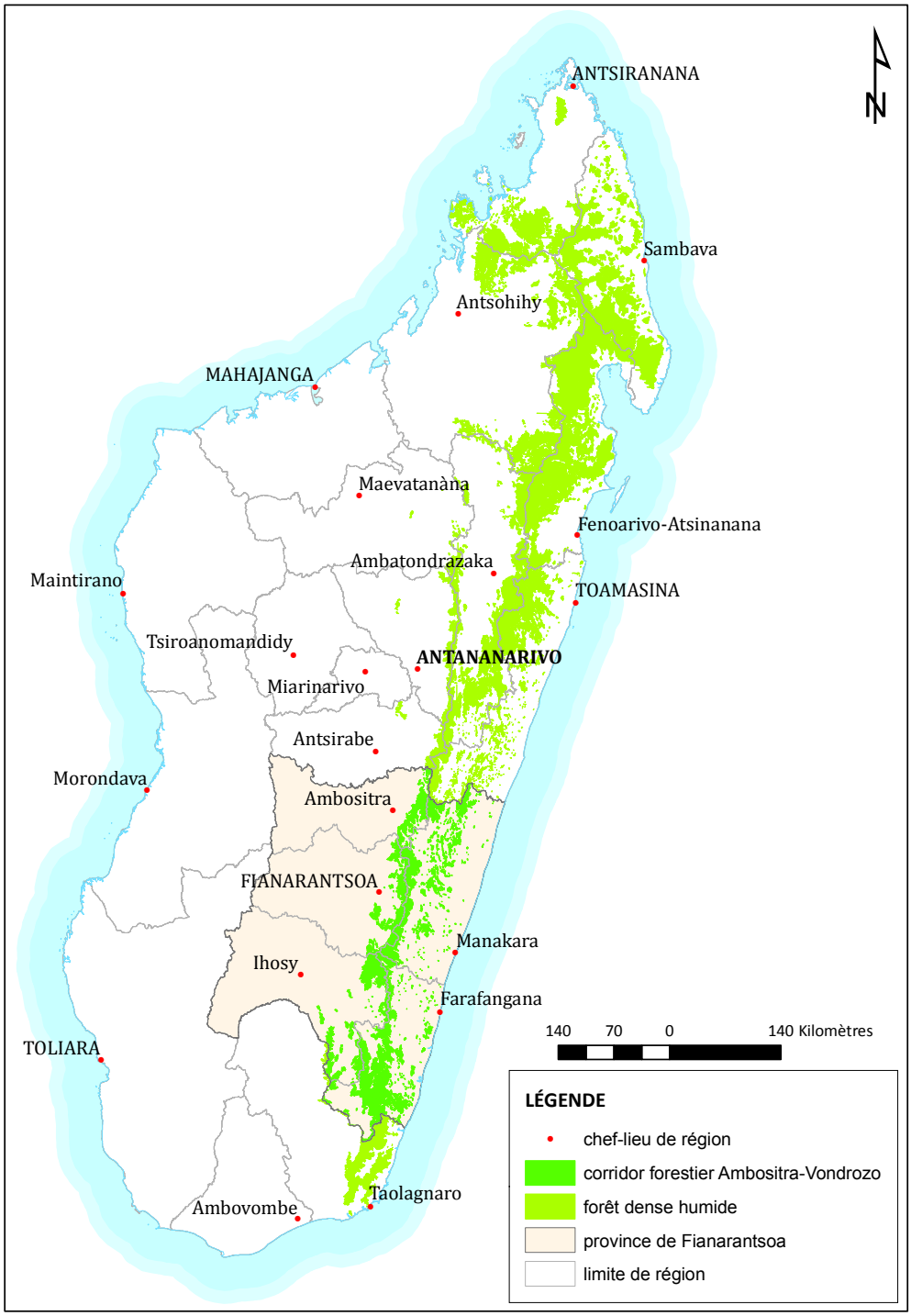

Figure 1.1: *Forêt dense et humide à l'est de Madagascar*

Les études des dynamiques d'usages des sols peuvent se faire à grande échelle (province, région) [1, 2, 12] ou à petite échelle (village, parcelle) [35, 48]. Notre étude se fait à l'échelle des parcelles : à partir des historiques des usages des parcelles dans le village d'Ambendrana (cf. Figure 2.1), nous proposons des modèles d'usages des parcelles qui seront ensuite utilisés comme outils d'inférence des dynamiques existantes. Contrairement à Agarwal *et al.* [2], nous ne tenons compte que des dynamiques temporelles afin de comprendre l'échelle de temps sur la dynamique d'utilisations des parcelles.

Nous travaillons sur les modèles markoviens à temps discret et à espace d'états fini. Ces modèles ont rencontré un vif succès depuis les années 50, notamment en économie agraire ainsi qu'en dynamique écologique ou forestière. L'approche markovienne permet :

- *de rendre compte de phénomènes aléatoires inhérents à l'écologie,*
- *de décrire des dynamiques admettant des équilibres stochastiques,*
- *d'améliorer la robustesse des modèles en encapsulant les erreurs des modèles dans des termes stochastiques.*

Le fait que les chaînes de Markov peuvent converger vers un équilibre, la mesure invariante, tout en conservant un comportement aléatoire a rapidement séduit la communauté des écologues. Les modèles markoviens permettent en effet de rendre compte de la convergence d'un système écologique vers son climax.

Plusieurs applications on été étudiées notamment en écologie [57, 56], en dynamique agraire [8, 53], en dynamique forestière [24, 17] ainsi qu'en séquestration du carbone [25, 39, 44, 46, 27]. Ces modèles peuvent aussi prendre en compte les dimensions socio-écologiques [45]. Ils sont actuellement aussi utilisés dans le contexte de l'exploitation des données satellitaires, notamment en conjonction avec les outils des systèmes d'informations géographiques [60]. Outre la modélisation, l'ingénierie markovienne se prête également à l'inférence statistique, au contrôle et à l'aide à la décision.

Il existe plusieurs extensions des modèles markoviens : les *modèles de Markov cachés* qui permettent de prendre en compte les dynamiques cachées d'un processus [54], les *modèles semi-markoviens* qui autorisent une plus grande souplesse concernant les temps de séjour dans un état donné et les champs spatio-temporels markoviens qui permettent de prendre en compte les dynamiques spatio-temporelles [61]. Il existe enfin de nombreuses généralisations au temps continu.

Partant d'un premier jeu de données, nous proposons différents modèles de Markov classiques. Nous estimons la matrice de transition de la chaîne de Markov d'abord à l'aide de la méthode de maximum de vraisemblance [3] puis à l'aide de l'approche bayésienne en utilisant la loi a priori non informative de Jeffreys. Le calcul de la loi a posteriori associée à la loi de Jeffreys n'est pas explicite,

nous faisons donc appel à une méthode de Monte Carlo par chaîne de Markov (MCMC).

A l'aide d'un deuxième jeu des données, nous inférons des modèles de types semi-markoviens. L'inférence se faît par une méthode empirique [7, 6], une méthode de maximum de vraisemblance et une méthode bayésienne. Pour la méthode bayésienne, nous utilisons à nouveau la loi de Jeffreys comme loi a priori et le calcul de la loi a posteriori se fait par une méthode MCMC. En général, les différentes méthodes d'inférences conduisent à des modèles différents, qui ont des propriétés asymptotiques différentes.

Au Chapitre 2, nous introduisons le problème de dynamique agraire et la démarche de modélisation que nous adoptons. Au Chapitre 3, nous présentons des généralités sur les chaînes de Markov. Au Chapitre 4, nous utilisons les chaînes de Markov afin de modéliser la dynamique d'utilisation des parcelles. Des tests statistiques sont proposés dans le Chapitre 5 pour évaluer les modèles markoviens étudiés précédemment. L'objectif de ces tests est de mesurer l'adéquation des modèles au jeu de données. Le Chapitre 6 est une présentation générale des chaînes semi-markoviennes et le Chapitre 7, leurs applications au problème de dynamique agraire.

Chapitre 2

Le problème de dynamique agraire sur le bord du corridor forestier Ranomafana–Andringitra

2.1 Contexte

Le corridor Ranomafana–Andringitra est le nom donnée au couloir forestier de 200 km de longueur qui relie les deux parcs nationaux de Ranomafana et Andringitra, voir la figure 2.1. Il joue un rôle primordial dans le maintien de la forêt de l'est du pays. Il assure en particulier les échanges biologiques entre les deux aires protégées.

Face aux pratiques agricoles paysannes comme le tavy (aménagement par défriche-brûlis pour cultures vivrières), aux feux et coupes servant aux activités de production (chasse, élevage, exploitation des bois forestiers, etc.) le corridor Ranomafana–Andringitra est en danger [41, 48]. Actuellement il fait l'objet de différents projets de conservation menés par différents intervenants dont l'état est le principal acteur.

Un compromis doit être établi entre d'une part les paysans qui ont pour objectif de maximiser la production et d'autre part les dirigeants (décideurs) qui doivent mettre en place des politiques d'exploitation pour accompagner les paysans dans l'exploitation de leurs ressources tout en respectant certaines contraintes (nationales ou internationales) sur le respect de l'environnement et les ressources naturelles.

Le but à long terme, qui dépasse ce travail de thèse, est de pouvoir développer des stratégies de gestion qui permettent de concilier les intérêts des agriculteurs

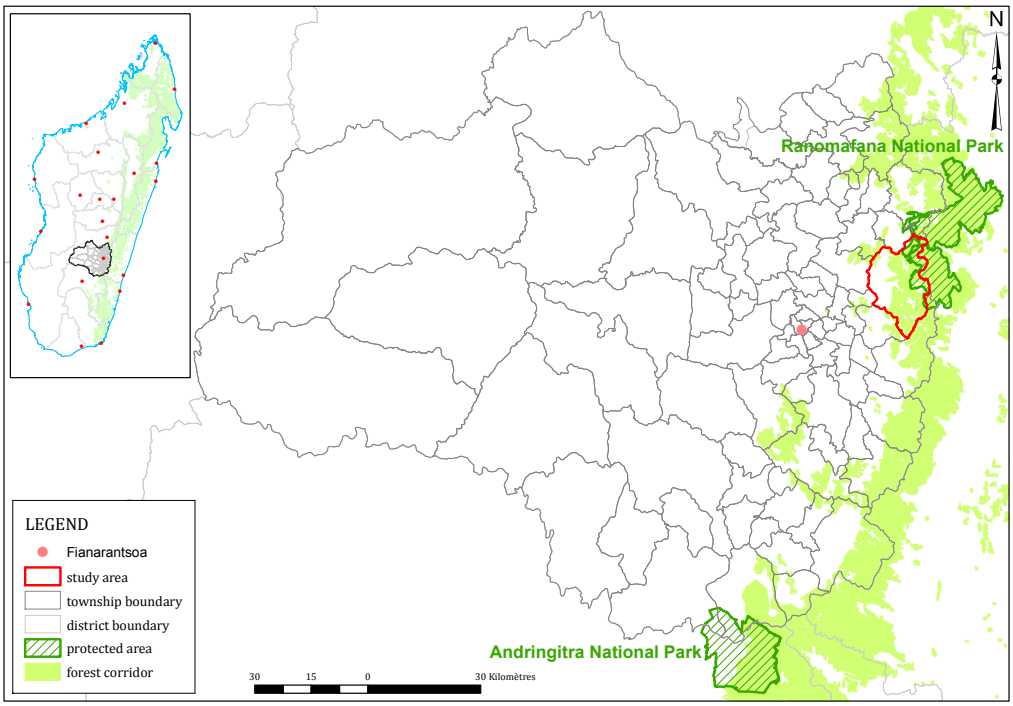

Figure 2.1: *Zone d'étude : Village d'Ambendrana, commune rurale Androy qui se situe au bord du corridor forestier Ranofamana-Andringitra.*

avec la protection de la forêt malgache et des enjeux associés en termes de protection de la biodiversité. La présente thèse constitue l'étape préalable dans ce processus : comprendre les pratiques agricoles existantes notamment ces dynamiques. Deux questions sont alors posées :

(*i*) les successions culturales convergent-elle vers une proportion d'usage des parcelles stable ?

(*ii*) au bout de combien de temps peut-on atteindre cette stabilité ?

À partir des données d'usage des parcelles, nous développons des modèles qui seront ensuite utilisés comme outils d'inférence statistique de la dynamique sous-jacente. Ces analyses statistiques permettent de comprendre la nature et les propriétés de cette dynamique.

2.2 Les données

Nous nous appuyons sur le travail de Carrière *et al.* [11]. Le village d'Ambendrana se caractérise par un paysage hétérogène composé de parcelles dédiés à des

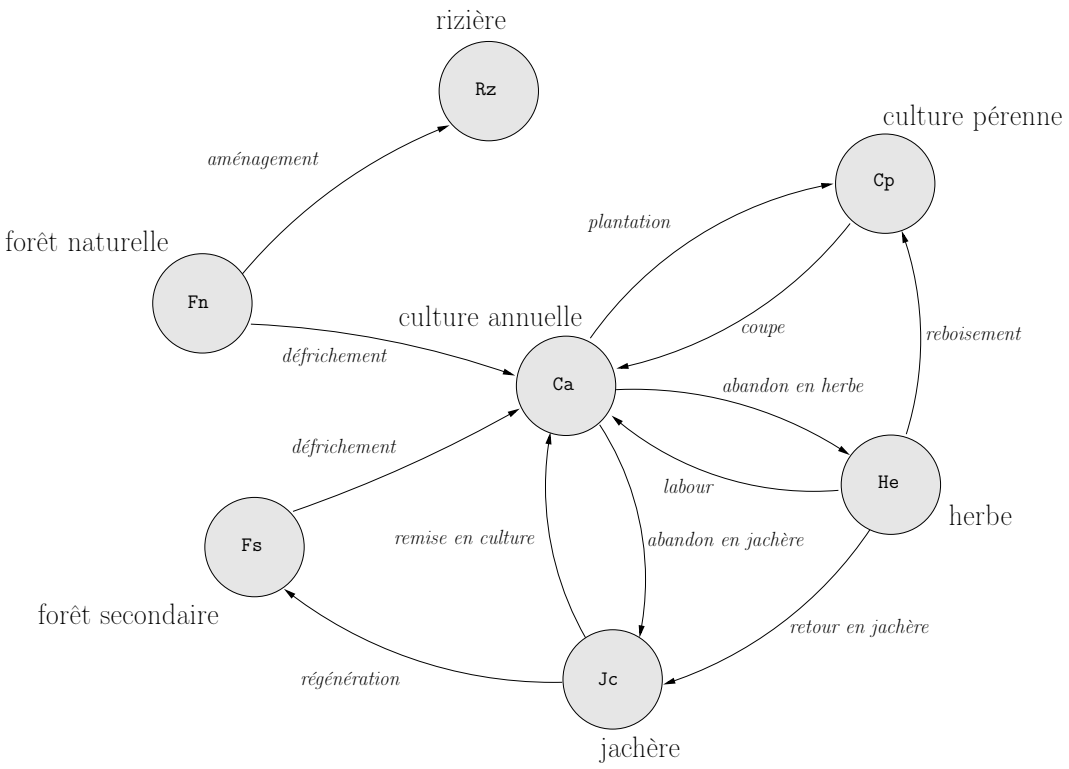

Figure 2.2: *Dynamique d' occupation des parcelles au bord du corridor forestier Ranomafana-Andringitra.*

usages des sols différents que nous appellerons les états de la parcelle :

- *Forêt naturelle* (`Fn`), *ou forêt primaire,*
- *Culture annuelle* (`Ca`) *: cultures vivrières (manioc, haricot etc.),*
- *Jachère* (`Jc`),
- *Rizière* (`Rz`),
- *Culture pérenne* (`Cp`) *: canne à sucre, boisement ou plantation d'arbres (pins, eucalyptus, etc.),*
- *Herbe* (`He`) *: prairie ou savane herbeuse,*
- *Forêt secondaire* (`Fs`) *: formation secondaire de plus de 25 ans (jachère arborée de plus de 25 ans).*

L'utilisation d'une parcelle évolue au cours du temps sous l'effet des pratiques paysannes mais aussi selon des phénomènes écologiques naturels. A l'aide de collectes de données et d'enquête auprès des paysans, les experts de l'IRD ont établi des règles de changement d'usage des parcelles, voir la figure2.2.

Initialement, toutes les parcelles sont à l'état forêt naturelle. Les habitants défrichent et aménagent la forêt pour en faire soit une rizière lorsque la parcelle est dans les bas-fonds, soit des cultures vivrières (manioc, maïs, haricot). Après

quelques années de culture, les parcelles sont laissées en jachère. L'état culture pérenne désigne les plantations pérennes comme la canne à sucre, le café, la banane ou des arbres (pins, eucalyptus, mimosa, bambous).

L'équipe animée par Dominique Hervé a reconstitué l'historique cultural de plusieurs parcelles situées dans le village d'Ambendrana. Dans cette thèse, nous travaillons sur deux jeux de données construits à partir de ces historiques.

Le premier jeu de données est composé de 42 parcelles suivies pendant 22 ans, de 1985 à 2006. L'historique d'occupation des sols pendant ces 22 années est représenté dans le tableau 2.1. Ces observations font apparaître seulement 4 états :

- Forêt naturelle (`Fn`),
- Culture annuelle (`Ca`),
- Jachère (`Jc`),
- Culture pérenne (`Cp`).

Le deuxième jeu des données est plus étendu. Il concerne les usages des 131 parcelles de 1956 à 2006, c'est-à-dire pendant 50 ans, voir la figure 2.3. Le nombre d'états a naturellement augmenté puisque le nombre de parcelles a augmenté. En plus des quatre états précédents, on observe l'apparition de trois autres états :

- Herbe (`He`),
- Rizière (`Rz`),
- Forêt secondaire (`Fs`).

2.3 Démarche de modélisation

Cette thèse n'a pas pour but de développer des stratégies d'exploitation des parcelles agricoles encore moins d'établir des nouvelles politiques agricoles. Elle vise à mettre en place les bases d'une démarche de modélisation pour aider les acteurs et les décisionnaires de ces politiques. Les modèles proposés constituent des outils d'analyse et de compréhension qui permettent d'exploiter les données de terrain à l'aide de traitements statistiques.

La toute première étape dans cette exploitation est de construire la matrice de transition empirique en comptant les fréquences relatives des différentes transitions observées. Ce simple comptage revient en fait à supposer que la dynamique sous-jacente est markovienne homogène en temps et indépendante d'une parcelle à l'autre. Cette hypothèse est évidemment irréaliste : il y a toujours des effets mémoire contredisant l'hypothèse de Markov et les dynamiques des parcelles ne peuvent pas être indépendantes les unes des autres. Cette hypothèse préliminaire permet néanmoins d'analyser les données, d'en extraire une certaine information notamment en termes d'échelles de temps. Ces échelles de temps concernent à la fois les dynamiques anthropiques (choix des agriculteurs, évolution des politiques

parcelles

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42
1985	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn
1986	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Ca
1987	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Ca
1988	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Ca
1989	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Jc	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Cp
1990	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Ca	Ca	Ca	Ca	Ca	Ca	Jc	Jc	Ca	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Cp
1991	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Ca	Ca	Jc	Ca	Ca	Jc	Jc	Ca	Ca	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Cp
1992	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Ca	Ca	Jc	Ca	Ca	Jc	Ca	Ca	Jc	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Cp
1993	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Ca	Ca	Jc	Ca	Jc	Ca	Jc	Jc	Jc	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Cp
1994	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Jc	Jc	Jc	Jc	Jc	Ca	Jc	Jc	Jc	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Cp
1995	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Jc	Jc	Jc	Jc	Ca	Ca	Jc	Ca	Ca	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Cp
1996	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Ca	Ca	Jc	Jc	Jc	Jc	Ca	Jc	Jc	Ca	Ca	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Cp
1997	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Ca	Ca	Jc	Jc	Jc	Ca	Ca	Jc	Ca	Ca	Jc	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Cp
1998	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Ca	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Ca	Ca	Jc	Jc	Jc	Ca	Ca	Ca	Ca	Jc	Jc	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Cp
1999	Fn	Fn	Fn	Fn	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Ca	Ca	Jc	Jc	Jc	Ca	Ca	Jc	Jc	Ca	Jc	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Fn	Cp
2000	Ca	Fn	Fn	Fn	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Jc	Jc	Ca	Jc	Jc	Jc	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Cp
2001	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Jc	Ca	Ca	Ca	Ca	Ca	Ca	Jc	Jc	Jc	Ca	Jc	Ca	Jc	Jc	Jc	Ca	Jc	Jc	Ca	Jc	Jc	Jc	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Cp
2002	Jc	Ca	Ca	Ca	Ca	Ca	Ca	Jc	Ca	Jc	Ca	Jc	Ca	Ca	Jc	Jc	Jc	Ca	Jc	Ca	Ca	Ca	Jc	Jc	Ca	Ca	Jc	Ca	Ca	Jc	Jc	Jc	Ca	Ca	Ca	Ca	Ca	Jc	Jc	Ca	Ca	Cp
2003	Jc	Ca	Ca	Ca	Ca	Ca	Cp	Ca	Ca	Ca	Ca	Jc	Ca	Ca	Jc	Jc	Jc	Ca	Jc	Jc	Ca	Ca	Ca	Jc	Ca	Ca	Jc	Ca	Ca	Jc	Jc	Jc	Ca	Ca	Ca	Ca	Ca	Jc	Ca	Ca	Ca	Cp
2004	Jc	Ca	Ca	Ca	Ca	Ca	Cp	Ca	Ca	Ca	Ca	Jc	Ca	Ca	Jc	Jc	Jc	Jc	Jc	Jc	Ca	Ca	Ca	Jc	Jc	Jc	Ca	Jc	Ca	Jc	Ca	Jc	Ca	Jc	Ca	Ca	Ca	Ca	Ca	Ca	Ca	Cp
2005	Ca	Cp	Ca	Ca	Ca	Jc	Cp	Ca	Jc	Jc	Jc	Jc	Jc	Jc	Ca	Ca	Ca	Jc	Jc	Jc	Ca	Jc	Ca	Ca	Jc	Ca	Ca	Jc	Ca	Ca	Ca	Ca	Jc	Ca	Jc	Ca	Ca	Jc	Ca	Ca	Ca	Cp
2006	Ca	Cp	Ca	Ca	Ca	Ca	Cp	Ca	Jc	Ca	Ca	Jc	Jc	Jc	Ca	Ca	Jc	Jc	Jc	Jc	Jc	Ca	Jc	Jc	Ca	Ca	Ca	Jc	Ca	Ca	Jc	Ca	Ca	Ca	Ca	Jc	Ca	Ca	Jc	Ca	Ca	Cp

Tableau 2.1: *Jeu de données n° 1 : États annuels des 42 parcelles qui se trouvent sur la lisière du corridor Ranomafana–Andringitra, cf. Figure 2.1. Chaque parcelle peut présenter 4 états possibles : “ forêt naturelle”* (`Fn`), *“culture annuelle”* (`Ca`), *“jachère”* (`Jc`), *et “culture pérenne”* (`Cp`)

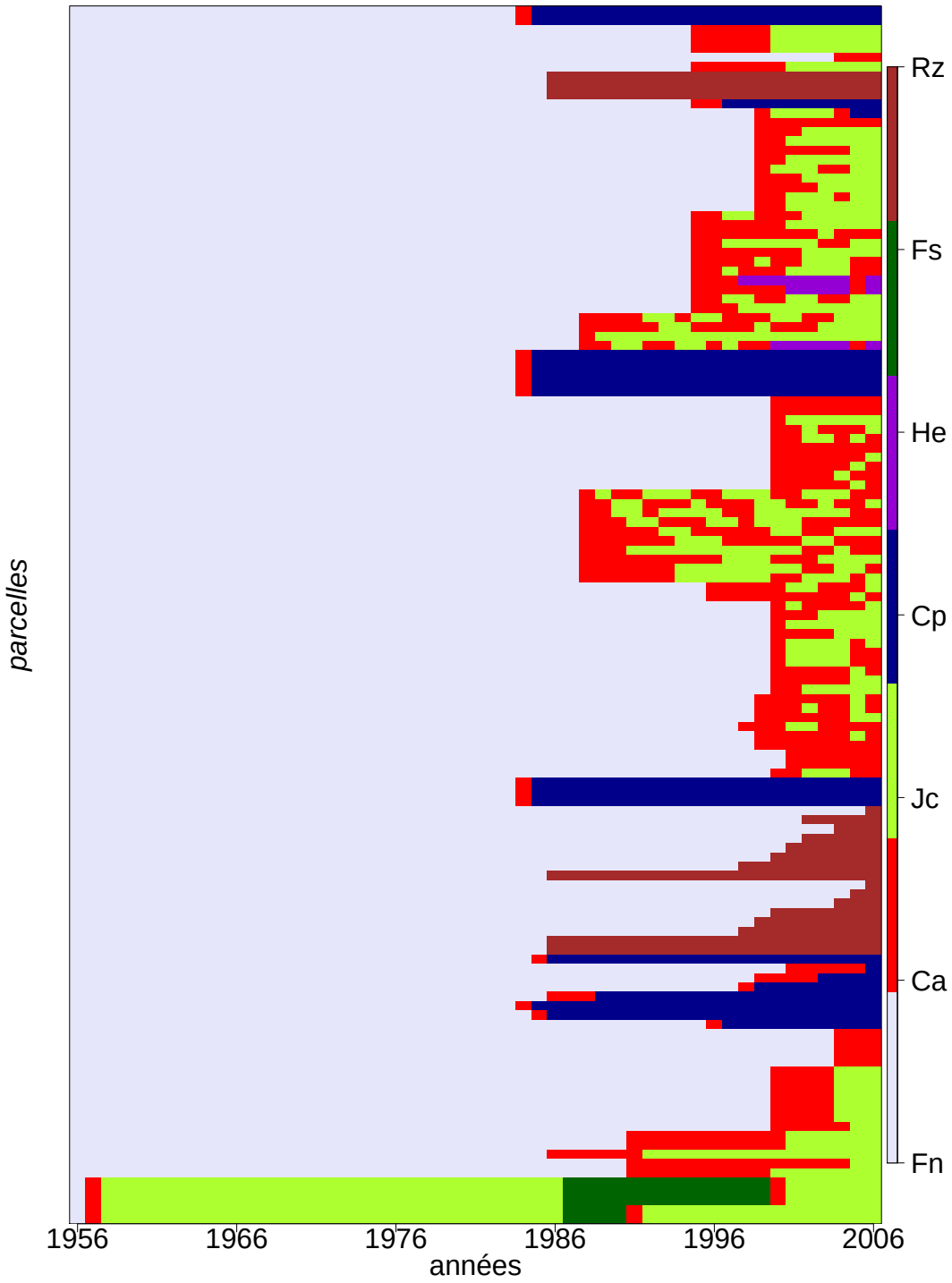

Figure 2.3: *Jeu de données n° 2 : Évolution annuelle de l'occupation des 131 parcelles sur la lisière ouest du corridor Ranomafana–Andringitra. On observe 7 états : Forêt naturelle* (`Fn`)*, Culture annuelle* (`Ca`)*, Jachère* (`Jc`)*, Culture pérenne* (`Cp`)*, Herbe* (`He`)*, Rizière* (`Rz`) *et Forêt secondaire* (`Fs`)*.*

agricoles) et les dynamiques écologiques (recolonisation par la forêt).

Sur le premier jeu de données, voir le tableau 2.1, nous travaillons dans le cadre des modèles markoviens. Dans un premier temps, on fait appel à la statistique empirique, correspondant à l'approche par maximum de vraisemblance. Cette approche ne permet pas d'intégrer la connaissance a priori des spécialistes en charge des relevés et des enquêtes de terrain, pour cela nous faisons appel à la statistique bayésienne. Les résultats obtenus sont relativement différents.

Les modèles de Markov ne sont pas adaptés au second jeu de données, cf. figure 2.3. Nous faisons appel à des modèles semi-markoviens pour inférer la dynamique d'usage des parcelles. Pour ces modèles, l'approche empirique ne correspond pas exactement à l'approche par maximum de vraisemblance. En plus de ces deux approches, nous faisons à nouveau appel à l'approche bayésienne afin d'intégrer la connaissance des spécialistes qui n'est pas présente dans le jeu des données.

Chapitre 3

Généralités sur les chaînes de Markov

Dans ce chapitre, nous reprenons quelques résultats classiques sur les chaînes de Markov. On peut se reporter à [37] pour plus de détails.

3.1 Concepts de base

3.1.1 Définitions et propriétés

Considérons un processus stochastique en temps discret $(X_n)_{n\in\mathbb{N}}$ défini sur un espace de probabilités $(\Omega, \mathcal{F}, \mathbb{P})$, à valeurs dans un espace d'état fini E. La loi d'un tel processus est la donnée de :

$$\mathbb{P}(X_0 = i_0, \ldots, X_n = i_n),$$

pour tout $n \geq 0$ et $i_0, \ldots, i_n \in E$.

Définition 3.1.1 (Chaîne de Markov)**.** *Un processus stochastique $(X_n)_{n\geq 0}$ est appelé chaîne de Markov si pour tout $n \in \mathbb{N}$, la loi conditionnelle de X_{n+1} sachant $X_0, X_1, \ldots, X_n$ est égale à la loi conditionnelle sachant X_n, c'est-à-dire, pour tout $i_0, i_1, \ldots, i_n \in E$:*

$$\begin{aligned}\mathbb{P}(X_{n+1} = i_{n+1} | X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) \\ = \mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n).\end{aligned} \tag{3.1}$$

Si de plus, la probabilité

$$\mathbb{P}(X_{n+1} = j | X_n = i)$$

ne dépend pas de n, c'est-à-dire

$$\forall n \in \mathbb{N},\ \mathbb{P}(X_{n+1} = j | X_n = i) = \mathbb{P}(X_1 = j | X_0 = i)\,,$$

alors la chaîne de Markov est dite homogène.

Dans la suite, toutes les chaînes de Markov seront supposées homogènes. Nous adoptons la notation :

$$\{p : q\} \stackrel{\text{déf}}{=} \{p, p+1, \ldots, q\}.$$

Ainsi $X_{n:m} = i_{n:m}$ signifie $X_n = i_n$, $X_{n+1} = i_{n+1}, \ldots, X_m = i_m$.

Définition 3.1.2 (Matrice de transition)**.** *Étant donnée une chaîne de Markov homogène $(X_n)_{n\in\mathbb{N}}$ à valeurs dans E, la matrice $Q = [Q_{ij}]$ définie par*

$$Q_{ij} \stackrel{\text{déf}}{=} \mathbb{P}(X_{n+1} = j | X_n = i) \tag{3.2}$$

est appelée matrice de transition.

Une matrice de transition Q vérifie $Q_{ij} \geq 0$ pour tout i, j, et $\sum_{j\in E} Q_{ij} = 1$ pour tout i. Chaque ligne de cette matrice est donc une loi de probabilité sur l'espace d'état E.

Définition 3.1.3 (Loi initiale)**.** *La loi μ définie par :*

$$\mu_i \stackrel{\text{déf}}{=} \mathbb{P}(X_0 = i)$$

est appelée loi initiale.

Si $(X_n)_{n\in\mathbb{N}}$ admet Q comme matrice de transition et μ comme loi initiale, alors on dira qu'elle est une chaîne de Markov (μ, Q).

Proposition 3.1.4. *La loi d'une chaîne de Markov homogène est entièrement déterminée par la donnée de sa matrice de transition et de sa loi initiale. En effet, pour tout $n \in \mathbb{N}$ et pour tous états $i_0, i_1, \ldots, i_n$ de E :*

$$\mathbb{P}(X_{0:n} = i_{0:n}) = \mu_{i_0}\, Q_{i_0 i_1} \cdots Q_{i_{n-1} i_n}. \tag{3.3}$$

Démonstration. Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov (μ, Q) homogène. Alors :

$$\begin{aligned}\mathbb{P}(X_{0:n} = i_{0:n}) =& \mathbb{P}(X_0 = i_0)\, \mathbb{P}(X_1 = i_1 | X_0 = i_0) \\ & \cdots \mathbb{P}(X_n = i_n | X_{0:n-1} = i_{0:n-1}) \\ &= \mu_{i_0}\, Q_{i_0 i_1} \cdots Q_{i_{n-1} i_n}.\end{aligned}$$

□

Posons :

$$Q^{(n)} \stackrel{\text{déf}}{=} \underbrace{Q \times \cdots \times Q}_{n \text{ fois}},$$

avec $Q^{(0)} = I$ (matrice identité).

Théorème 3.1.5. *Si $(X_n)_{n\geq 0}$ est une chaîne de Markov homogène de matrice de transition Q et de loi initiale μ, alors :*

(i) $\mathbb{P}(X_n = j|X_0 = i) = \mathbb{P}(X_{n+p} = j|X_p = i) = Q_{ij}^{(n)}$,

(ii) $\mathbb{P}(X_n = j) = [\mu\, Q^{(n)}]_j$,

(iii) $\mathbb{E}\,(f(X_n)|X_0 = i) = [Q^{(n)} f]_i$

pour tout $n \geq 0$ et $i, j \in E$.

Démonstration. Pour (i) nous pouvons écrire :

$$\begin{aligned}\mathbb{P}(X_n = j|X_0 = i) &= \sum_{i_{1:n-1}} \mathbb{P}(X_n = j, X_{1:n-1} = i_{1:n-1}|X_0 = i),\\ &= \sum_{i_{1:n-1}} \frac{\mu_i\, Q_{ii_1} \cdots Q_{i_{n-1}j}}{\mu_i} = [Q^{(n)}]_{ij}\,.\end{aligned}$$

Pour (ii)

$$\begin{aligned}\mathbb{P}(X_n = j) &= \sum_{i\in E} \mathbb{P}(X_n = j, X_0 = i)\\ &= \sum_{i\in E} \mathbb{P}(X_n = j|X_0 = i)\,\mu_i = [\mu\, Q^{(n)}]_j\,.\end{aligned}$$

Enfin (iii) s'obtient par

$$\begin{aligned}\mathbb{E}[f(X_n)|X_0 = i] &= \sum_{j\in E} f(j)\mathbb{P}(X_n = j|X_0 = i),\\ &= (Q^{(n)} f)_i\,.\end{aligned}$$

□

Théorème 3.1.6 (Equation de Chapman-Kolmogorov)**.** *Soit $(X_n)_{n\geq 0}$ une chaîne de Markov homogène. Pour tout couple (i, j) d'états de E et pour tout couple (n, m) d'entiers naturels :*

$$\mathbb{P}(X_{n+m} = j|X_0 = i) = \sum_{k\in E} \mathbb{P}(X_n = k|X_0 = i)\,\mathbb{P}(X_m = j|X_0 = k). \tag{3.4}$$

Démonstration.

$$\begin{aligned}\mathbb{P}(X_{n+m} = j|X_0 = i) &= \sum_{k\in E} \mathbb{P}(X_{n+m} = j|X_n = k, X_0 = i)\,\mathbb{P}(X_n = k|X_0 = i)\\ &= \sum_{k\in E} \mathbb{P}(X_{n+m} = j|X_n = k)\,\mathbb{P}(X_n = k|X_0 = i)\\ &= \sum_{k\in E} \mathbb{P}(X_m = j|X_0 = k)\,\mathbb{P}(X_n = k|X_0 = i)\\ &\text{(car } (X_n)_{n\geq 0} \text{ est chaîne de Markov homogène)}\\ &= \sum_{k\in E} \mathbb{P}(X_n = k|X_0 = i)\,\mathbb{P}(X_m = j|X_0 = k)\end{aligned}$$

□

En notation matricielle, (3.4) peut s'écrire :

$$Q^{(n+m)} = Q^{(n)}\, Q^{(m)},$$

où Q est la matrice de transition de $(X_n)_{n\geq 0}$.

3.1.2 Propriétés de Markov

Pour une chaîne de Markov $(X_n)_{n\geq 0}$ de matrice de transition Q, on note $\mathbb{P}_\mu$ la probabilité définie sur $(\Omega, \mathcal{F})$ telle que sous $\mathbb{P}_\mu$, $(X_n)_{n\in\mathbb{N}}$ admette μ comme loi initiale, c'est-à-dire :

$$\mathbb{P}_\mu(X_0 = j) \overset{\text{déf}}{=} \mu_j$$

pour tout $j \in E$.

Pour tout $i, j \in E$, on note $\mathbb{P}_i = \mathbb{P}_{\delta_i}$ où δ_i est la loi de probabilité définie par :

$$\delta_i(j) = \delta_{ij},$$

c'est-à-dire :

$$\mathbb{P}_i(X_0 = j) = \delta_{ij}.$$

Définition 3.1.7 (Temps d'arrêt)**.** *Soit :*

$$\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$$

la filtration naturelle associée au processus $(X_n)_{n\geq 0}$. Une variable aléatoire T à valeurs dans $\mathbb{N} \cup \{\infty\}$ est appelée temps d'arrêt si

$$\{T = n\} \in \mathcal{F}_n$$

pour tout n. L'observation de la chaîne jusqu'à l'instant n, permet de décider si T vaut n ou non.

Théorème 3.1.8 (Propriété de Markov faible)**.** *On considère une chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ de matrice de transition Q et de loi initiale μ. Pour tout $i \in E$ et $n \in \mathbb{N}$, conditionnellement à $X_n = i$, le processus $(\tilde{X}_p)_{p\in\mathbb{N}}$ défini par*

$$\tilde{X}_p \overset{\text{déf}}{=} X_{n+p}$$

est une chaîne de de Markov indépendante de $\mathcal{F}_n$. La chaîne $(\tilde{X}_p)_{p\in\mathbb{N}}$ a la même matrice de transition que $(X_n)_{n\in\mathbb{N}}$ et sa loi initiale est δ_i, c'est-à-dire :

$$\mathbb{P}(A \cap \{\tilde{X}_{0:m} = i_{n:n+m}\} | X_n = i) = \mathbb{P}_i(A | X_n = i)\, \mathbb{P}(\tilde{X}_{0:m} = i_{n:n+m} | X_n = i)$$

et

$$\mathbb{P}(\tilde{X}_{0:m} = i_{n:n+m} | X_n = i) = \delta_{ii_n}\, Q_{i_n i_{n+1}} \cdots Q_{i_{n+m-1} i_{n+m}}$$

pour tout $A \in \mathcal{F}_n$, $m > 0$ et $i_1, \ldots, i_{n+m} \in E$.

Démonstration. Il suffit de se limiter aux événements $A \in \mathcal{F}_n$ de la forme $A = \{X_{0:n} = i_{0:n}\}$ pour tout $i_0, \ldots, i_n \in E$. On a

$$\begin{aligned}
\mathbb{P}(A \cap \{\tilde{X}_{0:m} = i_{n:n+m}\}|X_n = i) &= \\
&= \mathbb{P}(\{X_{0:n} = i_{0:n}\} \cap \{X_{n:n+m} = i_{n:n+m}\}\,|X_n = i) \\
&= \mathbb{P}(X_{0:n+m} = i_{0:n+m}|X_n = i) \\
&= \frac{\mathbb{P}(\{X_{0:n+m} = i_{0:n+m}\} \cap \{X_n = i\})}{\mathbb{P}(X_n = i)} \\
&= \frac{\mathbb{P}(X_{0:n+m} = i_{0:n+m})}{\mathbb{P}(X_n = i)}\,\delta_{ii_n} \\
&= \frac{\mu_{i_0}\,Q_{i_0 i_1}\cdots Q_{i_{n-1}i_n}\,Q_{i_n i_{n+1}}\cdots Q_{i_{n+m-1}i_{n+m}}}{\mathbb{P}(X_n = i)}\,\delta_{ii_n} \\
&= \frac{\mu_{i_0}\,Q_{i_0 i_1}\cdots Q_{i_{n-1}i_n}}{\mathbb{P}(X_n = i)}\,\delta_{ii_n} \times Q_{i_n i_{n+1}}\cdots Q_{i_{n+m-1}i_{n+m}} \\
&= \frac{\mathbb{P}(X_{0:n} = i_{0:n})}{\mathbb{P}(X_n = i)}\,\delta_{ii_n} \times \mathbb{P}(X_{n:n+m} = i_{n:n+m}|X_n = i) \\
&= \mathbb{P}(X_{0:n} = i_{0:n}|X_n = i) \times \mathbb{P}(\tilde{X}_{0:m} = i_{n:n+m}|X_n = i) \\
&= \mathbb{P}(A|X_n = i) \times \mathbb{P}(\tilde{X}_{0:m} = i_{n:n+m}|X_n = i)\,.
\end{aligned}$$

On a donc démontré les deux identités. □

Théorème 3.1.9 (Propriété de Markov forte)**.** *Considérons une chaîne de Markov (μ, Q), $(X_n)_{n\in\mathbb{N}}$, et soit T un temps d'arrêt. Pour tout $i \in E$, conditionnellement à $\{T < \infty\} \cap \{X_T = i\}$, le processus $(\tilde{X}_n)_{n\in\mathbb{N}}$ défini par :*

$$\tilde{X}_n \overset{\text{déf}}{=} X_{T+n}$$

est une chaîne de Markov (δ_i, Q), indépendante de $\mathcal{F}_T$ pour tout $n \in \mathbb{N}$. Autrement dit :

$$\begin{aligned}
\mathbb{P}(A \cap \{\tilde{X}_{0:n} = \tilde{\imath}_{0:n}\}|X_T = i,\, T < \infty) \\
= \mathbb{P}(A|X_T = i, T < \infty)\,\mathbb{P}(\tilde{X}_{0:n} = \tilde{\imath}_{0:n}|X_T = i, T < \infty)
\end{aligned}$$

et

$$\mathbb{P}(\tilde{X}_{0:n} = \tilde{\imath}_{0:n}|X_T = i, T < \infty) = \delta_{i\,\tilde{\imath}_0}\,Q_{\tilde{\imath}_0\tilde{\imath}_1}\cdots Q_{\tilde{\imath}_{n-1}\tilde{\imath}_n}$$

pour tout $A \in \mathcal{F}_T$ et pour $n > 0$.

Démonstration. Soit T un temps d'arrêt $A \in \mathcal{F}_T$, alors $\forall m \in \mathbb{N},\ A \cap \{T = m\} \in \mathcal{F}_m$. D'après la propriété de Markov à l'instant m on a :

$$\begin{aligned}
\mathbb{P}(A \cap \{T = m\} \cap \{\tilde{X}_{0:n} = \tilde{\imath}_{0:n}\}|X_T = i,\, T < \infty) \\
= \mathbb{P}(A \cap \{T = m\}|X_T = i, T < \infty)\,\mathbb{P}(\tilde{X}_{0:n} = \tilde{\imath}_{0:n}|X_T = i, T < \infty)
\end{aligned}$$

et nous démontrons le résultat en faisant la somme sur $m \in \mathbb{N}$. □

3.1.3 Classification des états

Définition 3.1.10. *Soit i et j deux états d'une chaîne de Markov $(X_n)_{n\geq 0}$ de matrice de transition Q. L'état j est dit accessible à partir de i noté $i \to j$, si et seulement si :*

$$\exists\, n \in \mathbb{N}, \mathbb{P}(X_n = j | X_0 = i) = Q_{ij}^{(n)} > 0\,.$$

On dit que les états i et j communiquent, noté $i \leftrightarrow j$ si et seulement si j est accessible à partir de i et i est accessible à partir de j.

La relation $i \leftrightarrow j$ est une relation d'équivalence sur E. L'espace d'état E peut être donc partitionné en classes d'équivalence pour cette relation.

Définition 3.1.11 (Chaîne irréductible)**.** *Lorsque l'espace d'état E est réduit à une seule classe d'équivalence, on dit que la chaîne est irréductible. Dans le cas E fini, cela est équivalent au fait que tous les états communiquent.*

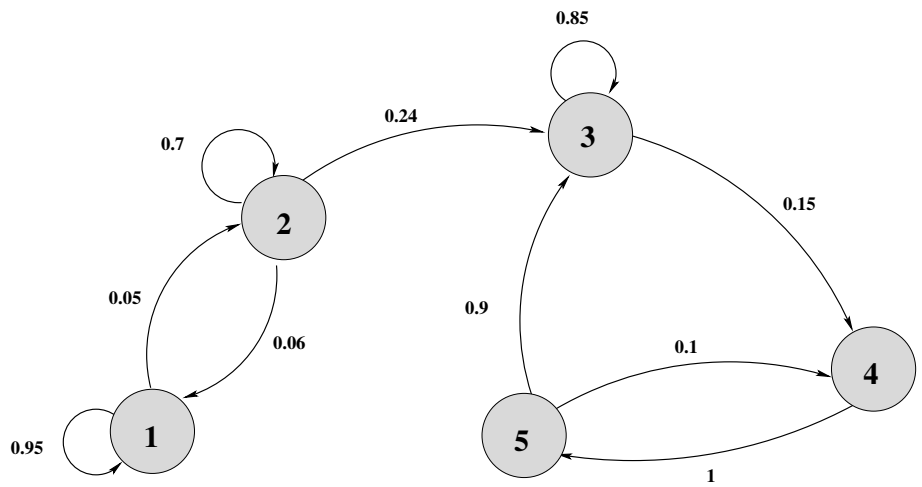

Figure 3.1: *Graphe d'une chaîne de Markov non irréductible, possédant deux classes :* $\{1, 2\}$ *et* $\{3, 4, 5\}$.

Définition 3.1.12 (État récurrent et état transient)**.** *Un état i est dit récurrent si et seulement si, partant de i, la chaîne revient $\mathbb{P}_i$-presque sûrement à l'état i, c'est-à-dire*

$$\mathbb{P}_i(X_n = i \text{ pour une infinité de } n) = 1\,.$$

Si de plus, l'espérance du temps de retour en i est finie, on dit qu'il est récurrent positif. Un état non récurrent est dit transient.

Définition 3.1.13. *On dit qu'une chaîne de Markov ou sa matrice de transition est récurrente (respectivement transiente) si tous les états sont récurrents (respectivement transients).*

Remarque 3.1.14. *Pour $i \in E$, soit T_i le premier temps de retour à l'état i :*

(*i*) *si i est récurrent, alors $\mathbb{P}_i(T_i < \infty) = 1$,*

(*ii*) *si i est transient, alors $\mathbb{P}_i(T_i < \infty) < 1$.*

Définition 3.1.15. *Pour $i \in E, r \in \mathbb{N}$ soit $T_i^{(r)}$ le temps du $r^{ème}$ retour à l'état i, c'est-à-dire $T_i^{(0)} = 0$, $T_i^{(1)} = T_i$ et*

$$T_i^{(r+1)} = \inf\{n \geq T_i^{(r)} + 1, X_n = i\}$$

pour tout $r \in \mathbb{N}$. La longueur de la $r^{ème}$ excursion à l'état i est :

$$S_i^{(r)} = \begin{cases} T_i^{(r)} - T_i^{(r-1)} & si \quad T_i^{(r-1)} < \infty, \\ 0 & sinon. \end{cases}$$

Lemme 3.1.16. *Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov de matrice de transition Q. Pour tout $r > 2$, conditionnellement à $T_i^{(r-1)} < \infty$, $S_i^{(r)}$ est indépendant de $\{X_m : m \leq T_i^{(r-1)}\}$ et*

$$\mathbb{P}(S_i^{(r)} = n | T_i^{(r-1)} < \infty) = \mathbb{P}_i(T_i = n).$$

Démonstration. D'après la propriété de Markov forte appliquée au temps d'arrêt $T = T_i^{(r-1)}$, conditionnellement à $T < \infty$ et $X_T = i$, le processus $(\tilde{X}_n = X_{T+n})_{n\geq 0}$ est une chaîne de Markov (δ_i, Q) indépendante de $\mathcal{F}_T$. Alors $S_i^{(r)} = \inf\{n \geq 1, X_{T+n} = i\}$, le premier temps de retour de $(\tilde{X}_n)_{n\geq 0}$ à l'état i, a la même loi que T_i sous $\mathbb{P}_i$. □

Proposition 3.1.17. *Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov de matrice de transition Q, on note V_i le nombre de retours à l'état $i \in E$, i.e :*

$$V_i = \sum_{n=0}^{\infty} 1_{\{X_n = i\}}.$$

Alors,

$$\mathbb{E}_i(V_i) = \sum_{n=0}^{\infty} Q_{ii}^{(n)}. \tag{3.5}$$

Si on note $f_i = \mathbb{P}_i(T_i < \infty)$, alors :

$$\mathbb{P}_i(V_i > r) = f_i^r \tag{3.6}$$

pour tout $r \in \mathbb{N}$.

Démonstration. Pour (3.5), on a :

$$\begin{aligned}\mathbb{E}_i(V_i) &= \mathbb{E}_i \sum_{n=0}^{\infty} 1_{X_n=i} = \sum_{n=0}^{\infty} \mathbb{E}_i(1_{X_n=i}) \\ &= \sum_{n=0}^{\infty} \mathbb{P}_i(X_n = i) = \sum_{n=0}^{\infty} (Q^{(n)})_{ii}\end{aligned}$$

Nous démontrons (3.6) par récurrence, pour $r = 0$ on a $\mathbb{P}_i(V_i > 0) = 1$. Notons que si $X_0 = i$, alors :

$$\{V_i > r\} = \{T_i^{(r)} < \infty\}. \tag{3.7}$$

Donc :

$$\begin{aligned}\mathbb{P}_i(V_i > r+1) &= \mathbb{P}_i(T_i^{(r+1)} < \infty) \\ &= \mathbb{P}_i(T_i^{(r)} < \infty, S_i^{(r+1)} < \infty) \\ &= \mathbb{P}_i(S_i^{(r+1)} < \infty | T_i^{(r)} < \infty)\, \mathbb{P}_i(T_i^{(r)} < \infty)\end{aligned}$$

En utilisant le Lemme 3.1.16 , on a :

$$\mathbb{P}_i(V_i > r+1) = f_i\, f_i^r = f_i^{r+1}\,.$$

□

Corollaire 3.1.18.

(*i*) *Si i est récurrent alors :*

$$\mathbb{P}_i(V_i = +\infty) = 1\,.$$

(*ii*) *Si i est transient alors :*

$$\mathbb{P}_i(V_i = k) = (1 - f_i)\, f_i^k\,.$$

Démonstration. (*i*) Si i est récurrent,

$$\begin{aligned}\mathbb{P}_i(V_i = \infty) &= \lim_{r\to\infty} \mathbb{P}_i(V_i > r) \\ &= \lim_{r\to\infty} [\mathbb{P}_i(T_i < \infty)]^r = 1.\end{aligned}$$

(*ii*) Si i est transient,

$$\begin{aligned}\mathbb{P}_i\,(V_i = k) &= \mathbb{P}_i\,(T_i^{(k)} < \infty, T_i^{(k+1)} = \infty) \\ &= \mathbb{P}_i(T_i^{(k+1)} = \infty | T_i^{(k)} < \infty)\, \mathbb{P}_i(T_i^{(k)} < \infty) \\ &= \mathbb{P}_i(T_i = \infty)\, \mathbb{P}_i(T_i^{(k)} < \infty) \\ &= (1 - \mathbb{P}_i(T_i < \infty))\, \mathbb{P}_i(T_i^{(k)} < \infty) \\ &= (1 - f_i)\, f_i^k.\end{aligned}$$

□

Définition 3.1.19 (état absorbant)**.** *Un état $i \in \mathbb{E}$ est dit absorbant lorsque la probabilité de rester dans ce même état lorsque l'on y parvient est égale à 1, i.e.*

$$Q_{ii} = \mathbb{P}(X_{n+1} = i | X_n = i) = 1\,.$$

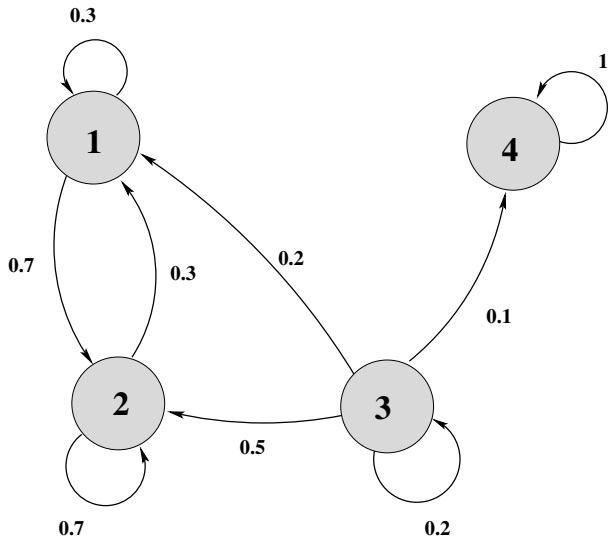

Figure 3.2: *Chaîne de Markov contenant 3 classes* $\{1, 2\}$, $\{3\}$ *et* $\{4\}$. *L'état* 4 *est absorbant, les états* $1, 2$ *et* 4 *sont récurrents et l'état* 3 *est transient*

Exemple 3.1.20. *Soit* $(X_n)_{n\in\mathbb{N}}$ *une chaîne de Markov de matrice de transition*

$$Q = \begin{pmatrix} 0.3 & 0.7 & 0 & 0 \\ 0.3 & 0.7 & 0 & 0 \\ 0.2 & 0.3 & 0.2 & 0.1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

La chaîne comporte 3 classes $\{1, 2\}$, $\{3\}$ *et* $\{4\}$. *L'état* 4 *est absorbant, les états* $1, 2$ *et* 4 *sont récurrents et l'état* 3 *est transient, cf. Figure 3.2.*

3.1.4 Période d'un état

Définition 3.1.21. *Soit* $(X_n)_{n\geq 0}$ *une chaîne de Markov* (μ, Q) *à valeurs dans* E. *La période d'un état* $i \in E$ *est l'entier* $d(i)$ *défini par :*

$$d(i) = \mathrm{PGCD}\{n \geq 1 \text{ tel que } Q_{ii}^{(n)} > 0\}.$$

Lorsque $d(i) = 1$, *l'état* i *est dit apériodique.*

On démontre aussi que :

(*i*) tous les états d'une même classe ont la même période ;

(*ii*) si la chaîne est irréductible et s'il existe un état apériodique, alors tous les états sont apériodiques.

3.2 Chaîne de Markov régulière

3.2.1 Définitions et propriétés

Dans la section précédente, nous avons vu qu'une chaîne de Markov est irréductible si tout état est atteignable depuis tout autre état. Dans ce cas, on dit aussi que la chaîne est ergodique.

Définition 3.2.1 (Chaîne régulière)**.** *Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov de matrice de transition Q. La matrice Q est dite régulière si il existe une puissance $Q^{(m)}$ de Q dont tous les éléments sont strictement positifs. Une chaîne de Markov qui a matrice de transition régulière est dite régulière.*

Remarque 3.2.2. (*i*) *Une chaîne de Markov qui est à la fois irréductible et apériodique est une chaîne de Markov régulière.*

(*ii*) *Une chaîne de Markov régulière est nécessairement irréductible (ou ergodique). En effet, tout état est atteignable depuis un autre en m pas au plus avec m plus petit entier tel que $Q^{(m)} > 0$. La réciproque n'est pas vraie en général.*

Les théorèmes suivants résument les propriétés spectrales d'une matrice régulière.

Théorème 3.2.3 (Théorème de Perron-Frobenius)**.** *Si Q est une matrice régulière, alors il existe une valeur propre $r > 0$ de Q telle que :*

(*i*) *r est une valeur propre simple,*

(*ii*) *pour toute autre valeur propre $\lambda \neq r$, on a $r > |\lambda|$ (r valeur propre dominante),*

(*iii*) *le vecteur propre associé à r est unique à une constante multicatives près.*

On peut trouver une démonstration de cette proposition dans [47].

Proposition 3.2.4. *Si Q est une matrice de transition d'une chaîne de Markov régulière, alors Q admet 1 comme valeur propre dominante.*

Démonstration. Par définition de la matrice de transition, on a

$$\sum_{j\in E} Q_{ij} = 1, \forall i \in E.$$

Le vecteur $(1, \ldots, 1)^*$ est donc un vecteur propre de Q correspondant à la valeur propre 1.

Soit λ une valeur propre de Q de vecteur propre v et soit k tel que :

$$|v_k| = \sup_{j \in E} |v_j| \,,$$

alors :

$$|\lambda\, v_k| = |\sum_j Q_{kj}\, v_j| \leq \sum_j Q_{kj}\, |v_j| \leq \sum_j Q_{kj}\, |v_k| = |v_k| \,,$$

d'où $|\lambda| \leq 1$. □

3.2.2 Loi stationnaire

Définition 3.2.5. *Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov (μ, Q) et $\pi = (\pi_i)_{i\in E}$ une loi de probabilité sur l'espace d'état E. Un vecteur de probabilité π est une loi stationnaire (ou mesure invariante) par la matrice de transition Q, si elle vérifie $\pi\, Q = \pi$. C'est-à-dire si :*

$$\sum_{j\in E} \pi_j\, Q_{ji} = \pi_i$$

pour tout $i \in E$.

Théorème 3.2.6. *Toute chaîne de Markov régulière admet une mesure invariante strictement positive unique à une constante multiplicative près, i.e. toutes les mesures invariantes sont proportionnelles entre elles.*

Démonstration. Si elle existe, la loi invariante est le vecteur propre à gauche associé à la valeur propre dominante 1. Le théorème de Perron-Frobenius appliqué à Q' (matrice transposée de Q) montre que l'espace associé à sa valeur propre 1 est engendré par un vecteur propre v dont toutes les composantes sont strictement positives. En le normalisant, on obtient un vecteur de probabilité $\pi > 0$, qui est la seule loi invariante. □

Théorème 3.2.7. *Si Q est la matrice de transition d'une chaîne de Markov régulière admettant π comme loi invariante, alors :*

$$Q_{ij}^{(n)} \underset{n\to\infty}{\longrightarrow} \pi_j \;\forall i, j \in E. \tag{3.8}$$

En d'autres termes :

$$\lim_{n\to\infty} \mathbb{P}(X_n = j) = \pi_j$$

pour tout $j \in E$.

Démonstration. Nous reprenons l'argument de couplage utilisé par Norris [37].

Soit $(Y_n)_{n\in\mathbb{N}}$ une chaîne de Markov régulière de matrice de transition Q, indépendante de $(X_n)_{n\in\mathbb{N}}$. Fixons un état de référence $b \in E$ et posons :

$$T = \inf\{n \geq 1 : X_n = Y_n = b\}\,.$$

La démonstration se fait en trois étapes.

Etape (i) – Montrons que $\mathbb{P}(T < \infty) = 1$. Le processus $(W_n)_{n\in\mathbb{N}}$ défini par :

$$W_n = (X_n, Y_n)_{n\,\in\,\mathbb{N}}$$

est une chaîne de Markov à valeurs dans $E \times E$ et de matrice de transition $\tilde{Q}$ définie par

$$\tilde{Q} = \tilde{Q}_{(i,k)(j,l)} = Q_{ij}\, Q_{kl},$$

et de loi initiale λ définie par :

$$\lambda_{(i,k)} = \mu_i\, \alpha_k$$

où μ est la loi initiale de $(X_n)_{n\in\mathbb{N}}$ et α celle de $(Y_n)_{n\in\mathbb{N}}$.

Comme Q est régulière, il existe $N \in \mathbb{N}$ tel que $\forall\, n \geq N$,

$$\tilde{Q}^{(n)}_{(i,k)(j,l)} = Q^{(n)}_{ij}\, Q^{(n)}_{kl} > 0$$

pour tout $i, j, k, l \in I$. Par conséquent $\tilde{Q}$ est aussi régulière et possède une loi invariante (Théorème 3.2.6). D'après la Remarque 3.1.14, le premier temps de retour T à l'état (b, b) est fini p.s.

Etape (ii) – Posons

$$Z_n = \begin{cases} X_n & \text{si } n < T, \\ Y_n & \text{si } n \geq T. \end{cases}$$

Montrons que le processus $(Z_n)_{n\in\mathbb{N}}$ est une chaîne de Markov (μ, Q). Notons d'abord que d'après la propriété de Markov, les processus $(X_{T+n})_{n\geq 0}$ et $(Y_{T+n})_{n\geq 0}$ sont des chaînes de Markov (δ_b, Q), indépendantes de $\mathcal{F}_T$. Par conséquent $(Z_n)_{n\geq 0}$ est une chaîne de Markov (μ, Q).

Etape (iii) – Comme $(Z_n)_{n\in\mathbb{N}}$ et $(X_n)_{n\in\mathbb{N}}$ sont deux chaînes de Markov qui ont même loi, alors :

$$\mathbb{P}(Z_n = j) = \mathbb{P}(X_n = j)\,.$$

Comme $\pi_j = \mathbb{P}(Y_n = j)$, donc :

$$\begin{aligned}
\mathbb{P}(X_n = j) - \pi_j &= \mathbb{P}(Z_n = j) - \mathbb{P}(Y_n = j) \\
&= \mathbb{P}(Z_n = j, n < T) + \mathbb{P}(Z_n = j, n \geq T) \\
&- \mathbb{P}(Y_n = j, n < T) - \mathbb{P}(Y_n = j, n \geq T), \\
&= \mathbb{P}(Z_n = j, n < T) - \mathbb{P}(Y_n = j, n < T) \\
&\leq \mathbb{P}(Z_n = j, n < T) \leq \mathbb{P}(n < T)
\end{aligned}$$

et

$$\mathbb{P}(Y_n = j, n < T) - \mathbb{P}(Z_n = j, n < T) \geq \mathbb{P}(Y_n = j, n < T) < \mathbb{P}(n < T),$$

alors :

$$|P(X_n = j) - \pi_j| \leq \mathbb{P}(n < T).$$

Mais $\mathbb{P}(T < \infty) = 1$, donc $\mathbb{P}(n < T) \to 0$ lorsque $n \to \infty$, d'où :

$$\mathbb{P}(X_n = j) \to \pi_j.$$

□

Théorème 3.2.8. *Soit $(X_n)_{n\geq 0}$ une chaîne de Markov de matrice de transition Q et π une mesure positive. Si :*

$$\lim_{n\to\infty} Q_{ij}^{(n)} = \pi_j$$

pour tout $i \in E$, alors π est une loi invariante pour Q.

Démonstration. On a :

$$\sum_{j\in E} \pi_j = \sum_{j\in E} \lim_{n\to\infty} (Q_{ij}^{(n)}) = \lim_{n\to\infty} \sum_{j\in E} Q_{ij}^{(n)} = 1$$

donc π est une mesure de probabilité. Par ailleurs :

$$\begin{aligned}
\pi_j &= \lim_{n\to\infty} Q_{ij}^{(n)} = \lim_{n\to\infty} \sum_{k\in E} Q_{ik}^{(n-1)}\, Q_{kj}, \\
&= \sum_{k\in E} \lim_{n\to\infty} Q_{ik}^{(n-1)}\, Q_{kj} = \sum_{k\in E} \pi_k\, Q_{kj}
\end{aligned}$$

c'est-à-dire $\pi = \pi\, Q$. □

Pour une chaîne de Markov, on peut aussi établir une extension de la loi des grands nombres.

Théorème 3.2.9 (Théorème ergodique)**.** *Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov (λ, Q). Si Q est irréductible, alors :*

$$\lim_{n\to\infty} \frac{V_i(n)}{n} = \frac{1}{\mathbb{E}_i(T_i)} \ p.s$$

où

$$V_i(n) = \sum_{k=0}^{n-1} 1_{\{X_k=i\}}$$

désigne le nombre de visites à l'état i avant d'arriver à l'état n. Si de plus Q est récurrente positive et $(\pi_i)_{i\in E}$ est son unique probabilité invariante, alors pour toute fonction bornée $f : E \to \mathbb{R}$ on a

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) = \bar{f} \ p.s,$$

où

$$\bar{f} = \sum_{i\in E} \pi_i f_i \,.$$

Démonstration. Etape (i) – Si Q est transient et si le nombre total de visites à l'état i V_i est finie p.s., alors :

$$\frac{V_i(n)}{n} \leq \frac{V_i}{n} \to 0 = \frac{1}{\mathbb{E}_i(T_i)} \,.$$

Etape (ii) – Supposons maintenant que Q est récurrente et fixons un état i quelconque ; pour $T = T_i$ on a $\mathbb{P}(T_i < \infty) = 1$ et $(X_{T+n})_{n\geq 0}$ est une chaîne de Markov (δ_i, Q) indépendante de $\mathcal{F}_T$.

Pour les deux chaînes $(X_n)_{n\in\mathbb{N}}$ et $(X_{T+n})_{n\in\mathbb{N}}$ le temps nécessaire pour atteindre i après le premier passage est le même, alors on peut prendre le cas $\lambda = \delta_i$.

Désignons par $S_i^{(r)}$ la longueur de la $r^{\text{ème}}$ excursion dans l'état i. D'après le Lemme 3.1.16, les variables aléatoires $(S_i^{(r)})_{r\in\mathbb{N}}$ sont indépendantes et identiquement distribuées. On a :

$$\mathbb{E}_i(S_i^{(r)}) = \mathbb{E}_i(T_i), \tag{3.9}$$

$$S_i^{(1)} + \cdots + S_i^{(V_i(n)-1)} \leq n-1, \tag{3.10}$$

$$S_i^{(1)} + \cdots + S_i^{(V_i(n))} \geq n. \tag{3.11}$$

De (3.10) et (3.11), nous pouvons écrire :

$$\frac{S_i^1 + \cdots + S_i^{(V_i(n)-1)}}{V_i(n)} \leq \frac{n}{V_i(n)} \leq \frac{S_i^{(1)} + \cdots + S_i^{(V_i(n))}}{V_i(n)}. \tag{3.12}$$

D'après la loi des grands nombres :

$$\frac{S_i^{(1)} + \cdots + S_i^{(n)}}{n} \underset{n\to\infty}{\longrightarrow} \mathbb{E}_i(T_i) \quad \text{p.s.}$$

Comme Q est récurrente, alors $V_i(n) \to \infty$ p.s, en laissant $n \to \infty$, la relation (3.12) nous permet d'écrire :

$$\frac{n}{V_i(n)} \underset{n\to\infty}{\longrightarrow} \mathbb{E}_i(T_i) \quad \text{p.s,}$$

d'où :

$$\frac{V_i(n)}{n} \underset{n\to\infty}{\longrightarrow} \frac{1}{\mathbb{E}_i(T_i)} \quad \text{p.s.}$$

Etape (iii) – Enfin si Q est récurrente positive, il existe une probabilité invariante $\pi_{i_{(i\in E)}}$. Soit $f : E \to \mathbb{R}$ une fonction bornée telle que $|f| \leq 1$. Posons :

$$\bar{f} = \sum_{i\in E} \pi_i \, f_i$$

pour tout $F \subset E$. Alors,

$$\begin{aligned}
\left|\frac{1}{n}\sum_{k=0}^{n-1} f(X_k) - \bar{f}\right| &= \left|\sum_{i\in F}(\frac{V_i(n)}{n} - \pi_i)f_i\right| \leq \sum_{i\in F}\left|\frac{V_i(n)}{n} - \pi_i\right| + \sum_{i\notin F}\left|\frac{V_i(n)}{n} - \pi_i\right|, \\
&\leq \sum_{i\in F}\left|\frac{V_i(n)}{n} - \pi_i\right| + \sum_{i\notin F}\left|\frac{V_i(n)}{n} + \pi_i\right|, \\
&\leq 2\sum_{i\in F}\left|\frac{V_i(n)}{n} - \pi_i\right| + 2\sum_{i\notin F}\pi_i. \qquad (3.13)
\end{aligned}$$

Comme on a prouvé que :

$$\frac{V_i(n)}{n} \underset{n\to\infty}{\longrightarrow} \pi_i \quad \text{, p.s,}$$

pour $\varepsilon > 0$ donné, choisissons F fini tel que :

$$\sum_{i\notin F} \pi_i \leq \frac{\varepsilon}{4},$$

puis $N = N(\omega)$ tel que

$$\sum_{i\in F}\left|\frac{V_i(n)}{n} - \pi_i\right| < \frac{\varepsilon}{4}.$$

pour tout $n \geq N$. La relation (3.13) devient :

$$\left|\frac{1}{n}\sum_{k=0}^{n-1} f(X_k) - \bar{f}\right| < \varepsilon$$

pour tout $n \geq N$. □

Définition 3.2.10 (Chaîne réversible)**.** *Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov de matrice de transition Q dans l'espace d'état E. Une probabilité π sur E est dite réversible pour Q, ou pour la chaîne de Markov $(X_n)_{n\in\mathbb{N}}$, si :*

$$\pi_i\, Q_{ij} = \pi_j\, Q_{ji} \quad \forall i, j \in E. \tag{3.14}$$

Proposition 3.2.11. *Si une loi de probabilité π est réversible par rapport à une matrice de transition Q, alors elle est invariante par Q.*

Démonstration. En sommant l'équation (3.14) sur $j \in E$, on déduit que :

$$\pi_i = \sum_j \pi_i Q_{ij} = \sum_j \pi_j Q_{ji}.$$

D'où, π est invariante. □

3.3 Chaîne de Markov absorbante

3.3.1 Définitions et propriétés

Définition 3.3.1. *Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov (μ, Q) dans E. La chaîne $(X_n)_{n\in\mathbb{N}}$ est dite absorbante si elle contient au moins un état absorbant.*

Dans la suite de cette section, nous supposons que $(X_n)_{n\in\mathbb{N}}$ admet $r \geq 1$ états absorbants. Réarrangeons les états de E de manière à placer d'abord les p états non absorbants, ensuite les r états absorbants. La matrice de transition Q peut s'écrire sous la forme :

$$Q = \begin{pmatrix} P & R \\ 0 & I \end{pmatrix} \tag{3.15}$$

où P est une matrice $p \times p$, R une matrice $p \times r$, 0 désigne la matrice nulle de taille $r \times p$ et I la matrice identité de taille r. Cette représentation s'appelle *forme canonique* de la matrice de transition Q.

Proposition 3.3.2. *Soit Q la matrice de transition d'une chaîne de Markov absorbante, écrite sous la forme canonique, alors :*

$$\lim_{n\to\infty} P^n = 0, \tag{3.16}$$

et la matrice $I - P$ est inversible et son inverse est :

$$[I - P]^{-1} = \sum_{k=0}^{\infty} P^k. \tag{3.17}$$

Démonstration. Pour tout $i \in E$, soit :

$$m_i \stackrel{\text{déf}}{=} \min\{n \geq 1 : \exists\, k > p,\ (Q^n)_{ik} > 0\}$$

le nombre minimal de pas nécessaire pour atteindre un état absorbant k depuis i. Notons :

$$q_i = \mathbb{P}_i(X_{m_i} \leq p) < 1$$

la probabilité de ne pas atteindre d'état absorbant en m_i pas, partant de i.

Remarquons que $P_{ij}^{(n)}$ est la probabilité de se trouver dans un état non absorbant j, après n pas, partant de i, qui est inférieure ou égale à p_i.

En prenant :

$$M = \max_{i=1\ldots p} m_i, \text{ et } q = \max_{i=1,\ldots,p} q_i,$$

on a $P^M \leq q$. En particulier, $\forall n \in \mathbb{N}$:

$$P^{Mn} \leq q^n \to 0 \text{ quand } n \to \infty.$$

Pour l'inverse de $I - P$, considérons la forme canonique de Q. En effectuant une multiplication itérative, on a :

$$Q^2 = \begin{pmatrix} P & R \\ 0 & I \end{pmatrix} \begin{pmatrix} P & R \\ 0 & I \end{pmatrix} = \begin{pmatrix} P^2 & R + PR \\ 0 & I \end{pmatrix},$$
$$Q^3 = \begin{pmatrix} P^2 & R + PR \\ 0 & I \end{pmatrix} \begin{pmatrix} P & R \\ 0 & I \end{pmatrix} = \begin{pmatrix} P^3 & R + PR + P^2R \\ 0 & I \end{pmatrix}.$$

et par récurrence, on obtient :

$$Q^n = \begin{pmatrix} P^n & (1 + P + P^2 + \cdots + P^{n-1})R \\ 0 & I \end{pmatrix}.$$

En multipliant $(I - P)$ par $\sum_{k=0}^{n-1} P^k$, on a :

$$[I - P] \sum_{k=0}^{n-1} P^k = I - P^n.$$

Comme $P^n \to 0$ quand $n \to \infty$, on a

$$(I - P) \sum_{k=0}^{\infty} P^k = I$$

d'où

$$[1 - P]^{-1} = \sum_{k=0}^{\infty} P^k.$$

□

La matrice $[I-P]^{-1}$, s'appelle *matrice fondamentale* de la chaîne de Markov absorbante, et on la notera F. En utilisant l'équation (3.15), on montre que :

$$\lim_{n\to\infty} Q^n = \begin{pmatrix} 0 & FR \\ 0 & I \end{pmatrix}. \tag{3.18}$$

La matrice FR représente les probabilités de transition d'un état non absorbant vers un autre état absorbant. Les éléments de F correspondent aux états non absorbants de la chaîne. Ainsi, l'équation (3.18) traduit bien le fait que la probabilité d'absorption tend vers 1 lorsque le temps tend vers l'infini.

Proposition 3.3.3. *Soit F la matrice fondamentale d'une chaîne de Markov absorbante $(X_n)_{n\in\mathbb{N}}$. L'élément f_{ij} de F est l'espérance du nombre de passages en j partant de i, c'est-à-dire :*

$$f_{ij} = \mathbb{E}_i(\sum_{n\geq 0} 1_{X_n=j}). \tag{3.19}$$

Démonstration. Posons $Y_n = 1_{X_n=j}$, alors :

$$\mathbb{E}_i(Y_n) = \mathbb{P}_i(Y_n = 1) = (Q^n)_{ij}, \text{ et } \mathbb{E}_i(\sum_{n\geq 0} Y_n) = \sum_{n\geq 0}(Q^n)_{ij} = f_{ij}.$$

□

Proposition 3.3.4 (Temps d'absorption)**.** *En écrivant la matrice de transition sous sa forme cannonique, soit $\tau \overset{\text{déf}}{=} \inf\{n \geq 1 : X_n > p\}$ le temps jusqu'à absorption, alors*

$$\mathbb{E}_i(\tau) = \sum_{j=1}^{p} f_{ij}. \tag{3.20}$$

Démonstration. Pour tout état non absorbant j, on a :

$$\begin{aligned}\sum_{j=1}^{p} f_{ij} &= \sum_{j=1}^{p} \mathbb{E}_i(\sum_{n\geq 0} Y_n) = \mathbb{E}_i(\sum_{n\geq 0} 1_{X_n\leq p}),\\ &= \sum_{n\geq 0} \mathbb{P}_i(\tau > n) = \sum_{n\geq 0} n\,\mathbb{P}_i(\tau = n) = \mathbb{E}_i(\tau).\end{aligned}$$

□

3.3.2 Loi quasi-stationnaire

Dans la forme canonique (3.15) de Q , la matrice P désigne la loi du processus $(X_n)_{n\in\mathbb{N}}$ avant l'absorption. Notons A l'ensemble des états non absorbants de E. Munie de la loi initiale μ et de la matrice $\tilde{P}$ définie par :

$$\tilde{P}_{ij} = \frac{P_{ij}}{\sum_{j\in F} P_{ij}} \; i \in F\,, \tag{3.21}$$

elle constitue un chaîne de Markov régulière $(\tilde{X}_n)_{n\in\mathbb{N}}$ sur A [37]. Cette chaîne admet une loi limite (invariante) qui représente les proportions des temps passés sur chaque état avant que le processus soit absorbé. Cette loi limite s'appelle *loi quasi-stationnaire* de la chaîne absorbante $(X_n)_{n\in\mathbb{N}}$ [15, 58].

3.4 Principe des méthodes de monte carlo par chaîne de Markov (MCMC)

Les chaînes de Markov ne sont pas seulement un outil de modélisation, elle peuvent également être utilisée comme outil d'approximation numérique. Dans cette thèse nous faisons appel aux méthodes de Monte Carlo par chaînes de Markov ou MCMC (Markov Chain Monte Carlo), nous en rappelons maintenant les grands principes, on pourra se référer à [32] pour plus de détails. Notons ici que les chaînes de Markov utilisées sont décrites dans le cas d'espaces d'état continu dans la mesure où elles vont nous permettre d'échantillonner selon des lois continues.

3.4.1 Principe de base

Les méthodes MCMC rassemblent les méthodes qui consistent à simuler une loi de probabilité, dite loi cible, à partir d'une chaîne de Markov ergodique. Ces techniques ont été introduites par Metropolis *et al.* [34]. Avec les progrès de l'informatique, elles ont grandement diffusé dans le domaine de la statistique bayésienne à partir des années 1990 [18], elles permettent en effet d'échantillonner numériquement selon la loi a posteriori.

On souhaite calculer l'intégrale :

$$\mathcal{I}(f) = \mathbb{E}[f(X)] = \int f(x)\pi(x), \tag{3.22}$$

où π est une densité connue à une constante multiplicative près et f une fonction bornée d'un espace d'état E dans $\mathbb{R}$ (ici E est supposé continu).

Les méthodes MCMC permettent de créer une chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ dont les échantillons sont distribués asymptotiquement selon la distribution requise π. La convergence est assurée par le théorème ergodique 3.2.9 appliqué à $(X_n)_{n\in\mathbb{N}}$:

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f(X_k) = \sum_{i\in E} \pi_i f_i = \mathbb{E}_\pi f(X), \quad X \sim \pi\,.$$

3.4.2 L'algorithme de Metropolis–Hastings

L'algorithme de Metropolis-Hastings[21] est l'un des algorithmes MCMC les plus utilisés. On se donne (la densité d') une "loi instrumentale" ou " loi de

proposition" q qui doit être simulable rapidement et disponible analytiquement à une constante près. Partant d'une valeur initiale x_0, on construit x_n, $n \geq 1$ en deux étapes :

(*i*) proposition : on simule $y \sim q(\cdot|x_{n-1})$,

(*ii*) acceptation/rejet :

$$x_n = \begin{cases} y & \text{avec la probabilité } \alpha(x_{n-1}, y)\,, \\ x_{n-1} & \text{avec la probabilité } 1 - \alpha(x_{n-1}, y) \end{cases}$$

où

$$\alpha(x, y) = \min\left(1, \frac{\pi(y)\, q(x|y)}{\pi(x)\, q(y|x)}\right)\,.$$

est la probabilité d'acceptation. Cette dernière probabilité est choisie de telle façon que π soit réversible pour la chaîne de Markov $(x_n)_{n\in\mathbb{N}}$. La loi de probabilité π sera donc invariante et la chaîne $(x_n)_{n\in\mathbb{N}}$ est irréductible, récurrente positive et apériodique.

Si la loi de proposition est symétrique, i.e. $q(y|x) = q(x|y)$, alors :

$$\alpha(x, y) = \min\left(1, \frac{\pi(y)}{\pi(x)}\right)$$

et on obtient la procédure de *Metropolis* décrite dans l'algorithme 3.1.

Si q n'est pas symétrique, on obtient la procédure de Metropolis-Hastings détaillée dans l'algorithme 3.2.

$X_0 \leftarrow x$,
pour $n = 1, 2, \ldots$ **faire**
  simuler $y \sim q(\cdot|X_{n-1})$
  simuler $u \sim \mathcal{U}[0, 1]$
  $\alpha \leftarrow \pi(y)/\pi(X_{n-1})$
  si $u \leq \alpha$ **alors**
    $X_n \leftarrow y$
  sinon
    $X_n \leftarrow X_{n-1}$
  fin si
fin pour

Alg. 3.1: *Algorithme de Metropolis.*

$X_0 \leftarrow x$
pour $n = 1, 2, \ldots$ **faire**
  simuler $y \sim q(\cdot|X_{n-1})$
  simuler $u \sim \mathcal{U}[0,1]$
  $\alpha \leftarrow (\pi(y)\, q(X_{n-1}|y))/(\pi(X_{n-1})\, q(y|X_{n-1}))$
  si $u \leq \alpha$ **alors**
    $X_n \leftarrow y$
  sinon
    $X_n \leftarrow X_{n-1}$
  fin si
fin pour

ALG. 3.2: *Algorithme de Metropolis-Hastings*

Chapitre 4

Modèles markoviens

L'utilisation des chaînes de Markov dans la modélisation des dynamiques d'occupation des sols n'est pas récente [54, 55, 59]. Le succès de cette approche s'explique par sa simplicité mais également par le fait qu'en agronomie ou en écologie, les dynamiques peuvent être représentées par une succession discrète d'un nombre fini d'états.

Dans ce chapitre, nous utilisons les chaînes de Markov pour modéliser la dynamique d'usages des sols correspondant aux parcelles dans la village d'Ambendrana. La matrice de transition de la chaîne sera estimée à partir de l'historique d'usages des parcelles pendant la période d'observation. Nous reprenons en premier lieu la technique proposée par Anderson et Goodman [3] basée sur le maximum de vraisemblance. Nous considérons également l'approche bayésienne [51, 33]. Ces deux méthodes nous permettent d'avoir deux modèles de Markov différents que l'on étudiera séparément.

4.1 Description du modèle

4.1.1 Les observations

Nous considérons le premier jeu de données du tableau 2.1. Le jeu de données représente la succession d'occupation des 42 parcelles pendant 22 ans (1986 à 2006). On a recensé quatre modes d'utilisation qui correspondent aux états des parcelles : forêt naturelle (`Fn`), culture annuelle (`Ca`), jachère (`Jc`), et culture pérenne (`Cp`). Dans un premier temps, nous avons considéré tout ces quatre états pour développer notre modèle [10, 42]. Puis, après avoir analyser le jeu de données concerné, nous constatons que l'état "forêt naturelle" (`Fn`) est un état totalement transient. Plus précisément, dès que le système quitte cet état, il n' y revient plus.

Ainsi nous avons décidé d'omettre cet état dans notre modèle[9]. On s'intéresse à la dynamique d'une parcelle à partir de sa première année de mise en culture. L'espace d'état sera alors :

$$E \stackrel{\text{déf}}{=} \{\texttt{Ca}, \texttt{Jc}, \texttt{Cp}\}. \tag{4.1}$$

Pendant les 22 ans d'observations, les parcelles n'ont pas toutes la même année de mise en culture ; par exemple, la parcelle n° 1 a été mise en culture en 2000, les parcelles n° 2, 3, 4 en 2001 etc. Par conséquent, les 42 parcelles n'ont pas la même longueur d'observation.

Soit N_p la longueur d'observation pour la parcelle p. En notant $e_n^{(p)}$, l'état de la parcelle p au temps n, l'ensemble des observations s'écrit :

$$\left(e_{n=1:N_p}^{(p)}\right)_{p=1:42}. \tag{4.2}$$

Nous faisons les hypothèses suivantes :

(h_1) Les parcelles sont indépendantes les unes des autres et ont la même dynamique.

(h_2) La succession d'états d'une parcelle est dirigée par une chaîne de Markov homogène.

En d'autres termes, ces deux hypothèses signifient que les observations sont issues d'une même $(Q, \delta_{\texttt{Ca}})$-chaîne de Markov $(X_n)_{n\geq 1}$ homogène, i.e. :

$$\mathbb{P}(X_n^{(p)} = e_n^{(p)} | X_{0:n-1}^{(p)} = e_{0:n-1}^{(p)}) = Q(e_{n-1}^{(p)}, e_n^{(p)})$$

et $X_0 \sim \delta_{\texttt{Ca}}$.

Il faut noter que ces hypothèses sont partiellement irréalistes. En effet, dans la réalité, la dynamique d'une parcelle donnée dépend de plusieurs mécanismes : les décisions des agriculteurs, l'exposition, la pente et la distance à la forêt, etc. Il y a forcement une dépendance entre parcelles voisines. Cependant, les hypothèses (h_1) et (h_2) permettent d'obtenir des résultats préliminaires importants.

4.1.2 Caractérisation du modèle

Pour caractériser le modèle, il faut définir la loi initiale μ et la matrice de transition Q de la chaîne de Markov $(X_n)_{n\geq 1}$. Comme l'observation commence à la première année de mise en culture, on peut prendre comme état initial l'état `Ca`, i.e. :

$$\mu \stackrel{\text{déf}}{=} \delta_{\texttt{Ca}}. \tag{4.3}$$

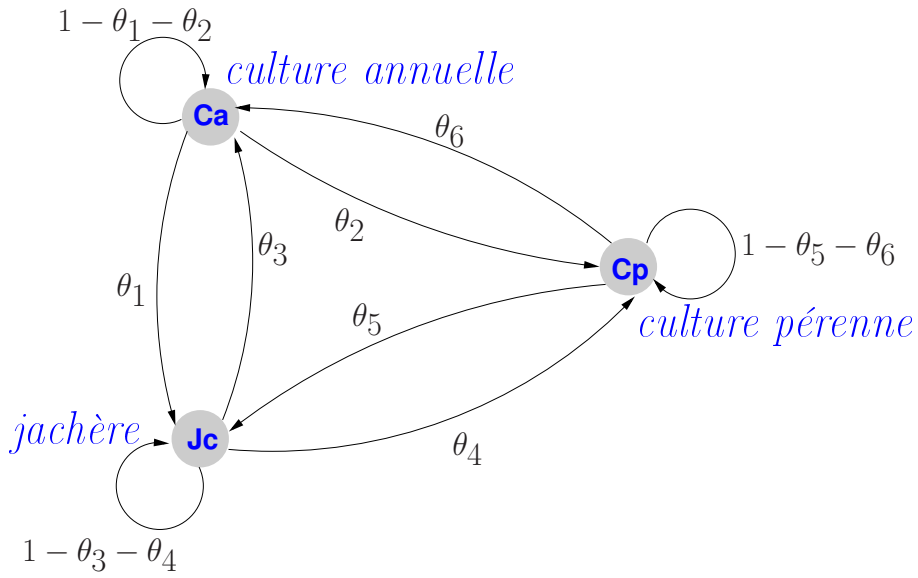

Figure 4.1: *Modèle markovien à 3 états : "culture annuelle"* (`Ca`), *"jachère"* (`Jc`), *et "culture pérenne"* (`Cp`) .

La loi de processus s'écrit alors :

$$\begin{aligned}\mathbb{P}(X_{1:N_p}^{(p)} = e_{1:N_p}^{(p)}\,,\ \forall p = 1,\dots,42) &= \prod_{p=1}^{42} \mathbb{P}(X_{1:N_p}^{(p)} = e_{1:N_p}) \\ &= \prod_{p=1}^{P} \delta_{\mathtt{Ca}}\,(e_1^{(p)})\; Q(e_1^{(p)}, e_2^{(p)}) \cdots Q(e_{N-2}^{(p)}, e_{N-1}^{(p)})\end{aligned} \tag{4.4}$$

pour tout $e_n^{(p)} \in E$, où $Q = [Q(i,j)]_{i,j\in\mathbb{E}}$.

La matrice de transition Q s'écrit sous la forme :

$$Q = [Q(i,j)]_{i,j\in E} = \begin{pmatrix} 1-\theta_1-\theta_2 & \theta_1 & \theta_2 \\ \theta_3 & 1-\theta_3-\theta_4 & \theta_4 \\ \theta_5 & \theta_6 & 1-\theta_5-\theta_6 \end{pmatrix}, \tag{4.5}$$

où $\theta = (\theta_i)_{1\le i\le 6}$ est le paramètre inconnu appartenant à l'ensemble Θ défini par

$$\Theta \overset{\text{déf}}{=} \{\theta \in [0,\,1]^6\,;\ \theta_1+\theta_2 \le 1,\, \theta_3+\theta_4 \le 1,\, \theta_5+\theta_6 \le 1\}. \tag{4.6}$$

Dans la suite, nous désignerons par $\mathbb{P}_\theta$ la probabilité sous laquelle, la chaîne de Markov $(X_n)_{n\ge1}$ admet Q de paramètre θ comme matrice de transition.

4.2 Inférence des modèles

La loi de processus (4.4) est caractérisée par la matrice de transition Q qui a pour paramètre $\theta = (\theta_i)_{1\leq i\leq 6} \in \Theta$. L'inférence du modèle consiste donc à estimer ce paramètre θ de Q à partir des observations $\left(e^{(p)}_{n=1:N_p}\right)_{p=1:42}$. Nous allons considérer deux méthodes différentes : la méthode de vraisemblance et la méthode bayésienne.

4.2.1 Méthode du maximum de vraisemblance

Dans cette section, nous allons utiliser les résultats de Anderson et Goodman [3], pour estimer la matrice de transition Q de la chaîne de Markov $(X_n)_{n\in\mathbb{N}}$.

Pour une parcelle p, soit :

$$n^{(p)}_{ij} \overset{\text{déf}}{=} \sum_{n=2}^{N-1} 1_{\{X^{(p)}_{n-1}=i\}}\, 1_{\{X^{(p)}_{n}=j\}} \qquad \forall i,j \in E \tag{4.7}$$

le nombre des transitions observées d'un état i vers un état j. Pour l'ensemble des parcelles, le nombre total des transitions entre i et j est :

$$n_{ij} \overset{\text{déf}}{=} \sum_{p=1}^{42} n^{(p)}_{ij}\,. \tag{4.8}$$

Fonction de vraisemblance

La fonction de vraisemblance associée aux observations $(e^{(p)}_{n=1:N_p})_{p=1:42}$ est la fonction L définie par :

$$L:\ \theta \longmapsto L(\theta) \overset{\text{déf}}{=} \prod_{p=1}^{42} \mathbb{P}_\theta(X^{(p)}_{1:N_p} = e^{(p)}_{1:N_p}) \tag{4.9}$$

et

$$\begin{aligned} L(\theta) &= \prod_{p=1}^{42} Q(1,1)^{n^{(p)}_{\texttt{CaCa}}}\, Q(1,2)^{n^{(p)}_{\texttt{CaJc}}}\, Q(1,3)^{n^{(p)}_{\texttt{CaCp}}}\, Q(2,1)^{n^{(p)}_{\texttt{JcCa}}}\, Q(2,2)^{n^{(p)}_{\texttt{JcJc}}} \\ &\qquad Q(2,3)^{n^{(p)}_{\texttt{JcCp}}}\, Q(3,1)^{n^{(p)}_{\texttt{CpCa}}}\, Q(3,2)^{n^{(p)}_{\texttt{CpJc}}}\, Q(3,3)^{n^{(p)}_{\texttt{CpCp}}} \\ &= \prod_{p=1}^{42} (1-\theta_1-\theta_2)^{n^{(p)}_{\texttt{CaCa}}}\, \theta_1^{n^{(p)}_{\texttt{CaJc}}}\, \theta_2^{n^{(p)}_{\texttt{CaCp}}}\, \theta_3^{n^{(p)}_{\texttt{JcCa}}}\, (1-\theta_3-\theta_4)^{n^{(p)}_{\texttt{JcJc}}} \\ &\qquad \theta_4^{n^{(p)}_{\texttt{JcCp}}}\, \theta_5^{n^{(p)}_{\texttt{CpCa}}}\, \theta_6^{n^{(p)}_{\texttt{CpJc}}}\, (1-\theta_5-\theta_6)^{n^{(p)}_{\texttt{CpCp}}}. \end{aligned}$$

En utilisant (4.8), on a :

$$L(\theta) = (1-\theta_1-\theta_2)^{n_{\texttt{CaCa}}}\ \theta_1^{n_{\texttt{CaJc}}}\ \theta_2^{n_{\texttt{CaCp}}}\ \theta_3^{n_{\texttt{JcCa}}}\ (1-\theta_3-\theta_4)^{n_{\texttt{JcJc}}} \\ \theta_4^{n_{\texttt{JcCp}}}\ \theta_5^{n_{\texttt{CpCa}}}\ \theta_6^{n_{\texttt{CpJc}}}\ (1-\theta_5-\theta_6)^{n_{\texttt{CpCp}}}\ . \tag{4.10}$$

La fonction de log–vraisemblance est :

$$\begin{aligned} l(\theta) \stackrel{\text{déf}}{=} \log L(\theta) = {} & n_{\texttt{CaCa}} \log(1-\theta_1-\theta_2) + n_{\texttt{CaJc}} \log(\theta_1) + n_{\texttt{CaCp}} \log(\theta_2) \\ & + n_{\texttt{JcCa}} \log(\theta_3) + n_{\texttt{JcJc}} \log(1-\theta_3-\theta_4) + n_{\texttt{JcCp}} \log(\theta_4) \\ & + n_{\texttt{CpCa}} \log(\theta_5) + n_{\texttt{CpJc}} \log(\theta_6) + n_{\texttt{CpCp}} \log(1-\theta_5-\theta_6)\,. \end{aligned}$$

Estimateur du maximum de vraisemblance (EMV)

L'estimateur du maximum de vraisemblance $\hat{\theta}$ est la valeur θ qui maximise la fonction de vraisemblance $L(\theta)$, c'est-à-dire :

$$L(\hat{\theta}) = \max_{\theta \in \Theta} L(\theta).$$

En considérant la fonction log-vraisemblance, on peut calculer $\hat{\theta}$ en résolvant l'équation de la vraisemblance :

$$\frac{\partial l(\theta)}{\partial \theta}|_{\theta=\hat{\theta}} = 0, \tag{4.11}$$

avec :

$$\begin{aligned} \frac{\partial l(\theta)}{\partial \theta_1} &= -\frac{n_{\texttt{CaCa}}}{(1-\theta_1-\theta_2)} + \frac{n_{\texttt{CaJc}}}{\theta_1}, & \frac{\partial l(\theta)}{\partial \theta_2} &= -\frac{n_{\texttt{CaCa}}}{(1-\theta_1-\theta_2)} + \frac{n_{\texttt{CaCp}}}{\theta_2}, \\ \frac{\partial l(\theta)}{\partial \theta_3} &= -\frac{n_{\texttt{JcJc}}}{(1-\theta_3-\theta_4)} + \frac{n_{\texttt{JcCa}}}{\theta_3}, & \frac{\partial l(\theta)}{\partial \theta_4} &= -\frac{n_{\texttt{JcJc}}}{(1-\theta_1-\theta_2)} + \frac{n_{\texttt{JcCp}}}{\theta_2}, \\ \frac{\partial l(\theta)}{\partial \theta_5} &= -\frac{n_{\texttt{CpCp}}}{(1-\theta_5-\theta_6)} + \frac{n_{\texttt{CpCa}}}{\theta_5}, & \frac{\partial l(\theta)}{\partial \theta_6} &= -\frac{n_{\texttt{CpCp}}}{(1-\theta_5-\theta_6)} + \frac{n_{\texttt{CpJc}}}{\theta_6}. \end{aligned}$$

Après calcul, on trouve :

$$\begin{aligned} \hat{\theta}_1 &= \frac{n_{\texttt{CaJc}}}{(n_{\texttt{CaCa}} + n_{\texttt{CaJc}} + n_{\texttt{CaCp}})}\,, & \hat{\theta}_2 &= \frac{n_{\texttt{CaCp}}}{(n_{\texttt{CaCa}} + n_{\texttt{CaJc}} + n_{\texttt{CaCp}})}\,, \\ \hat{\theta}_3 &= \frac{n_{\texttt{JcCa}}}{(n_{\texttt{JcCa}} + n_{\texttt{JcJc}} + n_{\texttt{JcCp}})}\,, & \hat{\theta}_4 &= \frac{n_{\texttt{JcCp}}}{(n_{\texttt{JcCa}} + n_{\texttt{JcJc}} + n_{\texttt{JcCp}})}\,, \\ \hat{\theta}_5 &= \frac{n_{\texttt{CpCa}}}{(n_{\texttt{CpCa}} + n_{\texttt{CpCa}} + n_{\texttt{CpCp}})}\,, & \hat{\theta}_6 &= \frac{n_{\texttt{CpJc}}}{(n_{\texttt{CpCa}} + n_{\texttt{CpJc}} + n_{\texttt{CpCp}})}. \end{aligned}$$

En comptant les transitions dans le jeu des données, cf. tableau 2.1, l'estimateur du maximum de vraisemblance de la matrice de transition est :

$$\hat{Q} \stackrel{\text{déf}}{=} \begin{pmatrix} 0.7447699 & 0.2426778 & 0.0125523 \\ 0.3385827 & 0.6614173 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{4.12}$$

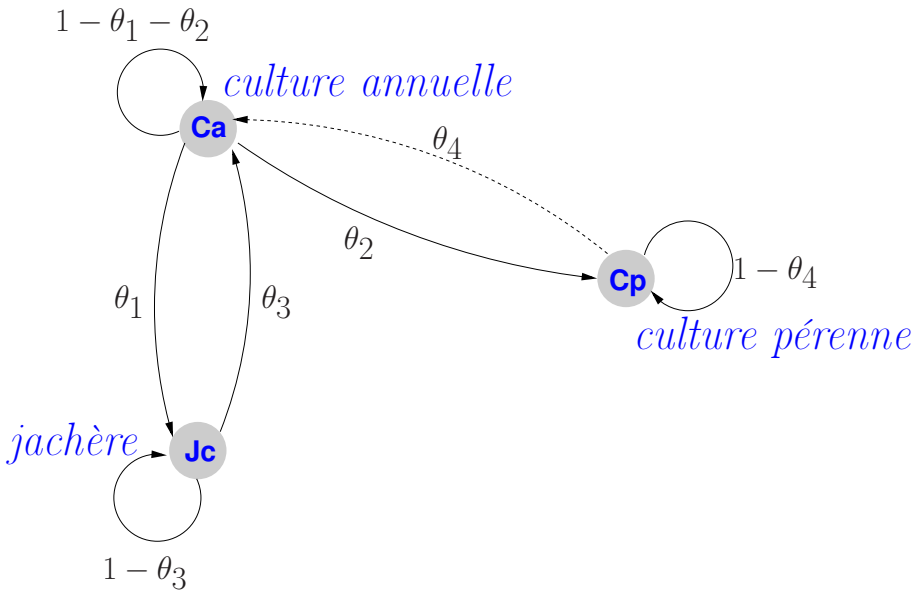

Figure 4.2: *Modèle 2. Chaîne de Markov régulière à trois états : culture annuelle* (`Ca`)*, jachère* (`Jc`)*, culture pérenne* (`Cp`)*.*

4.2.2 Méthode bayésienne

Avec l'estimateur du maximum de vraisemblance (4.12), les transitions `Cp` $\to$ `Ca`, `Jc` $\to$ `Cp` et `Cp` $\to$ `Jc` ne sont pas permises. Selon les spécialistes, les transitions `Jc` $\to$ `Cp` et `Cp` $\to$ `Jc` ne sont pas réalistes dans le contexte géographique de cette étude, tandis que la transition `Cp` $\to$ `Ca` peut être observée sur une échelle de temps de plusieurs décennies.

Dans cette situation, nous supposons que :

(h_3) Les transitions `Jc` $\to$ `Cp` et `Cp` $\to$ `Jc` sont impossibles et que toutes les autres transitions sont possibles.

Cette hypothèse conduit la matrice de transition suivante :

$$Q = \begin{pmatrix} 1-\theta_1-\theta_2 & \theta_1 & \theta_2 \\ \theta_3 & 1-\theta_3 & 0 \\ \theta_4 & 0 & 1-\theta_4 \end{pmatrix}, \tag{4.13}$$

avec $\theta = (\theta_1, \theta_2, \theta_3, \theta_4) \in \Theta = \{\theta \in [0,1]^4 ;\ \theta_1 + \theta_2 \leq 1\}$, qui correspond au modèle décrit dans la Figure 4.2. Cette matrice est régulière à condition que θ appartienne à l'intérieur strict de Θ.

Dans ce cas, l'utilisation de la méthode de maximum de vraisemblance mènera toujours à (4.12).

Estimateur bayésien

Avec l'approche bayésienne, on suppose que le paramètre θ est aléatoire et distribué selon une densité a priori $\pi_{\text{prior}}(\theta)$ qui résume les informations disponibles sur θ. On calcule la loi a posteriori $\pi_{\text{post}}(\theta)$ par la formule de Bayes :

$$\pi_{\text{post}}(\theta) \propto L_2(\theta)\, \pi_{\text{prior}}(\theta) \tag{4.14}$$

où $L_2(\theta)$ est la fonction de vraisemblance définie par

$$\begin{aligned} L_2(\theta) &= \textstyle\prod_{p=1}^{P} \{Q_{11}^{n_{\texttt{CaCa}}^{(p)}}\, Q_{12}^{n_{\texttt{CaJc}}^{(p)}}\, Q_{13}^{n_{\texttt{CaCp}}^{(p)}}\, Q_{21}^{n_{\texttt{JcCa}}^{(p)}}\, Q_{22}^{n_{\texttt{JcJc}}^{(p)}}\, Q_{31}^{n_{\texttt{CpCa}}^{(p)}}\, Q_{33}^{n_{\texttt{CpCp}}^{(p)}}\} \\ &= \textstyle\prod_{p=1}^{P} \{(1-\theta_1-\theta_2)^{n_{\texttt{CaCa}}^{(p)}}\, \theta_1^{n_{\texttt{CaJc}}^{(p)}}\, \theta_2^{n_{\texttt{CaCp}}^{(p)}}\, (1-\theta_3)^{n_{\texttt{JcJc}}^{(p)}}\, \theta_3^{n_{\texttt{JcCa}}^{(p)}}\, (1-\theta_4)^{n_{\texttt{CpCp}}^{(p)}}\, \theta_4^{n_{\texttt{CpCa}}^{(p)}}\} \\ &= (1-\theta_1-\theta_2)^{n_{\texttt{CaCa}}}\, \theta_1^{n_{\texttt{CaJc}}}\, \theta_2^{n_{\texttt{CaCp}}}\, (1-\theta_3)^{n_{\texttt{JcJc}}}\, \theta_3^{n_{\texttt{JcCa}}}\, (1-\theta_4)^{n_{\texttt{CpCp}}}\, \theta_4^{n_{\texttt{CpCa}}}. \end{aligned}$$

L'estimateur bayésien $\tilde{\theta}$ du paramètre θ est la moyenne de la loi a posteriori :

$$\tilde{\theta} \overset{\text{déf}}{=} \int_{\Theta} \theta\, \pi_{\text{post}}(\theta)\, \mathrm{d}\theta = \frac{\int_{\Theta} \theta\, L_2(\theta)\, \pi_{\text{prior}}(\theta)\, \mathrm{d}\theta}{\int_{\Theta} L_2(\theta)\, \pi_{\text{prior}}(\theta)\, \mathrm{d}\theta}\,. \tag{4.15}$$

Loi a priori de Jeffreys

Le choix de la loi a priori reste primordial dans l'utilisation de l'approche bayésienne. Les simulations numériques, présentés à la partie "Simulation : Choix de la loi a priori" de cette section, montrent que parmi différentes lois a priori (loi uniforme, loi beta), celle de Jeffreys est la plus adaptée à notre modèle.

La loi de Jeffreys est une loi a priori non-informative, déterminée à l'aide de la matrice d' information de Fisher [32] par :

$$\pi_{\text{prior}}(\theta) \propto \sqrt{\det[\mathcal{I}(\theta)]}. \tag{4.16}$$

Cette matrice d'information de Fisher $\mathcal{I}(\theta)$ s'obtient par :

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

où $l_2(\theta)$ est la fonction de log-vraisemblance :

$$\begin{aligned} l_2(\theta) = {} & n_{\texttt{CaCa}} \log(1-\theta_1-\theta_2) + n_{\texttt{CaJc}} \log(\theta_1) + n_{\texttt{CaCp}} \log(\theta_2) \\ & + n_{\texttt{JcJc}} \log(1-\theta_3) + n_{\texttt{JcCa}} \log(\theta_3) + n_{\texttt{CpCp}} \log(1-\theta_4) + n_{\texttt{CpCa}} \log(\theta_4). \end{aligned} \tag{4.17}$$

$\mathcal{I}(\theta)$ peut s'écrire :

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

avec

$$A_{1,2} \overset{\text{déf}}{=} -\mathbb{E}_\theta \begin{pmatrix} \frac{\partial^2 l_2(\theta)}{\partial^2 \theta_1} & \frac{\partial^2 l_2(\theta)}{\partial \theta_1 \partial \theta_2} \\ \frac{\partial^2 l_2(\theta)}{\partial \theta_1 \partial \theta_2} & \frac{\partial^2 l_2(\theta)}{\partial^2 \theta_2} \end{pmatrix}, \quad a_k \overset{\text{déf}}{=} -\mathbb{E}_\theta \left(\frac{\partial^2 l_2(\theta)}{\partial^2 \theta_k} \right).$$

D'où $\det \mathcal{I}(\theta) = (\det A_{1,2}) \times a_3 \times a_4$ et

$$\pi_{\text{prior}}(\theta) \propto \sqrt{\det A_{1,2} \times a_3 \times a_4}\,. \tag{4.18a}$$

En utilisant l'expression de la fonction de log-vraisemblance dans (4.17), on a :

$$\begin{aligned}
\frac{\partial^2 l_2(\theta)}{\partial^2 \theta_1} &= -\left(\frac{n_{\texttt{CaCa}}}{(1-\theta_1-\theta_2)^2} + \frac{n_{\texttt{CaJc}}}{\theta_1^2} \right), & \frac{\partial^2 l_2(\theta)}{\partial \theta_1 \, \partial \theta_2} &= -\frac{n_{\texttt{CaCa}}}{(1-\theta_1-\theta_2)^2}, \\
\frac{\partial^2 l_2(\theta)}{\partial^2 \theta_2} &= -\left(\frac{n_{\texttt{CaCa}}}{(1-\theta_1-\theta_2)^2} + \frac{n_{\texttt{CaCp}}}{\theta_2^2} \right), & \frac{\partial^2 l_2(\theta)}{\partial \theta_2 \, \partial \theta_1} &= -\frac{n_{\texttt{CaCa}}}{(1-\theta_1-\theta_2)^2}, \\
\frac{\partial^2 l_2(\theta)}{\partial^2 \theta_3} &= -\left(\frac{n_{\texttt{JcJc}}}{(1-\theta_3)^2} + \frac{n_{\texttt{JcCa}}}{\theta_3^2} \right), & \frac{\partial^2 l_2(\theta)}{\partial^2 \theta_4} &= -\left(\frac{n_{\texttt{CpCp}}}{(1-\theta_4)^2} + \frac{n_{\texttt{CpCa}}}{\theta_4^2} \right)
\end{aligned}$$

et

$$\det A_{1,2} = \left(\frac{\mathbb{E}_\theta[n_{\texttt{CaCa}}]}{(1-\theta_1-\theta_2)^2} + \frac{\mathbb{E}_\theta[n_{\texttt{CaJc}}]}{\theta_1^2} \right) \left(\frac{\mathbb{E}_\theta[n_{\texttt{CaCa}}]}{(1-\theta_1-\theta_2)^2} + \frac{\mathbb{E}_\theta[n_{\texttt{CaCp}}]}{\theta_2^2} \right) - \left(\frac{\mathbb{E}_\theta[n_{\texttt{CaCa}}]}{(1-\theta_1-\theta_2)^2} \right)^2, \tag{4.18b}$$

$$a_3 = \frac{\mathbb{E}_\theta[n_{\texttt{JcJc}}]}{(1-\theta_3)^2} + \frac{\mathbb{E}_\theta[n_{\texttt{JcCa}}]}{\theta_3^2}, \tag{4.18c}$$

$$a_4 = \frac{\mathbb{E}_\theta[n_{\texttt{CpCp}}]}{(1-\theta_4)^2} + \frac{\mathbb{E}_\theta[n_{\texttt{CpCa}}]}{\theta_4^2}. \tag{4.18d}$$

D'après (4.7) et (4.8) :

$$\begin{aligned}
\mathbb{E}_\theta[n_{ee'}] &= \sum_{p=1}^{P} \sum_{n=2}^{N_p} \mathbb{P}_\theta(X_n^{(p)} = e', X_{n-1}^{(p)} = e) \\
&= \sum_{p=1}^{P} \sum_{n=2}^{N_p} \mathbb{P}_\theta(X_n^{(p)} = e' | X_{n-1}^{(p)} = e)\, \mathbb{P}_\theta(X_{n-1} = e) \\
&= Q_{ee'} \sum_{p=1}^{P} \sum_{n=2}^{N_p} \mathbb{P}_\theta(X_{n-1}^{(p)} = e) \\
&= Q_{ee'} \sum_{p=1}^{P} \sum_{n=2}^{N_p} \left(\delta_{\texttt{Ca}} Q^{n-1} \right)_e = Q_{ee'} \sum_{p=1}^{P} \sum_{n=2}^{N_p} [Q^{n-1}]_{\texttt{Ca}e}
\end{aligned} \tag{4.18e}$$

pour tout $e, e' \in E$.

Assoudou et Essebbar [4] proposent une autre méthode plus complexe pour calculer la loi a priori de Jeffreys.

Méthode de Monte Carlo par chaîne de Markov

Bien que le calcul de la loi a priori de Jeffreys soit explicite, le calcul de la loi a posteriori associée (4.15) ne peut pas se faire analytiquement. Nous faisons

appel à une méthode de Monte Carlo par chaînes de Markov (MCMC). L'algorithme utilisé est l'algorithme de Metropolis-Hastings à noyau gaussien décrit par l'algorithme 4.1.

---

choisir θ
$\bar{\theta} \leftarrow \theta$
pour $k = 2, 3, 4, \ldots$ **faire**
  $\varepsilon \sim \mathcal{N}(0, \sigma^2)$
  $\theta^{\text{prop}} \leftarrow \theta + \varepsilon$
  $u \sim U[0, 1]$
  $\alpha \leftarrow \min\left\{1, \dfrac{\pi_{\text{post}}(\theta^{\text{prop}})\, g(\theta - \theta^{\text{prop}})}{\pi_{\text{post}}(\theta)\, g(\theta^{\text{prop}} - \theta)}\right\}$ % g PDF of $\mathcal{N}(0, \sigma^2)$
  si $u \leq \alpha$ **alors**
    $\theta \leftarrow \theta^{\text{prop}}$ % acceptation
  fin si
  $\bar{\theta} \leftarrow \dfrac{k-1}{k}\,\bar{\theta} + \dfrac{1}{k}\,\theta$
fin pour

Alg. 4.1: *Algorithme de Metropolis-Hastings à noyau gaussien : la loi cible est* $\pi_{\text{post}}(\theta)$ *définie par* (4.14)*, et la loi de proposition est* $g(\cdot - \theta)$*, où* g *est la densité associée à la loi normale centrée*$\mathcal{N}(0, \sigma^2)$*.*

La figure 4.3 représente les échantillons obtenus par la méthode MCMC et l'estimateur bayésien $\hat{\theta}$ du paramètre θ. L'estimateur bayésien de Q est alors :

$$\tilde{Q} \stackrel{\text{déf}}{=} \begin{pmatrix} 0.74216087 & 0.2410902 & 0.01674893 \\ 0.32459209 & 0.67540791 & 0 \\ 0.02460644 & 0 & 0.97539356 \end{pmatrix}. \tag{4.19}$$

Simulation : Choix de la loi a priori

Pour avoir une bonne loi a priori, nous avons considéré un modèle simplifié contenant deux états possibles $\{0, 1\}$.

Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov dans $E = \{0, 1\}$ qui a pour matrice de transition :

$$Q \stackrel{\text{déf}}{=} \begin{pmatrix} \theta_1 & 1 - \theta_1 \\ 1 - \theta_2 & \theta_2 \end{pmatrix},$$

où $\theta = (\theta_1, \theta_2) \in [0, 1]^2$.

On suppose que la loi initiale μ de $(X_n)_{n\in\mathbb{N}}$ est égale à sa loi invariante, qui est une solution de l'équation :

$$\mu\, Q = \mu\,. \tag{4.20}$$

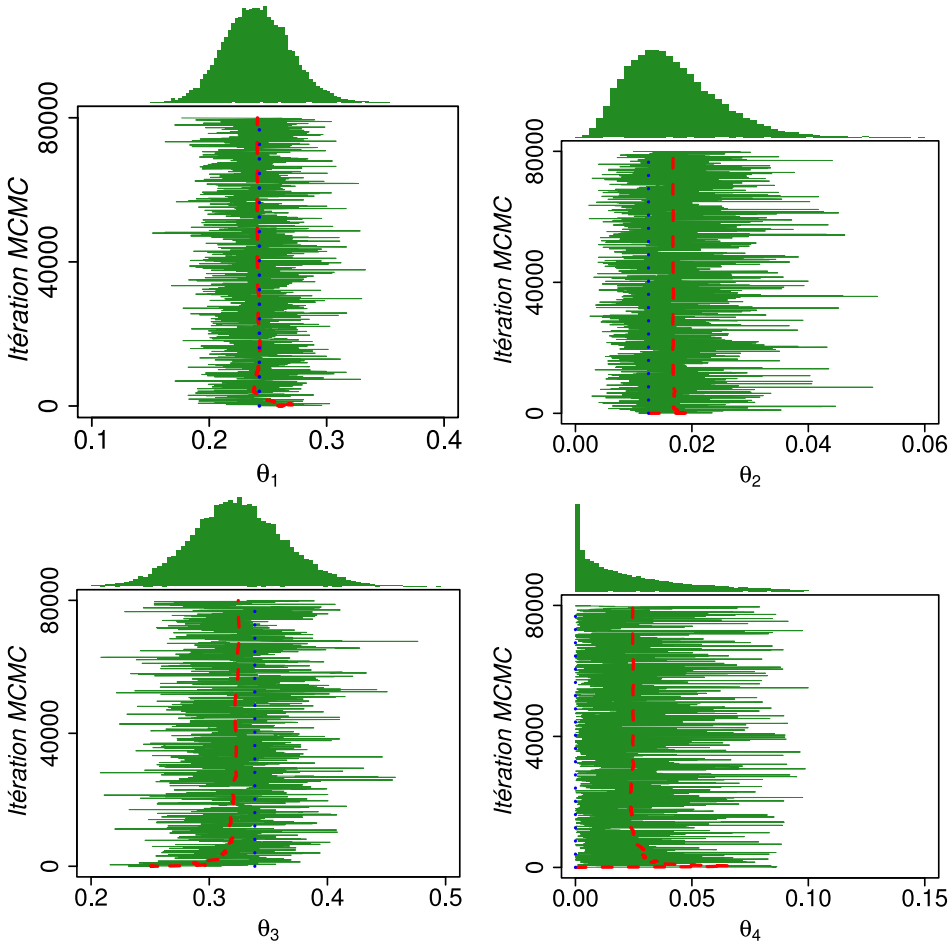

Figure 4.3: *Estimateur bayésien et estimateur du maximum de vraisemblance du paramètre θ de la matrice de transition : échantillon de la loi a posteriori obtenu par la méthode MCMC en "—/vert", la moyenne associée en "- - -/rouge" et l'estimateur du maximum de vraisemblance en "$\cdots$/bleu".*

En résolvant (4.20), on trouve :

$$\mu = \left(\frac{\theta_1}{\theta_1 + \theta_2}, \qquad \frac{\theta_2}{\theta_1 + \theta_2} \right) . \tag{4.21}$$

On souhaite calculer l'estimateur bayésien du paramètre inconnu $\theta = (\theta_1, \theta_2) \in [0,1]^2$ à partir d'un échantillon $(e_0, e_1, \ldots, e_N)$ de $(X_n)_{n \in \mathbb{N}}$.

La fonction de vraisemblance associée au modèle est :

$$L(\theta) = \mathbb{P}_\theta(X_{0:N} = e_{0:N}) = (1-\theta_1)^{n_{00}} \, \theta_1^{n_{01}} \, \theta_2^{n_{10}} \, (1-\theta_2)^{n_{11}},$$

où

$$n_{ij} = \sum_{n=0}^{N-1} 1_{X_n = i} \, 1_{X_{n+1} = j}$$

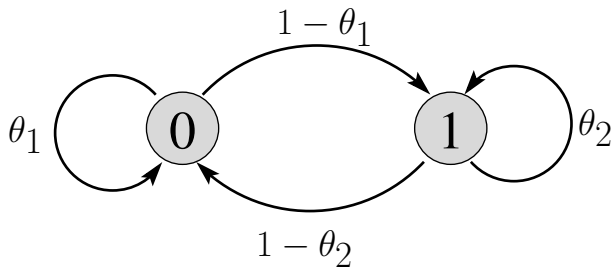

Figure 4.4: Modèle markovien à 2 états. $\theta = (\theta_1, \theta_2) \in [0,1]^2$ est le paramètre de la matrice de transition.

est le nombre de transitions de l'état i vers j.

La fonction de log-vraisemblance est :

$$l(\theta) = n_{00} \log(1-\theta_1) + n_{01} \log \theta_1 + n_{10} \log \theta_2 + n_{11} \log(1-\theta_2).$$

On considère les trois lois a priori suivantes :

(*i*) la loi uniforme sur [0,1] : π^{U},

(*ii*) la loi beta de paramètre (a, b) :

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

où $\beta(a,b)$ est la fonction beta :

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

avec $\Gamma(z) \stackrel{\text{déf}}{=} \int_0^{+\infty} t^{z-1}\, e^{-t}\, \mathrm{d}t$.

En choisissant $a = b = 1/2$, on a $\Gamma(1) = 1$ et $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

(*iii*) la loi non informative de Jeffreys :

$$\pi^{\mathrm{J}}(\theta) \propto \sqrt{\det[\mathcal{I}(\theta)]}$$

avec

$$\mathcal{I}(\theta) \stackrel{\text{déf}}{=} \mathbb{E}_\theta \begin{pmatrix} -\frac{\partial^2 l(\theta)}{\partial^2 \theta_1} & -\frac{\partial^2 l(\theta)}{\partial \theta_1 \partial \theta_2} \\ -\frac{\partial^2 l(\theta)}{\partial \theta_1 \partial \theta_2} & -\frac{\partial^2 l(\theta)}{\partial^2 \theta_2} \end{pmatrix}.$$

En tenant compte de :

$$\begin{aligned} \mathbb{E}_\theta(n_{ij}) &= \sum_{n=1}^{N} \mathbb{P}_\theta(X_{n-1} = i, X_n = j) \\ &= \sum_{n=1}^{N} \mathbb{P}_\theta(X_n = j | X_{n-1} = i)\, \mathbb{P}_\theta(X_{n-1} = i) \end{aligned}$$

c'est-à-dire

$$\begin{aligned}
\mathbb{E}_\theta(n_{00}) &= N\,(1-\theta_1)\,\nu_0 = N\,\frac{(1-\theta_1)\,\theta_2}{\theta_1+\theta_2}\,,\\
\mathbb{E}_\theta(n_{01}) &= N\,\nu_0 = N\frac{\theta_1\,\theta_2}{\theta_1+\theta_2}\,,\\
\mathbb{E}_\theta(n_{10}) &= N\,\theta_2\,\nu_1 = N\,\frac{\theta_1\,\theta_2}{\theta_1+\theta_2},\\
\mathbb{E}_\theta(n_{11}) &= N\,(1-\theta_2)\,\nu_1 = N\,\frac{(1-\theta_1)\,\theta_2}{\theta_1+\theta_2},
\end{aligned}$$

on a :

$$\begin{aligned}
\pi_J(\theta) &\propto \sqrt{\frac{\mathbb{E}_\theta(n_{01})}{\theta_1^2}+\frac{\mathbb{E}_\theta(n_{00})}{(1-\theta_1)^2}} \times \sqrt{\frac{\mathbb{E}_\theta(n_{10})}{\theta_2^2}+\frac{\mathbb{E}_\theta(n_{11})}{(1-\theta_2)^2}}\\
&= \frac{N}{(\theta_1+\theta_2)\,\sqrt{1-\theta_1}\,\sqrt{1-\theta_2}}\,.
\end{aligned}$$

Finalement, on peut écrire :

$$\pi^{\mathrm{J}}(\theta) \propto \frac{1}{(\theta_1+\theta_2)\sqrt{1-\theta_1}\sqrt{1-\theta_2}}\,.$$

Les lois a posteriori correspondantes aux trois lois a priori sont :

$$\begin{aligned}
\tilde{\pi}^{\mathrm{J}}(\theta) &\propto L(\theta)\pi^{\mathrm{J}}(\theta) = \frac{(1-\theta_1)^{n_{00}-\frac{1}{2}}\theta_1^{n_{01}}\,\theta_2^{n_{10}}\,(1-\theta_2)^{n_{11}-\frac{1}{2}}}{\theta_1+\theta_2}\,,\\
\tilde{\pi}^{\mathrm{U}}(\theta) &\propto L(\theta) = (1-\theta_1)^{n_{00}}\,\theta_1^{n_{01}}\,\theta_2^{n_{10}}\,(1-\theta_2)^{n_{11}}\,,\\
\tilde{\pi}^{\mathrm{B}}(\theta) &\propto L(\theta)\,\pi^{\mathrm{B}}(\theta) = (1-\theta_1)^{n_{00}-\frac{1}{2}}\,\theta_1^{n_{01}-\frac{1}{2}}\,\theta_2^{n_{10}-\frac{1}{2}}\,(1-\theta_2)^{n_{11}-\frac{1}{2}},
\end{aligned}$$

et les estimateurs bayésiens sont :

$$\tilde{\theta}^{\mathrm{U}} = \int_{[0,1]^2} \theta\;\tilde{\pi}^{\mathrm{U}}(\theta)d\theta, \tag{4.22}$$

$$\tilde{\theta}^{\mathrm{B}} = \int_{[0,1]^2} \theta\;\tilde{\pi}^{\mathrm{B}}(\theta)d\theta, \tag{4.23}$$

$$\tilde{\theta}^{\mathrm{J}} = \int_{[0,1]^2} \theta\;\tilde{\pi}^{\mathrm{J}}(\theta)d\theta. \tag{4.24}$$

Pour (4.22) et (4.23), les calculs sont explicites avec :

$$\tilde{\theta_1}^{\mathrm{U}} = \frac{n_{01}+1}{n_{01}+n_{00}+2}, \qquad \tilde{\theta_2}^{\mathrm{U}} = \frac{n_{10}+1}{n_{10}+n_{11}+2}$$

et

$$\tilde{\theta_1}^{\mathrm{B}} = \frac{n_{01}+\frac{1}{2}}{n_{01}+n_{00}+1} \qquad \tilde{\theta_2}^{\mathrm{B}} = \frac{n_{10}+\frac{1}{2}}{n_{10}+n_{11}+1}\,.$$

Le calcul de (4.22) se fait à l'aide de la méthode monte carlo par chaîne de Markov suivant l'algorithme 4.1.

Nous comparons les trois estimateurs bayésiens et l'estimateur du maximum de vraisemblance $\hat{\theta}$, où :

$$\hat{\theta}_1 = \frac{n_{00}}{n_{01} + n_{01}}, \qquad \hat{\theta}_2 = \frac{n_{11}}{n_{11} + n_{10}}$$

par une méthode de Monte carlo. Plus précisément, nous échantillonnons 1000 valeurs $\theta^{(\ell)}$ de θ suivant une loi uniforme $\Theta = [0,1]^2$, $\ell = 1, \ldots, 1000$, et nous posons :

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

Pour tout $\ell = 1, \ldots, 1000$, nous simulons des observations $(e_{0:19}^{(p)})_{p=1:42}$ avec Q_{θ^ℓ}. Puis nous calculons l'estimateur du maximum de vraisemblance $\hat{\theta}^{(\ell)}$ et les estimateurs bayésiens $\tilde{\theta}_{\mathrm{U}}^{(\ell)}$, $\tilde{\theta}_{\mathrm{B}}^{(\ell)}$, $\tilde{\theta}_{\mathrm{J}}^{(\ell)}$ associés respectivement aux trois lois a priori : uniforme, beta et Jeffreys.

La comparaison consiste à calculer la loi empirique des erreurs d'estimation pour chaque estimateur qui est une distance entre la matrice de transition réelle et son estimation, avec :

$$\hat{\varepsilon}^\ell = \|Q_{\theta^{(\ell)}} - Q_{\hat{\theta}^{(\ell)}}\|, \tag{4.25}$$

$$\tilde{\varepsilon}_{\mathrm{U}}^\ell = \|Q_{\theta_{\mathrm{U}}^{(\ell)}} - Q_{\tilde{\theta}_{\mathrm{U}}^{(\ell)}}\|, \tag{4.26}$$

$$\tilde{\varepsilon}_{\mathrm{B}}^\ell = \|Q_{\theta_{\mathrm{B}}^{(\ell)}} - Q_{\tilde{\theta}_{\mathrm{B}}^{(\ell)}}\|, \tag{4.27}$$

$$\tilde{\varepsilon}_{\mathrm{J}}^\ell = \|Q_{\theta_{\mathrm{J}}^{(\ell)}} - Q_{\tilde{\theta}_{\mathrm{J}}^{(\ell)}}\|. \tag{4.28}$$

Deux normes matricielles ont été utilisées, la norme de Frobenius :

$$\|A\|_F^2 \stackrel{\text{déf}}{=} \mathrm{trace}(A^* A) \tag{4.29}$$

et la norme-2 :

$$\|A\|_2 \stackrel{\text{déf}}{=} \sqrt{\lambda_{\max}(A^* A)} \tag{4.30}$$

où $\lambda_{\max}(A^* A)$ est la plus grande valeur propre de la matrice $A^* A$.

Dans la figure 4.5, nous avons tracé les lois empiriques des erreurs d'estimation $\hat{\varepsilon}^\ell$ et $\tilde{\varepsilon}^\ell$, $\ell = 1, \ldots, 1000$, pour les deux différentes normes. Nous constatons que l'estimateur bayésien associé à la loi a priori de Jeffreys donne des résultats légèrement meilleurs pour le cas correspondant à la norme-2 et de meilleurs résultats pour la norme de Frobenius. Par conséquent, on a choisi d'utiliser la loi a priori de Jeffreys dans notre étude.

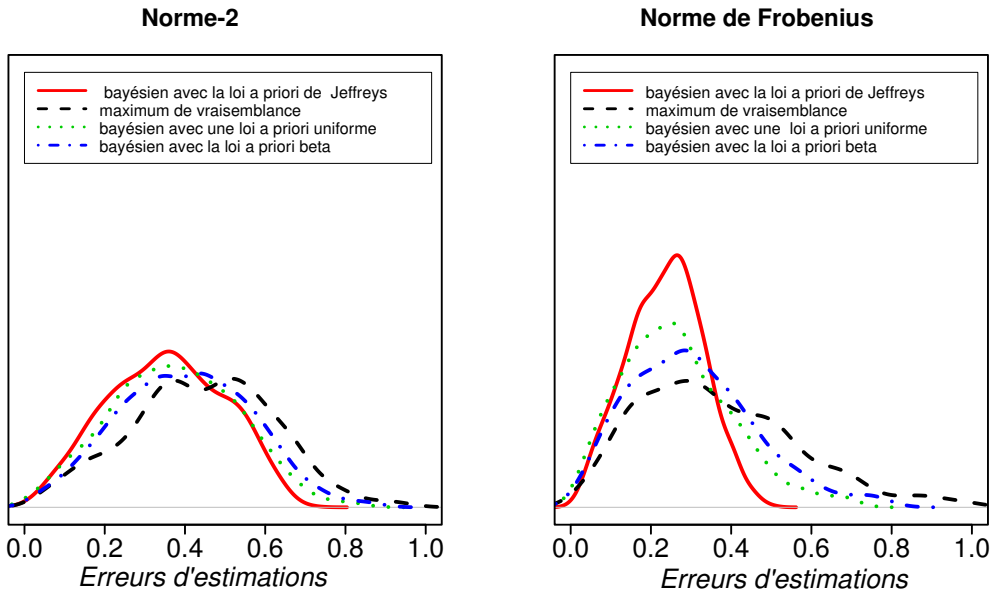

Figure 4.5: *Lois empiriques des erreurs d'estimations* (4.25) *associées à la norme-2* (4.30) *et la norme de Frobenius* (4.29) *basées sur 1000 simulations du paramètre* θ.

4.3 Résultats et discussions

4.3.1 Modèle 1 : chaîne de Markov absorbante déduite de la méthode de vraisemblance

L' estimateur du maximum de vraisemblance (EMV) (4.12) de la matrice transition Q correspond à une chaîne de Markov absorbante, cf. figure 4.6. En effet quand le processus atteint l'état $\mathtt{Cp}$, il y reste pour toujours. La loi limite est donc $\delta_{\mathtt{Cp}}$.

Une première question est de pouvoir décrire le comportement du processus avant qu'il n'atteigne cet état $\mathtt{Cp}$, ensuite analyser la durée de temps pour l'atteindre.

La réponse à la première question est donnée par la loi quasi-stationnaire de la chaîne de Markov absorbante associée à $\hat{Q}$ et la deuxième revient à étudier la loi du temps d'absorption, cf. proposition 3.3.4.

Loi du temps d'absorption

Notons $\tau_{\mathtt{Cp}}$ la variable aléatoire désignant le temps d'atteinte de l'état absorbant $\mathtt{Cp}$, avec :

$$\tau_{\mathtt{Cp}} \overset{\text{déf}}{=} \inf\{n \geq 1 \ : \ X_n = \mathtt{Cp}\} \ \text{ si } \exists\, n \geq 1 \text{ tel que } X_n = \mathtt{Cp}, \ +\infty \text{ sinon}. \tag{4.31}$$

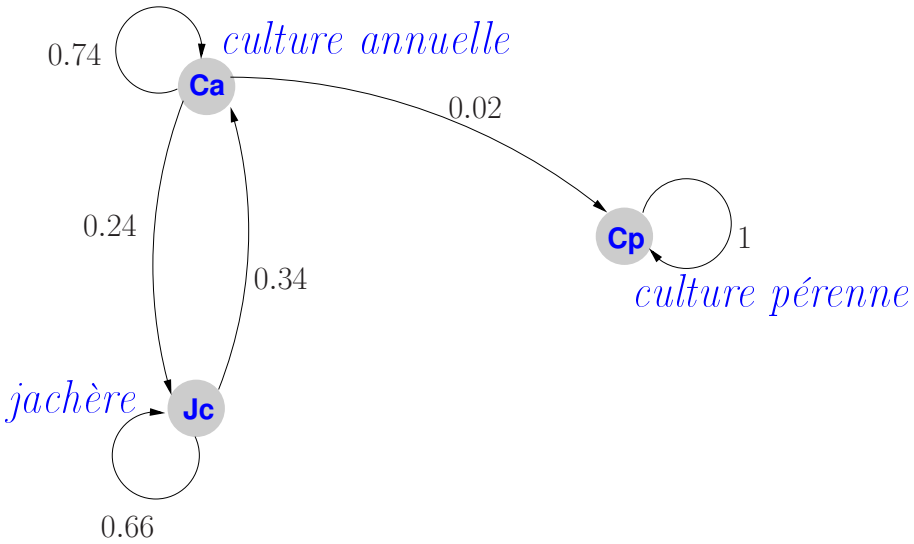

Figure 4.6: *Modèle 1 : Chaîne de Markov absorbante, déduite directement du jeu des données à l'aide de l'estimateur du maximum de vraisemblance.*

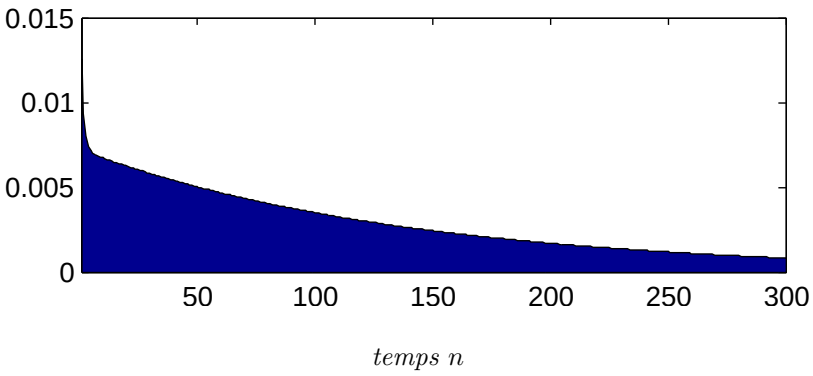

Figure 4.7: *Loi du temps d'atteinte de l'état absorbant* Cp.

Nous nous intéressons à la probabilité :

$$\begin{aligned} f_{\mathtt{CaCp}}(n) &\overset{\text{déf}}{=} \mathbb{P}(\tau_{\mathtt{Cp}} = n | X_0 = \mathtt{Ca}) \\ &= \mathbb{P}(X_n = \mathtt{Cp},\ X_m \neq \mathtt{Cp},\ m = 1, \ldots, n-1 | X_0 = \mathtt{Ca}). \end{aligned}$$

Notons que pour $n > 1$:

$$\begin{aligned}
\hat{Q}^{(n_{\mathtt{CaCp}})} &= \mathbb{P}(X_n = \mathtt{Cp}|X_0 = \mathtt{Ca}),\\
&= \mathbb{P}(X_n = \mathtt{Cp},\, \tau_{\mathtt{Cp}} = 1|X_0 = \mathtt{Ca}) + \mathbb{P}(X_n = \mathtt{Cp},\, \tau_{\mathtt{Cp}} = 2|X_0 = \mathtt{Ca}) + \cdots\\
&\qquad \cdots + \mathbb{P}(X_n = \mathtt{Cp},\, \tau_{\mathtt{Cp}} = n-1|X_0 = \mathtt{Ca})\\
&\qquad\qquad + \mathbb{P}(X_n = \mathtt{Cp},\, \tau_{\mathtt{Cp}} = n|X_0 = \mathtt{Ca})\\
&= \mathbb{P}(\tau_{\mathtt{Cp}} = 1|X_0 = \mathtt{Ca})\,\mathbb{P}(X_n = \mathtt{Cp}|X_1 = \mathtt{Ca})\\
&\qquad + \mathbb{P}(\tau_{\mathtt{Cp}} = 2|X_0 = \mathtt{Ca})\,\mathbb{P}(X_n = \mathtt{Cp}|X_2 = \mathtt{Cp})\\
&\qquad + \mathbb{P}(\tau_{\mathtt{Cp}} = n-1|X_0 = \mathtt{Ca})\,\mathbb{P}(X_n = \mathtt{Cp}|X_{n-1} = \mathtt{Cp}) + \cdots\\
&\qquad\qquad \cdots + \mathbb{P}(\tau_{\mathtt{Cp}} = n|X_0 = \mathtt{Ca})\\
&= f_{\mathtt{CaCp}}(1)\,\hat{Q}^{(n-1)}_{\mathtt{CpCp}} + f_{\mathtt{CaCp}}(2)\,\hat{Q}^{(n-2)}_{\mathtt{CpCp}} + \cdots + f_{\mathtt{CaCp}}(n-1)\,\hat{Q}^{(1)}_{\mathtt{CpCp}} + f_{\mathtt{CaCp}}(n).
\end{aligned}$$

Ainsi $f_{\mathtt{CaCp}}(n)$ peut être calculé par récurrence :

$$f_{\mathtt{CaCp}}(n) = \hat{Q}^{(n)}_{\mathtt{CaCp}} - \sum_{k=1}^{(n-1)} f_{\mathtt{CaCp}}(k)\,\hat{Q}^{(n-k)}_{\mathtt{CpCp}} \tag{4.32}$$

avec $f_{\mathtt{CaCp}}(1) = \hat{Q}_{\mathtt{CaCp}}$.

Le calcul de la moyenne peut se faire à l'aide de l'équation (3.20) ou bien directement par :

$$\mathbb{E}(\tau_{\mathtt{Cp}}|X_0 = \mathtt{Ca}) = \sum_{n=1}^{\infty} n\, f_{\mathtt{CaCp}}(n).$$

D'après notre calcul, la moyenne du temps d'atteinte de l'état `Cp` est de 136 ans et la moyenne et son écart-type est de 135 ans.

Loi quasi-stationnaire

La loi quasi-stationnaire est la loi "limite" du système avant d'être absorbé par `Cp`. Considérons la probabilité pour que le système soit dans l'un des deux états non absorbants $e \in \{\mathtt{Ca}, \mathtt{Jc}\}$ avant d'atteindre l'état `Cp`, c'est-à-dire :

$$\mu_n(e) = \mathbb{P}(X_n = e|\tau_{\mathtt{Cp}} > n,\, X_0 = \mathtt{Ca}).$$

Si

$$\mu_n(e) \xrightarrow[n\to\infty]{} \sigma_{\mathrm{qs}}(e)\,,\ e \in \{\mathtt{Ca}, \mathtt{Jc}\},$$

alors, $(\sigma_{\mathrm{qs}}(e))_{e\in\{\mathtt{Ca},\mathtt{Jc}\}}$ est appelée loi quasi-stationnaire de la chaîne absorbante $(X_n)_{n\in\mathbb{N}}$.

La loi quasi-stationnaire $\sigma_{\mathrm{qs}} = [\sigma_{\mathrm{qs}}(\mathtt{Ca}), \sigma_{\mathrm{qs}}(\mathtt{Jc})]$ est la solution de :

$$\sigma_{\mathrm{qs}}\,\hat{Q}_{\mathrm{qs}} = \lambda\,\sigma_{\mathrm{qs}}$$

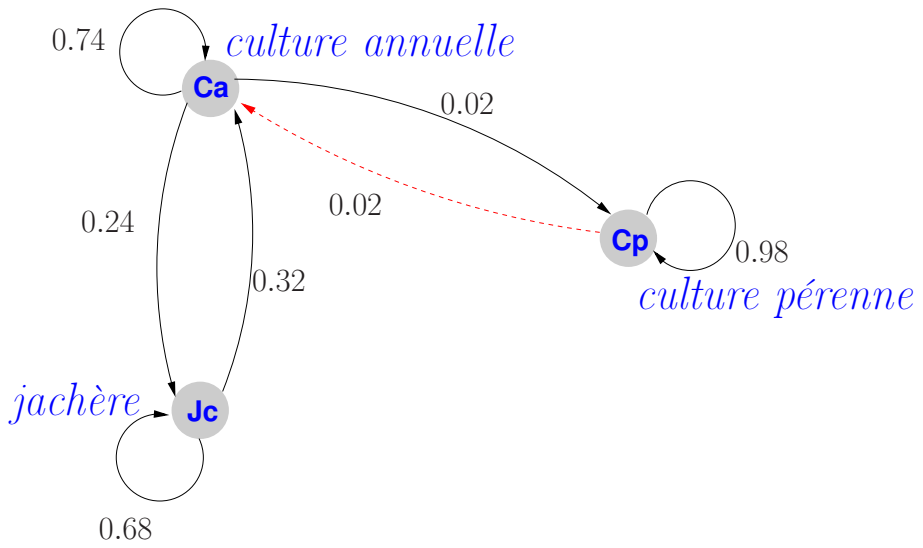

Figure 4.8: *Modèle 2 : Chaîne de Markov régulière, déduite par l'estimateur bayésien de la matrice de transition.*

avec $\sigma_{\text{qs}}(e) \geq 0$, $\sigma_{\text{qs}}(\texttt{Ca}) + \sigma_{\text{qs}}(\texttt{Jc}) = 1$, $\hat{Q}_{\text{qs}}$ est la sous matrice 2×2 définie par :

$$\hat{Q} = \left(\begin{array}{c|c} \hat{Q}_{\text{qs}} & R \\ \hline 0 & 1 \end{array}\right)$$

et λ est le rayon spectrale $\hat{Q}_{\text{qs}}$.

On trouve :

$$\sigma_{\text{qs}} = (\sigma_{\text{qs}}(\texttt{Ca}), \sigma_{\text{qs}}(\texttt{Jc})) = (0.5794, 0.4206)\,. \tag{4.33}$$

Ceci veut dire que, conditionnellement au fait que le processus n'atteint pas `Cp`, il va passer 58% de son temps dans l'état `Ca` et 42% dans l'état `Jc`.

4.3.2 Modèle 2 : chaîne de Markov régulière déduite de l'approche bayésienne

On vérifie facilement que l'estimateur bayésien $\tilde{Q}$ de la matrice de transition Q défini dans (4.19) est irréductible et apériodique, donc régulière. D'après le théorème 3.2.6, elle admet une loi invariante.

Loi invariante

La loi invariante de la chaîne de Markov associée à $\tilde{Q}$ est une solution de l'équation

$$\sigma = \sigma\,\tilde{Q}. \tag{4.34}$$

Selon le théorème de Perron–Frobenius, σ existe et unique. C'est le vecteur propre associé à la valeur propre 1 qui est la plus grande valeurs propre de $\tilde{Q}$. On trouve :

$$\sigma = (0.4127\,,\, 0.3065\,,\, 0.2808)\,. \tag{4.35}$$

C'est-à-dire à l'équilibre, 41.27% des parcelles sont en "culture annuelle", 30.65% en "jachère" et 28.08% en "culture pérenne".

Vitesse de convergence

Dans cette section nous allons étudier la vitesse de convergence du processus vers sa loi invariante [23]. Cette vitesse se traduit par un taux qui représente la proportion des parcelles qui atteignent l'équilibre par an. Nous allons considérer deux indications sur cette vitesse : le coefficient d'amortissement [13, p. 95] et le coefficient de Dobrushin [16].

Coefficient d'amortissement

Le rapport entre les deux plus grandes valeurs propres :

$$\alpha = \lambda_1/|\lambda_2|, \tag{4.36}$$

appelé *coefficient d'amortissement*, donne une première indication sur la vitesse de convergence vers la loi stationnaire.

Proposition 4.3.1. *Le processus converge exponentiellement vers l'équilibre au moins à la vitesse* $\log\alpha$ *par an.*

Démonstration. La matrice $\tilde{Q}$ est régulière et ses trois vecteurs propres w_1, w_2 et w_3 sont linéairement indépendants. La loi initiale μ peut être écrite sous forme d'une combinaison linéaire de ces vecteurs :

$$\mathrm{loi}(X_0) = \mu = c_1 w_1 + c_2 w_2 + c_3 w_3,$$

où c_1, c_2, c_3 sont des coefficients réels. La loi de X_1 peut s'écrire comme :

$$\begin{aligned}\mathrm{loi}(X_1) &= \mu\,\tilde{Q}\\ &= c_1 w_1 \tilde{Q} + c_2 w_2 \tilde{Q} + c_3 w_3 \tilde{Q}\\ &= c_1 \lambda_1 w_1 + c_2 \lambda_2 w_2 + c_3 \lambda_3 w_3\end{aligned}$$

avec λ_1, λ_2, λ_3 valeurs propres de $\tilde{Q}$.

De même pour X_2 :

$$\begin{aligned}\mathrm{loi}(X_2) &= \mathrm{loi}(X_1)\,\tilde{Q}\\ &= c_1\lambda_1 w_1\tilde{Q} + c_2\lambda_2 w_2\tilde{Q} + c_3\lambda_3 w_3\tilde{Q}\\ &= c_1\lambda_1^2 w_1 + c_2\lambda_2^2 w_2 + c_3\lambda_3^2 w_3.\end{aligned}$$

Pour tout $n \in \mathbb{N}$, nous pouvons écrire :

$$\mathrm{loi}(X_n) = \mu\,\tilde{Q}^{(n)} = c_1\lambda_1^n w_1 + c_2\lambda_2^n w_2 + c_3\lambda_3^n w_3. \tag{4.37}$$

Divisons les deux côtés par λ_1^n :

$$\frac{\tilde{Q}^{(n)}}{\lambda_1^n} = c_1 w_1 + c_2(\frac{\lambda_2}{\lambda_1})^n w_2 + c_3(\frac{\lambda_3}{\lambda_1})^n w_3.$$

En considérant le fait que $1 = \lambda_1 > |\lambda_2| > |\lambda_3|$, et $w_1 = \sigma$ on a

$$\lim_{n\to\infty}\left(\frac{\tilde{Q}^{(n)}}{\lambda_1^n} - c_1 w_1 - c_2\alpha^{-t} w_2\right) = 0.$$

Par conséquent

$$\|\tilde{Q}^{(n)} - c_1\sigma\| \leq k\,\alpha^{-n} = k\,e^{-n\log\alpha}, \tag{4.38}$$

où k est une constante. □

D'après la proposition 4.3.1, les proportions d'usages des parcelles convergent asymptotiquement vers ses équilibres au moins à la vitesse $\log\alpha = 0.036$, c'est-à-dire 3.6% par an.

Coefficient d'ergodicité de Dobrushin

Le coefficient d'ergodicité de Dobrushin [16] pour la matrice $\tilde{Q}$ est défini par :

$$\mathcal{D}(\tilde{Q}) \overset{\text{déf}}{=} \tfrac{1}{2}\max_{i,k\in E}\sum_{j\in\mathbb{E}}|\tilde{Q}_{ij} - \tilde{Q}_{kj}|. \tag{4.39}$$

En partant des deux lois initiales différentes μ^1 et μ^2, les proportions après une itération en temps sont $\mu_1\tilde{Q}$ et $\mu_2\tilde{Q}$ et le taux de contraction est :

$$\begin{aligned}\frac{\|\mu^1\tilde{Q} - \mu^2\tilde{Q}\|_1}{\|\mu^1 - \mu^2\|_1} &\leq \sup_{\mu^1,\mu^2}\frac{\|\mu^1\tilde{Q} - \mu^2\tilde{Q}\|_1}{\|\mu^1 - \mu^2\|_1}\\ &= \tfrac{1}{2}\max_{i,k\in E}\sum_{j\in\mathbb{E}}|\tilde{Q}_{ij} - \tilde{Q}_{kj}| = \mathcal{D}(\tilde{Q})\end{aligned}$$

où $\|\mu\|_1 = \sum_{i\in\mathbb{E}}\mu_i$ est la norme-1.

Ce taux de contraction $-\log\mathcal{D}(\tilde{Q})$ donne aussi une borne inférieure de la vitesse de convergence des proportions d'usages des parcelles vers ses équilibres. Dans notre cas $\mathcal{D}(\tilde{Q}) = 0.975$ et $-\log(\mathcal{D}(\tilde{Q})) = 0.025$, c'est-à-dire 2.5% par an. Contrairement à la proposition 4.3.1, cette borne inférieure ne se limite pas à des conditions asymptotiques.

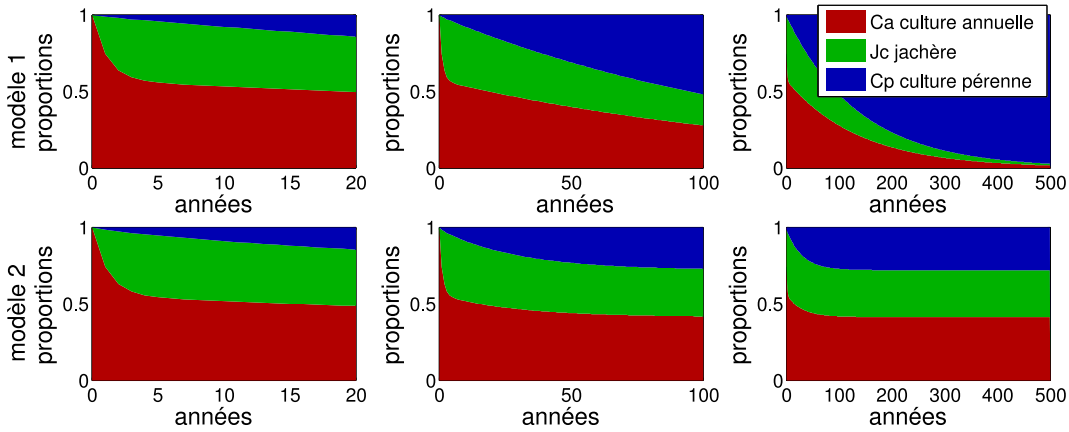

Figure 4.9: *Évolution des proportions des parcelles occupées par les trois états* `Ca`*,* `Jc` *et* `Cp` *pendant 20, 100 et 500 ans selon les deux modèles.*

4.3.3 Discussions

L'application d'une approche markovienne au jeu des données conduit à deux modèles. Le premier modèle contient un état absorbant `Cp` et la loi limite du processus est $1_{\texttt{Cp}}$. Le deuxième modèle est une chaîne de Markov régulière admettant une loi invariante σ. La figure 4.9 montre l'évolution des proportions des parcelles occupées par les trois états `Ca`, `Jc` et `Cp`, en partant de l'état initial `Ca`.

Pour les deux modèles, les évolutions sont presque identiques dans un horizon de 20 ans. Des différences sont perceptibles après quelques décennies. Les deux modèles sont radicalement différents à une longue échelle, après quelques siècles) : Pour le modèle 1, le processus est absorbé par l'état `Cp` alors qu'une équilibre s'établit pour le modèle 2.

En général, les deux modèles convergent vers l'équilibre des proportions sur les états autres que `Cp` au bout de 100 ans. Pour le premier modèle, le temps moyen pour atteindre l'état absorbant `Cp` est de 136 ans. Dans le second modèle, l'équilibre est atteint après un siècle et le processus passe 41% de sont temps à l'état `Ca`, 31% à l'état `Jc` et 28% à l'état `Cp`.

Le modèle 1 admet une loi quasi-stationnaire. Si on considère l'évolution des proportions relatives des parcelles dans l'état `Ca` et dans l'état `Jc`, les deux modèles présentent des comportements presque identiques, voir figure 4.10.

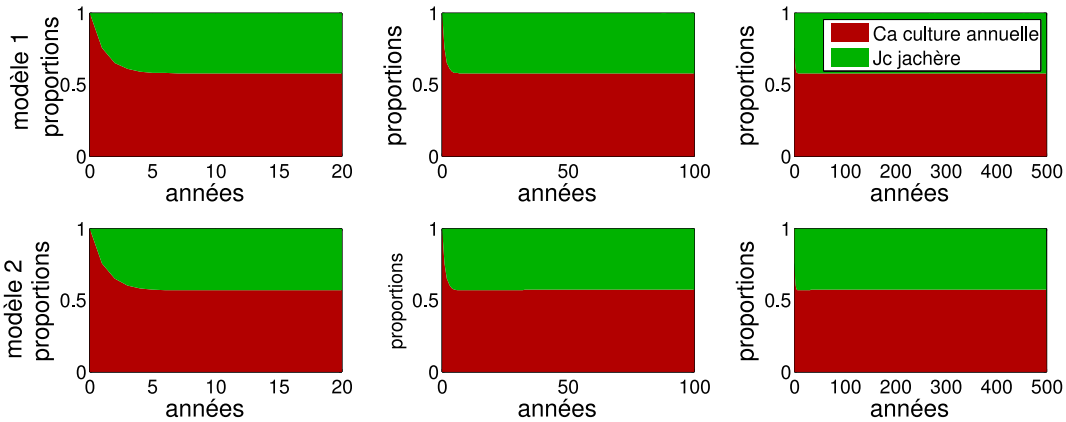

Figure 4.10: *Évolution des proportions des parcelles occupées par les états* `Ca` `Jc` *pendant 20, 100 et 500 années selon les deux modèles. Après 5 ans, les proportions des parcelles aux états culture annuelle* (`Ca`) *et jachère* (`Jc`) *approchent l'équilibre, c'est-à-dire* 58% *et* 48%.

Chapitre 5

Evaluation des modèles markoviens

Dans ce chapitre, nous allons effectuer quelques tests statistiques qui permettent d'évaluer l'adéquation du modèle markovien au jeu de données. Le premier test concerne les lois des temps de séjour sur les états et le deuxième l'homogénéité de la dynamique.

5.1 Test d'adéquation des lois de temps de séjour

5.1.1 Lois des temps de séjour

Considérons une chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ de matrice de transition Q sur un espace d'états E fini. Pour un état $i \in E$, le temps de séjour $S(i)$ est le temps que la chaîne passe en i avant de passer à un autre état :

$$S(i) \stackrel{\text{déf}}{=} \inf\{n \in \mathbb{N};\ X_n \neq i | X_0 = i\} \tag{5.1}$$

qui a pour loi :

$$\begin{aligned}
\mathbb{P}(S(i) = n | X_0 = i) &= \mathbb{P}(X_1 = \cdots = X_n = i, X_{n+1} \neq i | X_0 = i) \\
&= \sum_{j\neq i} \mathbb{P}(X_1 = \cdots = X_n = i, X_{n+1} = j | X_0 = i) \\
&= \sum_{j\neq i} \underbrace{Q_{ii} \cdots Q_{ii}}_{n-1 \text{ fois}} Q_{ij} = (Q_{ii})^{n-1} (1 - Q_{ii})
\end{aligned}$$

pour $n \geq 1$ et 0 pour $n = 0$. La loi de probabilité de $S(i)$ est donc la loi géométrique de paramètre :

$$1 - Q_{ii} = 1 - \mathbb{P}(X_{n+1} = i | X_n = i)\,,$$

son espérance est :

$$\mathbb{E}(S(i)|X_0 = i) = 1/(1 - Q_{ii})$$

et sa variance :

$$\mathrm{var}(S(i)|X_0 = i) = Q_{ii}/(1 - Q_{ii})^2 \,.$$

Le test d'adéquation entre le modèle et le jeu de données consiste à recenser les temps de séjour en $i \in E$ parmi les observations $(e^{(p)}_{1:N_p})_{p=1:42}$ et de tester si leurs lois correspondent à une loi géométrique ou non. La technique que nous avons choisi pour mener ce test est le bootstrap paramétrique [22].

5.1.2 Test d'adéquation par bootstrap paramétrique

Soit un k-échantillon $S_1, \ldots, S_k$ d'une fonction de répartition $(F(n))_{n\in\mathbb{N}}$. Nous souhaitons tester si la loi de cet échantillon appartient à une famille des fonctions de répartition indexées par un paramètre $p \in \Theta$:

$$H_0 \,:\, F \in \mathbb{F} \stackrel{\text{déf}}{=} \{F_p : p \in \Theta\} \,. \tag{5.2}$$

Dans notre cas, F_p est la fonction de répartition d'une loi géométrique de paramètre $p \in [0\ 1]$.

On se donne un estimateur :

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

du paramètre p et on calcule une distance entre la fonction de répartition théorique $F_{\hat{p}}$ et la fonction de répartition empirique :

$$\hat{F}_{S_{1:k}}(n) \stackrel{\text{déf}}{=} \frac{1}{k}\sum_{\ell=1}^{k} 1_{\{S_\ell \le n\}} \,. \tag{5.3}$$

L'estimateur du maximum de vraisemblance du paramètre p est donné par :

$$\hat{p} \stackrel{\text{déf}}{=} \frac{1}{1 + \frac{1}{k}\sum_{\ell=1}^{k} S_\ell} \,. \tag{5.4}$$

En effet, la fonction de vraisemblance associée à l'échantillon s'écrit :

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

et $L'(p) = 0$ conduit à $k - p\,(\sum_{\ell=1}^{k} S_\ell + k) = 0$ et (5.4).

Nous considérons la distance de Kolmogorov-Smirnov :

$$K^* = K(S_{1:k}) \stackrel{\text{déf}}{=} \sup_{n\in\mathbb{N}} \sqrt{k}\,|\hat{F}_{S_{1:k}}(n) - F_{T(S_{1:k})}(n)| \tag{5.5}$$

de $\hat{F}_{S_{1:k}}$ à $F_{S_{1:k}}$.

La question est donc de savoir si la distance K^* est significativement différente de 0 ou non. Pour cela, on simule M échantillons indépendants de $F_{\hat{p}}$ de taille k :

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

et on calcule :

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

où K est la distance calculée dans (5.5).

La p-valeur associée au test est :

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

On se donne une probabilité α de rejeter à tort H_0. Si $\rho \leq \alpha$, alors on rejette H_0 sinon on l'accepte. Cette procédure est décrite dans l'algorithme 5.1.

$\hat{p} \leftarrow T(S_{1:k})$
$\bar{n} \leftarrow \sup(S_1, \ldots, S_k)$
$K^* \leftarrow \sup\{\sqrt{k}|\hat{F}_{S_{1:k}}(n) - F_{T(S_{1:k})}(n)|\,,\ 0 \leq n \leq \bar{n}\}$
pour $m = 1, 2, \ldots, M$ **faire**
  $S_1^m, \ldots, S_k^m \overset{\text{iid}}{\sim} F_{\hat{p}}$
  $\bar{n} \leftarrow \sup(S_1^m, \ldots, S_k^m)$
  $K^m \leftarrow \sup\{\sqrt{k}|\hat{F}_{S_{1:k}^m}(n) - F^{T(S_{1:k}^m)}(n)|\,,\ 0 \leq n \leq \bar{n}\}$
  $\rho \leftarrow \frac{1}{M}\sum_{m=1}^{M} 1_{\{K^m \geq K^*\}}$
fin pour
si $\rho \leq \alpha$ **alors**
  acceptation de H_0
sinon
  rejet de H_0
fin si

ALG. 5.1: *Test d'adéquation d'un échantillon $S_1, \ldots, S_k$ à une loi géométrique par la méthode de bootstrap paramétrique. La distance utilisée est la distance de Kolmogorov-Smirnov calculée à l'aide de l'équation* (5.5).

5.1.3 Résultats

Dans le comptage du temps de séjour, les états qui apparaissent à la fin des observations ne seront pas traités. Ils sont considérés comme des données censurées. Par conséquent, le temps de séjour sur l'état `Cp` (culture pérenne) n'est pas considéré dans ce test. Les temps de séjour observés sur les autres états sont donnés dans le tableau 5.1.

`Ca` (culture annuelle)	valeurs du temps de séjour	1	2	3	4	5	6	
	nombre d'occurrences	11	17	12	5	9	7	
`Jc` (Jachère)	valeurs du temps de séjour	1	2	3	4	6	8	11
	nombre d'occurrences	16	12	7	4	2	1	1

Tableau 5.1: *Valeurs des temps de séjour (par an) observées sur l'état* `Ca` *et* `Jc`.

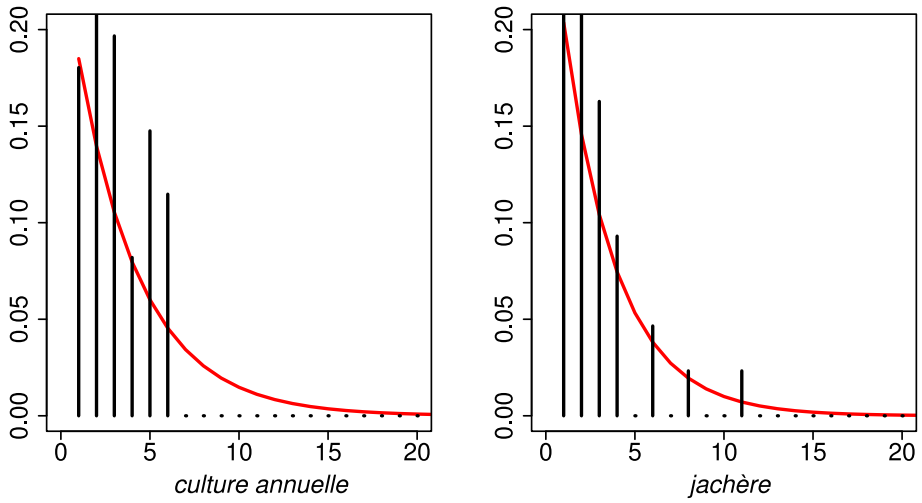

Figure 5.1: *Densités empiriques des temps de séjour sur les deux états culture annuelle (*`Ca`*) et jachère (*`Jc`*) calculées avec le jeu des données du tableau 2.1. La ligne continue rouge correspond à la loi géométrique dont le paramètre est $\hat{p}$, calculé avec (5.4)*

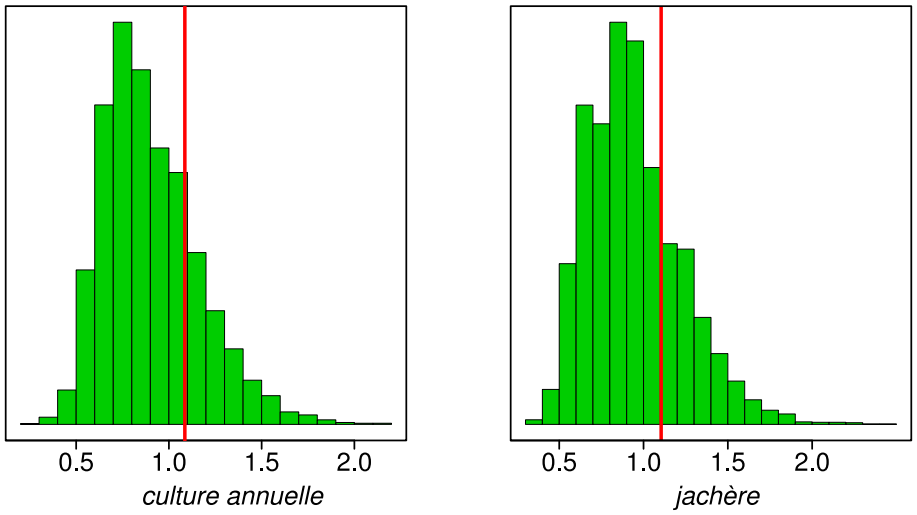

Figure 5.2: *Densités empiriques des échantillons K^m pour les deux états* `Ca` *et* `Jc`*. La ligne rouge verticale est la valeur de K^* calculée avec les données réelles.*

Les densités empiriques et les estimateurs des temps de séjours observés pour les états "culture annuelle" et "jachère" sont représentés dans la figure 5.1.

Dans la figure 5.2, nous avons tracé la densité empirique de l'échantillon K^m pour les états `Ca` et `Jc`. Pour l'état "culture annuelle", $K^* = 1.086$ et la p-valeur associée est 0.2240. Pour l'état "jachère", K^* vaut 1.104 et la p-valeur est 0.2555. Avec ces valeurs, on peut conclure que l'hypothèse de la loi géométrique est en accord avec les temps de séjour observés sur les deux états `Ca` et `Jc`.

5.2 Tests d'homogénéité

Selon l'hypothèse (h_2) du chapitre 4 page 36, la dynamique du processus $(X_n)_{n\in\mathbb{N}}$ est markovienne homogène. Cette hypothèse signifie que la matrice de transition Q qui dirige la succession des états d'une parcelle est constante au cours du temps (cf. chapitre 3). Cette matrice est la même pour tout l'ensemble des parcelles d'après l'hypothèse (h_1). Ce sont des hypothèses fortes car beaucoup de facteurs liés au temps et à l'espace peuvent influencer l'utilisation d'une parcelle donnée : par exemple l'utilisation des parcelles voisines, la mise en vigueur d'une loi sur la conservation de la forêt, accident ou catastrophes naturelles (feux de brousse, cyclone, etc.). Pour vérifier la consistance de ces deux hypothèses avec les données réelles, nous faisons appel à des tests que nous présentons maintenant.

5.2.1 Test d'homogénéité en espace selon la date de mise en culture

D'après le jeu de données du tableau 2.1, on voit clairement que l'on peut diviser l'ensemble des parcelles en deux groupes suivant l'année de mise en culture, d'un côté les parcelles :

$$\mathcal{P}_1 \overset{\text{déf}}{=} \{24, 25, \ldots, 32, 42\}$$

qui sont défrichées tôt, entre la deuxième et la quatrième année d'observation, et de l'autre, les parcelles :

$$\mathcal{P}_2 = \{1, 2, \ldots, 23, 33, 34, \ldots, 41\}$$

qui sont tardivement défrichées, au-delà de la $12^{\text{ème}}$ année d'observation. On a donc deux groupes de parcelles et on souhaite tester l'hypothèse selon laquelle la dynamique sur ces deux groupes est la même.

On calcule l'estimateur $\hat{Q}^{(g)}$ du maximum de vraisemblance de la matrice de transition associée à chacun de ces deux groupes $g = 1, 2$. Avec les deux groupes d'observations, on calcule l'estimateur de la matrice de transition par la méthode de maximum de vraisemblance. L'hypothèse nulle est donc :

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

où $\hat{Q}$ est l'estimateur du maximum de vraisemblance obtenu dans la section 4.3.1.

Sous l'hypothèse nulle, la fonction de vraisemblance maximisée est :

$$\prod_{i,j\in E} \hat{Q}_{ij}^{n_{ij}}.$$

Sous l'hypothèse alternative, cette fonction vaut :

$$\prod_{g=1}^{2} \prod_{i,j\in E} [\hat{Q}_{ij}^{(g)}]^{n_{ij}^{(g)}},$$

avec $n_{ij}^{(g)}$ est le nombre des transitions de l'état i à j dans le groupe d'observations $\mathcal{P}_g$.

Le critère utilisé pour effectuer le test est le rapport de vraisemblance qui s'écrit :

$$\Lambda^{(1)} \stackrel{\text{déf}}{=} \prod_{g=1}^{2} \prod_{i,j\in E} \left(\frac{\hat{Q}_{ij}}{\hat{Q}_{ij}^{(g)}} \right)^{n_{ij}^{(g)}}.$$

La statistique associée est :

$$T^{(1)} \stackrel{\text{déf}}{=} -2\log\Lambda^{(1)} = 2\sum_{i\in E}\sum_{j\in E}\sum_{g=1}^{2} n_{ij}^{(g)} \log\left[\frac{\hat{Q}_{ij}^{(g)}}{\hat{Q}_{ij}}\right] \tag{5.6}$$

qui est distribuée suivant une loi de χ^2 à 6 degrés de liberté [3].

Les estimations de la matrice de transitions pour les deux groupes donnent :

$$\hat{Q}^{(1)} = \begin{pmatrix} 0.6818 & 0.3068 & 0.0114 \\ 0.3117 & 0.6883 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \hat{Q}^{(2)} = \begin{pmatrix} 0.7815 & 0.2053 & 0.0132 \\ 0.38 & 0.62 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

En calculant (5.6) avec ces matrices, on trouve :

$$T^{(1)} = 3.6879.$$

Cette valeur correspond à une probabilité égale à 0.1345, voir figure 5.3. L'hypothèse nulle $H_0^{(1)}$ ne peut pas être rejetée et on retient l'hypothèse que les deux groupes d'observations $\mathcal{P}_1$ et $\mathcal{P}_1$ sont issues d'une même chaîne de Markov.

5.2.2 Test d'homogénéité en espace selon l'exposition

Géographiquement, on peut diviser les parcelles en trois groupes selon leurs expositions :

$$\begin{aligned} \textit{Ouest et sud–ouest} :&\{1,7,8,9,10,11,17,18,19,20\} = \mathcal{P}_1, \\ \textit{Nord–ouest} :&\{1,7,8,9,10,11,17,18,19,20\} = \mathcal{P}_2 \\ \textit{Nord–est} :&\{28,29,30,31,32,33,34,35,36,37,38,39,40,41,42\} = \mathcal{P}_3. \end{aligned}$$

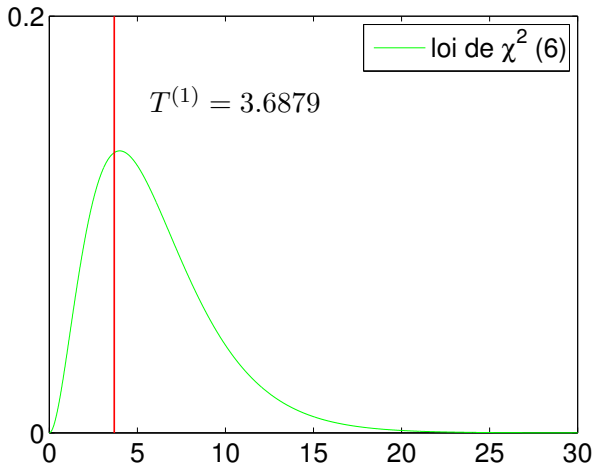

Figure 5.3: *Test d'homogénéité en espace selon l'année de mise en culture. La courbe verte est la densité de la loi de χ^2 à 6 degrés de liberté et la ligne rouge verticale est la valeur de la statistique $T^{(1)}$ calculée dans (5.6). On trouve $T^{(1)} = 3.6879$ correspondant à une probabilité de 0.1345. L'hypothèse nulle sur l'homogénéité n'est pas rejetée.*

On souhaite savoir si la position géographique influe sur la dynamique des parcelles. Pour cela, nous allons effectuer le même test d'homogénéité que précédemment.

Les EMVs de la matrice de transition pour $\mathcal{P}_g,\ g = 1, 2, 3$ sont :

$$\hat{Q}^{(1)} = \begin{pmatrix} 0.7 & 0.275 & 0.025 \\ 0.2609 & 0.7391 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \hat{Q}^{(2)} = \begin{pmatrix} 0.78 & 0.21 & 0.01 \\ 0.2679 & 0.7321 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\hat{Q}^{(3)} = \begin{pmatrix} 0.7273 & 0.2626 & 0.0101 \\ 0.4583 & 0.5417 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

L'hypothèse nulle correspondant au test est :

$$H_0^{(2)} : \hat{Q}^{(1)} = \hat{Q}^{(2)} = \hat{Q}^{(3)}, \tag{5.7}$$

et les matrices de transition sont égales à $\hat{Q}$ défini dans (4.12).

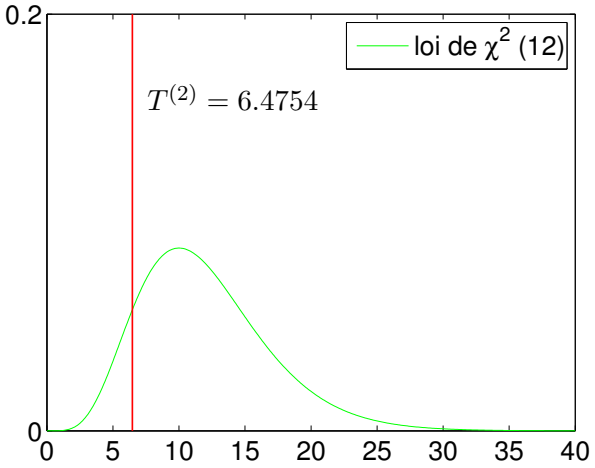

Figure 5.4: *Test d'homogénéité en espace selon l'exposition. L'hypothèse nulle sur l'homogénéité de la dynamique est acceptée.*

Le critère du rapport de vraisemblance est :

$$T^{(2)} = 2 \sum_{i \in E} \sum_{j \in E} \sum_{g=1}^{3} n_{ij}^{(g)} \, log \left[\frac{\hat{Q}_{(ij)}^{(g)}}{\hat{Q}_{ij}} \right] \tag{5.8}$$

qui suit une loi du χ^2 à 12 degrés de liberté [3]. Les résultats du test sont représentés dans la figure 5.4. On obtient $T^{(2)} = 6.4754$ avec une probabilité égale à 0.05819. Sur ce test, l'hypothèse nulle est acceptée et les dynamiques sur chacun des groupes de parcelles sont considérées comme égales.

5.2.3 Test d'homogénéité en temps

Pour tester l'hypothèse sur l'homogénéité en temps, nous divisons les observations en trois périodes :

(*i*) $\mathcal{P}_1$: observation des parcelles entre 1988 et 1995,

(*ii*) $\mathcal{P}_2$: observation des parcelles entre 1996 et 2000,

(*iii*) $\mathcal{P}_3$: observation des parcelles entre 2001 et 2006.

Soit :

$$\hat{Q}^{(1)} = \begin{pmatrix} 0.7207 & 0.2432 & 0.0270 \\ 0.4444 & 0.5556 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \hat{Q}^{(2)} = \begin{pmatrix} 0.8333 & 0.1667 & 0 \\ 0.2727 & 0.7273 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\hat{Q}^{(3)} = \begin{pmatrix} 0.7344 & 0.2500 & 0.0156 \\ 0.4110 & 0.5890 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Les estimateurs du maximum de vraisemblance pour les trois groupes d'observations. L'hypothèse nulle est :

$$H_0^{(3)} : \hat{Q}^{(1)} = \hat{Q}^{(2)} = \hat{Q}^{(3)},$$

et chacun de ces estimateurs est égal à $\hat{Q}$ définie dans (4.12).

La statistique adaptée :

$$T^{(3)} = 2 \sum_{i \in E} \sum_{j \in E} \sum_{g=1}^{3} n_{ij}^{(g)} \log \left[\frac{\hat{Q}_{(ij)}^{(g)}}{\hat{Q}_{ij}} \right]$$

suit une loi du χ^2 à 12 degrés de liberté.

Après calcul, $T^{(3)} = 2.5919$ qui correspond à une probabilité égale à 0.0042, voir figure 5.5. Avec cette valeur, l'hypothèse nulle $H_0^{(3)}$ ne peut pas être acceptée. Donc, d'après les données, la dynamique du processus n'est pas homogène en temps.

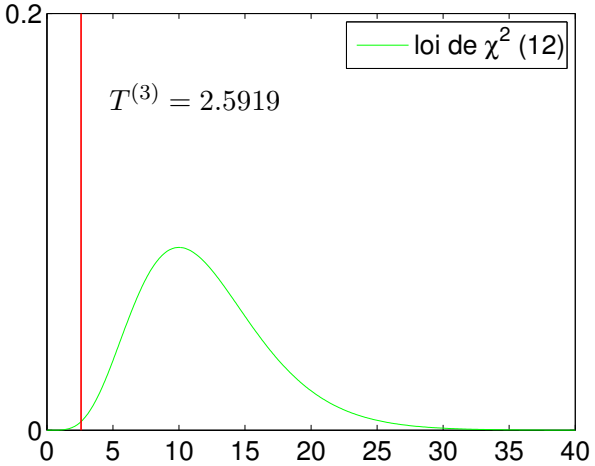

Figure 5.5: *Test d'homogénéité en temps. L'hypothèse nulle est rejetée : selon les données, la dynamique n'est pas homogène en temps.*

Chapitre 6

Généralités sur les chaînes semi-markoviennes

Introduit indépendamment par Lévy [30], Takás [52] et Smith [49], les processus semi-markoviens sont des généralisations des processus de Markov en enlevant les contraintes sur les lois des temps de séjour. On peut trouver plusieurs applications de ces processus [40, 14, 31] dans la littérature. Contrairement aux nombreux travaux sur les processus semi-markoviens à temps continu [29, 35, 38, 26], les processus semi-markoviens à temps discret ou chaînes semi-markoviennes n'ont pas suscité beaucoup d'attention. Une des rares références sur ce sujet est le livre de Barbu et Limnios [7]. Nous nous inspirons de ce livre dans la rédaction de ce présent chapitre.

6.1 Définitions et propriétés

On considère un processus en temps discret $(X_n)_{n\in\mathbb{N}}$ à valeurs dans un espace fini E. On définit la suite $(T_k)_{k\in\mathbb{N}}$ des instants de changement d'état, c'est-à-dire :

$$\begin{aligned} T_0 &\overset{\text{déf}}{=} 0, \\ T_k &\overset{\text{déf}}{=} \inf\{n > T_{k-1}; X_n \neq X_{T_{k-1}}\}, \end{aligned}$$

ainsi que la suite $(Y_k)_{k\in\mathbb{N}}$ des états successivement visités par $(X_n)_{n\in\mathbb{N}}$, c'est-à-dire :

$$Y_k \overset{\text{déf}}{=} X_{T_k}\,.$$

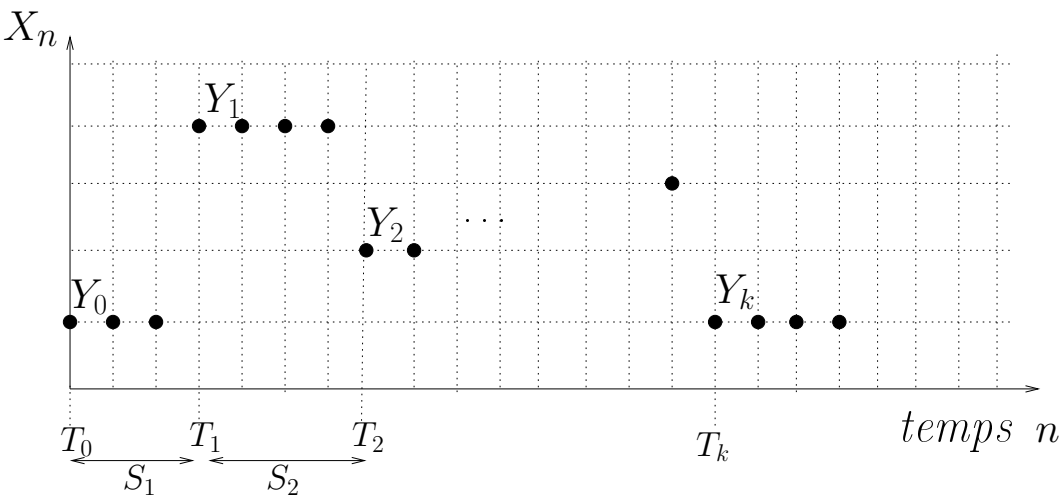

Figure 6.1: *Une trajectoire d'une chaîne semi-markovienne* $(X_n)_{n\in\mathbb{N}}$. *Le processus* $(Y_k)_{k\in\mathbb{N}}$ *est la chaîne de Markov incluse et* $(T_k)_{k\in\mathbb{N}}$ *les instants de changement d'état.*

On définit également les temps de séjour :

$$S_0 \overset{\text{déf}}{=} 0$$
$$S_k \overset{\text{déf}}{=} T_k - T_{k-1}.$$

On peut "inverser" la fonction $k \to T_k$ en définissant le nombre $M(n)$ de changements d'états à l'instant n, c'est-à-dire :

$$M(n) \overset{\text{déf}}{=} \max\{k \in \mathbb{N}; T_k \leq n\}.$$

On vérifie alors aisément que $X_n = Y_{M(n)}$.

6.1.1 Chaîne semi-markovienne

Définition 6.1.1. *On suppose que le couple* $(Y_k, T_k)_{k\in\mathbb{N}}$ *est une chaîne de renouvellement markovienne [14], c'est-à-dire :*

$$\begin{aligned}\mathbb{P}(Y_k =& i_k, T_k - T_{k-1} = s_k | Y_{0:k-1} = i_{0:k-1}, T_{0:k-1} = t_{0:k-1}) \\ &= \mathbb{P}(Y_k = i_k, T_k - T_{k-1} = s_k | Y_{k-1} = i_{k-1}).\end{aligned} \tag{6.1}$$

pour tout $k \geq 1,\ i_0, \ldots, i_k \in E,\ 0 = t_0 < t_1 < \cdots < t_{k-1} < t_k \in \mathbb{N}$ *et* $s_k \geq 1$. *Le processus* $(X_n)_{n\in\mathbb{N}}$ *est alors appelée chaîne semi-markovienne. Si l'équation* (6.1) *est indépendante de* n, *on dit que* $(Y_k, T_k)_{n\in\mathbb{N}}$ *est une chaîne de renouvellement homogène.*

On vérifie facilement que si $(Y_k, T_k)_{k\in\mathbb{N}}$ est homogène, alors le processus $(Y_k)_{k\in N}$ est une chaîne de Markov homogène, appelée chaîne incluse, de matrice de transition

$$Q_{ij} \stackrel{\text{déf}}{=} \mathbb{P}(Y_{k+1} = j | Y_k = i),\ i, j \in E, k \in \mathbb{N}.$$

Lorsque $M(n) = n\ \forall n \in \mathbb{N}$, alors $X_n = Y_n$, et $(X_n)_{n\in\mathbb{N}}$ est une chaîne de Markov. Mais en général, $(X_n)_{n\in\mathbb{N}}$ n'est pas une chaîne de Markov. Dans la suite, nous supposons que (Y_k, T_k) est une chaîne de renouvellement homogène.

Soit $\mathcal{M}_E(\mathbb{N})$ l'ensemble des applications de $\mathbb{N}$ dans $\mathcal{M}_E$ où $\mathcal{M}_E$ est l'ensemble des matrices carrées dans $E \times E$.

Définition 6.1.2. *On définit le noyau semi-markovien $q \in \mathcal{M}_E(\mathbb{N})$ par :*

$$q_{ij}(n) \stackrel{\text{déf}}{=} \mathbb{P}(Y_k = j,\ T_k - T_{k-1} = n | Y_{k-1} = i). \tag{6.2}$$

Il vérifie :

(i) $q_{ij}(n) \geq 0$, *pour tout* $i, j \in E$, $n \in \mathbb{N}$,

(ii) $q_{ij}(0) = 0$, *pour tout* $i, j \in E$,

(iii) $\sum_{n\in\mathbb{N}} \sum_{j\in E} q_{ij}(n) = 1$ *pour tout* $i \in E$.

La valeur $q_{ij}(n)$ est la probabilité de passer de i en j après un temps de séjour de n unités de temps en i. En utilisant l'équation (6.2), on vérifie qu'une chaîne de Markov de matrice de transition $(p_{ij})_{i,j\in E}$ est un cas particulier d'une chaîne semi-markovienne, dont le noyau est

$$q_{ij}(n) = \begin{cases} p_{ij}\,(p_{ii})^{n-1} & \text{si } i \neq j \text{ et } n \in \mathbb{N}^*, \\ 0 & \text{sinon.} \end{cases}$$

La matrice de transition de la chaîne incluse $(Y_k)_{k\in\mathbb{N}}$ peut être calculée à l'aide du noyau semi-markovien q par :

$$Q_{ij} = \sum_{n=0}^{\infty} q_{ij}(n)\,.$$

Définition 6.1.3. *On définit le produit de convolution $q * p$ des deux noyaux semi-markoviens $p, q \in \mathcal{M}_E(\mathbb{N})$ par :*

$$[q * p]_{ij}(n) \stackrel{\text{déf}}{=} \sum_{r\in E} \sum_{l=0}^{n} q_{ir}(n-l)\ p_{rj}(l)\,.$$

L'élément neutre du produit de convolution est :

$$[\delta I]_{ij}(n) = \begin{cases} 1 & \textit{si } i = j \textit{ et } n = 0, \\ 0 & \textit{sinon.} \end{cases}$$

Proposition 6.1.4. *Pour tout $i, j \in \mathbb{E}$ et $n, k \in \mathbb{N}$, on a :*

$$\mathbb{P}(Y_k = j, T_k = n | Y_0 = i) = q_{ij}^{(k)}(n), \forall i, j \in E \; n \in \mathbb{N}$$

où $q^{(k)}$ est la puissance de q au sens du produit de convolution :

$$q^{(k)} \overset{\text{déf}}{=} \underbrace{q * \cdots * q}_{k \text{ fois}},$$

en posant $q^{(0)} = \delta I$.

Démonstration. Démontrons par récurrence cette proposition. La proposition est évidente pour $k = 0$:

$$\mathbb{P}(Y_0 = j, T_0 = n | Y_0 = i) = q_{ij}^{(0)}(n) = \delta I$$

Pour $k = 1$, la proposition est vraie d'après la définition de q :

$$\mathbb{P}(Y_1 = j, \; T_1 - T_0 = n | Y_0 = i) = q_{ij}(n) = q_{ij}^{(1)}(n).$$

Pour $k \geq 2$, nous avons

$$\begin{aligned}
\mathbb{P}(Y_k = j, T_k = n | Y_0 = i) &= \sum_{r \in E} \sum_{l=1}^{n-1} \mathbb{P}(Y_k = j, T_k = n, Y_1 = r, T_1 = l | Y_0 = i)\,, \\
&= \sum_{r \in E} \sum_{l=1}^{n-1} \mathbb{P}(Y_k = j, T_k = n | Y_1 = r, T_1 = l, Y_0 = i)\; \mathbb{P}(Y_1 = r, T_1 = l | Y_0 = i), \\
&= \sum_{r \in E} \sum_{l=1}^{n-1} \mathbb{P}(Y_{k-1} = j, T_{k-1} = n - l | Y_0 = r)\; \mathbb{P}(Y_1 = r, T_1 = l | Y_0 = i), \\
&= \sum_{r \in E} \sum_{l=1}^{n-1} q_{rj}^{(k-1)}(n - l)\; q_{ir}(l) = q_{ij}^{(k)}(n)\,.
\end{aligned}$$

□

En particulier,

$$q_{ij}^{(2)}(n) = \sum_{r \in E} \sum_{l=0}^{n} q_{ir}(n - l)\, q_{rj}(l)$$

s'interprète comme la probabilité de passer de i en j en deux changements d'états, et en un temps de séjour cumulé n dont $(n - l)$ passés en i et l passés en r. De même $q_{ij}^{(n)}(k)$ est la probabilité de passer de i en j en n changements d'états et un temps de séjour cumulé de k unités de temps. Par conséquent :

$$q_{ij}^{(k)}(n) = 0, \; \forall k \geq n + 1. \tag{6.3}$$

6.1.2 Lois des temps de séjour

Pour tout $i, j \in E$, on définit :

(i) la loi conditionnelle des temps de séjour sur $\bar{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$ est donnée par :

$$f_{ij}(n) \overset{\text{déf}}{=} \mathbb{P}(S_{k+1} = n | Y_k = i, Y_{k+1} = j) = \begin{cases} \frac{q_{ij}(n)}{Q_{ij}}, n \in \mathbb{N} & 0 \text{ si } n = \infty, \text{ lorsque } Q_{ij} \neq 0, \\ 1_\infty(n) & \text{lorsque } Q_{ij} = 0. \end{cases} \tag{6.4}$$

(ii) la fonction de répartition conditionnelle des temps de séjour :

$$F_{ij}(n) \overset{\text{déf}}{=} \mathbb{P}(S_{k+1} \leq n | Y_k = i, Y_{k+1} = j) = \sum_{l=0}^{n} f_{ij}(l), \ n \in \mathbb{N}, \tag{6.5}$$

(iii) la loi des temps de séjour :

$$h_i(n) \overset{\text{déf}}{=} \mathbb{P}(S_{k+1} = n | Y_k = i) = \sum_{j \in E} q_{ij}(n), \ n \in \mathbb{N}, \tag{6.6}$$

(iv) la fonction de répartition des temps de séjour :

$$H_i(n) \overset{\text{déf}}{=} \mathbb{P}(S_{k+1} \leq n | Y_k = i) = \sum_{l=1}^{n} h_i(l), \ n \in \mathbb{N}, \tag{6.7}$$

(v) la fonction de survie conditionnelle des temps de séjour :

$$\bar{F}_{ij}(n) \overset{\text{déf}}{=} 1 - F_{ij}(n) = \mathbb{P}(S_{k+1} > n | Y_k = i, Y_{k+1} = j), \ n \in \mathbb{N}, \tag{6.8}$$

(vi) la fonction de survie des temps de séjour :

$$\bar{H}_i(n) \overset{\text{déf}}{=} 1 - H_i(n) = \mathbb{P}(S_{k+1} > n | Y_k = i) \ , n \in \mathbb{N}. \tag{6.9}$$

En se basant sur l'équation (6.7), la moyenne des temps de séjour sur un état j se calcule par :

$$\nu_j \overset{\text{déf}}{=} \mathbb{E}(T_1 | Y_0 = j) = \sum_{n \geq 0} n h_j(n) = \sum_{n \geq 0} (1 - H_j(n)). \tag{6.10}$$

6.1.3 Loi de transition semi-markovienne

Définition 6.1.5. *La loi de transition $P \in \mathcal{M}_E(\mathbb{N})$ de $(X_n)_{n \in \mathbb{N}}$ est définie par*

$$P_{ij}(n) \overset{\text{déf}}{=} \mathbb{P}(X_n = j | X_0 = i), \ i, j \in E, \ n \in \mathbb{N}. \tag{6.11}$$

Notons que la loi initiale d'une chaîne semi-markovienne est la même que celle de la chaîne incluse :

$$\alpha_i = \mathbb{P}(X_0 = i) = \mathbb{P}(Y_0 = i), \forall i \in E.$$

Dans la suite, nous allons étudier les outils qui permettent de calculer cette loi de transition.

Proposition 6.1.6. *La loi de transition $P \in \mathcal{M}_E(\mathbb{N})$ est une solution de l'équation dite de renouvellement :*

$$P(n) = (I - H)(n) + (q * P)(n), \forall n \in \mathbb{N} \tag{6.12}$$

où H est la matrice diagonale dont les éléments diagonaux sont H_i pour $i \in E$ et $q \in \mathcal{M}_E(\mathbb{N})$ le noyau semi-markovien défini dans (6.2).

Démonstration. Pour tout $i, j \in E$, $\forall\, n \in N$, on a

$$\begin{aligned}
P_{ij}(n) &= \mathbb{P}(X_n = j | X_0 = i), \\
&= \mathbb{P}(X_n = j, T_1 \leq n | X_0 = i) + \mathbb{P}(X_n = j, T_1 > n | X_0 = i), \\
&= \sum_{r \in E} \sum_{l=0}^{n} \mathbb{P}(X_n = j, X_{T_1} = r, T_1 = l | X_0 = i) \; + \; \delta_{ij} \; (1 - H_i(n)), \\
&= \sum_{r \in E} \sum_{l=0}^{n} \mathbb{P}(X_n = j | X_{T_1} = r, T_1 = l, X_0 = i) \; \mathbb{P}(Y_1 = r, T_1 = l | Y_0 = i) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + \delta_{ij} \; (1 - H_i(n)), \\
&= \sum_{r \in E} \sum_{l=0}^{n} \mathbb{P}(X_{n-l} = j | X_0 = r) \; \mathbb{P}(Y_1 = r, S_1 = l | Y_0 = i) + \delta_{ij} \; (1 - H_i(n)), \\
&= \delta_{ij} \; (1 - H_i(n)) + \sum_{r \in E} \sum_{l=0}^{n} P_{rj}(n - l) \; q_{ir}(l),
\end{aligned}$$

On peut de plus démontrer que P est l'unique solution de cette équation, voir [7]. □

En faisant une petite transformation, l'équation de renouvellement (6.12) peut s'écrire :

$$(\delta I - q) * P(n) = (I - H)(n), n \in \mathbb{N}.$$

Si $(\delta I - q)$ est inversible au sens du produit de convolution, alors :

$$P(n) = (\delta I - q)^{-1} * (I - H)(n), n \in \mathbb{N}.$$

Les propositions suivantes permettent de calculer l'inverse de $(\delta I - q)$ afin d'avoir une expression de la loi de transition semi-markovienne P.

Proposition 6.1.7. *Soit $q \in \mathcal{M}_E(\mathbb{N})$ un noyau semi-markovien, $q^{(-1)}$ existe et unique si et seulement si* $\det q(0) \neq 0$. *De plus, $\forall n \in \mathbb{N}$:*

$$q^{(-1)}(n) = \begin{cases} [q(0)]^{-1} & \textit{si } n = 0, \\ -[q(0)]^{-1} \sum_{l=0}^{n-1} q^{(-1)}(l) \; q(n - l) & \textit{sinon.} \end{cases} \tag{6.13}$$

Démonstration. Soit $p \in \mathcal{M}_E(\mathbb{N})$ l'inverse de q au sens du produit de convolution, alors :

$$\sum_{l=0}^{n} p(n-l)\, q(l) = \delta I(n),\ n \in \mathbb{N}.$$

Pour $n = 0$, $p(0)q(0) = \delta I(0) = I$; si $q(0)$ est inversible, i.e $\det q(0) \neq 0$, on a $p(0) = [q(0)]^{-1}$. Pour tout $n \geq 1$:

$$\sum_{l=0}^{n} p(n-l)\, q(l) = \delta I(n) = 0\,.$$

C'est-à-dire :

$$p(n)q(0) + \sum_{l=0}^{n-1} p(n-l)\, q(l) = 0,$$

d'où :

$$p(n) = -\Big[\sum_{l=0}^{n-1} q^{(-1)}(l)\, q(n-l)\Big]\, [q(0)]^{-1}\,.$$

□

Par conséquent, on peut montrer que $(\delta I - q)$ est inversible car :

$$\det(\delta I - q)(0) \neq 0\,.$$

En effet, $q(0) = 0$ et $(\delta I - q)(0) = \delta I(0) - q(0) = I_E$.

Proposition 6.1.8. *L'inverse ψ de $(\delta I - q)$ peut être calculée à l'aide du noyau semi-markovien par :*

$$\psi(n) = \sum_{m=0}^{n} q^{(m)}(n),\ n \in \mathbb{N}.$$

Démonstration. Pour tout $n \in \mathbb{N}$,

$$\begin{aligned}
\Big(\sum_{m=0}^{\infty} q^{(m)}\Big) * (\delta I - q)(n) &= \Big(\sum_{m=0}^{\infty} q^{(m)}\Big)(n) - \Big(\sum_{m=0}^{\infty} q^{(m)}\Big) * q(n) \\
&= \sum_{m=0}^{\infty} q^{(m)}(n) - \sum_{m=1}^{\infty} q^{(m)}(n), \\
&= q^{(0)}(n) = \delta I(n).
\end{aligned}$$

Comme l'inverse est unique d'après la Proposition 6.1.7, on a :

$$\psi(n) = \sum_{m=0}^{\infty} q^{(m)}(n).$$

D'après (6.3), nous avons $q^{(m)}(n) = 0$, $\forall m > n,\ n, m \in \mathbb{N}$, d'où :

$$\psi(n) = \sum_{m=0}^{n} q^{(m)}(n).$$

□

En utilisant la Proposition 6.1.4, on peut vérifier que, pour tout $i, j \in E$:

$$\psi_{ij}(n) = \mathbb{P}\big(\cup_{l=0}^{n} \{Y_l = j, T_l = n\} | Y_0 = i\big), \tag{6.14}$$

qui est la probabilité, partant de i, de faire un saut en j au temps n. $\psi_{ii}(n)$ est la probabilité que la chaîne se renouvelle en i au temps n.

L'équation (6.15) propose donc une méthode de calcul de la loi de transition semi-markovienne :

$$P(n) = \sum_{m=0}^{n} q^{(m)} * (I - H)(n), \ \forall n \in \mathbb{N}. \tag{6.15}$$

6.2 Comportement asymptotique

6.2.1 Classification des états

Pour un état $j \in E$, notons $(T_k^j)_{k>0}$ les temps de retour successifs du processus en j et T_0^j le temps du premier passage sur cet état, c'est-à-dire :

$$\begin{aligned} T_0^j &\overset{\text{déf}}{=} T_m, \text{ avec } m = \min\{l \in \mathbb{N} | Y_l = j\}, \\ T_k^j &\overset{\text{déf}}{=} T_m, \text{ avec } m = \min\{l \in \mathbb{N} | Y_l = j, T_l > T_{k-1}^j\}, k \geq 1. \end{aligned}$$

Pour tout $i \neq j$, soit :

(*i*) $g_{ij}(n) \overset{\text{déf}}{=} \mathbb{P}_i(T_0^j = n)$, $n \in \mathbb{N}$, la probabilité, partant de i, que le processus passe en j au temps n, avec $g_{ij}(0) = 0$. Sa moyenne est :

$$\mu_{ij} \overset{\text{déf}}{=} \mathbb{E}_i(T_0^j) = \sum_{n \geq 0} g_{ij}(n)\,;$$

(*ii*) $g_{jj}(n) = \mathbb{P}_j(T_1^j = n)$, $n \in \mathbb{N}$, la probabilité, partant j, que le processus repasse en j au temps n (temps de récurrence sur j), avec $g_{jj}(0) = 0$ et :

$$\mu_{jj}(n) \overset{\text{déf}}{=} \mathbb{E}_i(T_1^j) = \sum_{n \geq 0} g_{jj}(n)$$

sa moyenne.

(*iii*) $G_{jj}(n) = \sum_{l=0}^{n} g_{jj}(l)$, $n \in \mathbb{N}$, la probabilité, partant de j, que le processus repasse en j au plus tard au temps n.

Ainsi, pour deux états différents i et j de E, $G_{ij}(\infty)$ est la probabilité que le processus passe de i en j.

Définition 6.2.1. *Si $G_{ij}(\infty) \times G_{ji}(\infty) > 0$, on dit que i et j communiquent, noté $i \leftrightarrow j$. La relation "$i \leftrightarrow j$" est une relation d'équivalence sur E. On appelle classes les éléments de l'espace quotient $E|_{i \leftrightarrow j}$. S'il n'existe qu'une seule classe, c'est-à-dire si tous les états communiquent, on dira que la chaîne semi-markovienne est irréductible.*

Définition 6.2.2. *Un état i est dit récurrent si $G_{ii}(\infty) = 1$. Dans le cas contraire, c'est à dire si $G_{ii}(\infty) < 1$, on dit qu'il est transitoire.*

Si tous les états sont récurrents, on dit que la chaîne semi-markovienne est récurrente.

Un état récurrent i est dit récurrent positif si $\mu_{ii} < \infty$. Il est récurrent nul si $\mu_{ii} = \infty$.

Si tous les états sont récurrents positifs (resp. nuls), on dit que la chaîne semi-markovienne est récurrente positive (resp. nulle).

Définition 6.2.3. *Une chaîne semi-markovienne est dite ergodique si elle est irréductible et récurrente positive.*

Soit $d > 1$ un entier, un état $i \in E$ est d-périodique si $g_{ii}(.)$ est périodique et que sa période est égale à d. Une chaîne semi-markovienne irréductible est d-périodique ($d > 1$) si tous les états sont d-périodiques. Dans le cas contraire, on dit qu'elle est apériodique.

6.2.2 Loi stationnaire

Proposition 6.2.4. *Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne semi-markovienne apériodique. La moyenne des temps de récurrence pour un état $j \in E$ est :*

$$\mu_{jj} = \frac{\sum_{i \in E} \pi(i)\, \nu_i}{\pi(j)},$$

où π est la loi stationnaire de la chaîne incluse $(Y_k)_{k \in \mathbb{N}}$ et ν le vecteur dont les éléments sont les moyennes des temps de séjour sur les états.

Si on pose :

$$\bar{m} \stackrel{\text{déf}}{=} \sum_{i \in E} \pi(i)\, \nu_i,$$

alors :

$$\mu_{jj} = \frac{\bar{m}}{\pi(j)}.$$

Démonstration. Pour tout $i, j \in E$

$$\begin{aligned}
\mu_{ij} &= \mathbb{E}_i(T_0^j 1_{\{Y_1=j\}}) + \mathbb{E}_i(T_0^j 1_{\{Y_1\neq j\}}) \\
&= Q_{ij}\,\nu_i + \sum_{r\neq j} \mathbb{E}_i(T_0^j | 1_{\{Y_1=r\}})\mathbb{P}_i(Y_1 = r) \\
&= Q_{ij}\,\nu_i + \sum_{r\neq j} \left(\mathbb{E}_r(T_0^j) + \mathbb{E}_i(T_1)\right) \\
&= Q_{ij}\,\nu_i + \sum_{r\neq j} Q_{ir}\,\mu_{rj} + (1 - Q_{ij})\,m_i.
\end{aligned}$$

On obtient alors $\mu_{ij} = \nu_i + \sum_{r\neq j} Q_{ir}\,\mu_{rj}$.

En multipliant les deux membres par $\pi(i)$ et en sommant suivant i on obtient :

$$\begin{aligned}
\sum_{i\in E} \pi(i)\mu_{ij} &= \sum_{i\in E} \pi(i)\,\nu_i + \sum_{i\in E} \pi(i) \sum_{r\neq j} Q_{ir}\,\mu_{rj} \\
&= \bar{m} + \sum_{r\neq j} \left(\sum_{i\in E} \pi(i)\,Q_{ir} \right) \mu_{rj} \\
&= \bar{m} + \sum_{r\in E} \pi(r)\,\mu_{rj} - \pi(j)\,\mu_{jj}.
\end{aligned}$$

Par conséquent $\bar{m} = \pi(j)\,\mu_{jj}$. □

Proposition 6.2.5. *Une chaîne semi-markovienne $(X_n)_{n\in\mathbb{N}}$ irréductible et apériodique admet une loi limite :*

$$\Pi \overset{\text{déf}}{=} \lim_{n\to\infty} P(n),$$

telle que :

$$\Pi_j = \frac{1}{\mu_{jj}}\,\nu_j = \frac{\pi(j)\,\nu_j}{\sum_{i\in E} \pi(i)\,\nu_i}, \quad j \in E. \tag{6.16}$$

où π est la loi stationnaire de la chaîne de Markov incluse et ν un vecteur contenant les moyennes des temps de séjour sur les états.

La démonstration de la proposition 6.2.5 fait appel à la théorie du renouvellement markovien que nous n'abordons pas ici. On peut se reporter à [7, p. 65].

6.3 Exemple d'application

On considère une chaîne semi-markovienne $(X_n)_{n\in\mathbb{N}}$ dans un espace d'état E contenant 3 états :

$$E = \{1, 2, 3\}.$$

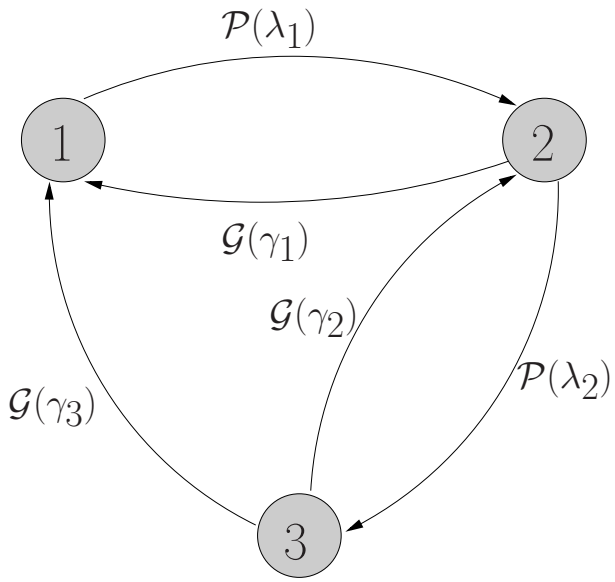

Figure 6.2: *Modèle semi-markovien à trois états.*

Soit $\alpha = (\alpha_1,\ \alpha_2,\ \alpha_3)$ sa loi initiale,

$$Q = \begin{pmatrix} 0 & 1 & 0 \\ \theta_1 & 0 & 1-\theta_1 \\ \theta_2 & 1-\theta_2 & 0 \end{pmatrix}, \qquad \theta = (\theta_1,\ \theta_2) \in [0,1]^2$$

sa matrice de transition, et enfin :

$$f(k) = \begin{pmatrix} 0 & \mathcal{P}(\lambda_1, k) & 0 \\ \mathcal{G}(\gamma_1, k) & 0 & \mathcal{P}(\lambda_2, k) \\ \mathcal{G}(\gamma_3, k) & \mathcal{G}(\gamma_2, k) & 0 \end{pmatrix}, k \in \mathbb{N}$$

la loi conditionnelle des temps de séjour où $\mathcal{G}(\gamma_i, .)_{i=1,2}$ est la loi géométrique de paramètre γ_i et $\mathcal{P}(\lambda_i, .)_{i=1,2}$ la loi de Poisson de paramètre λ_j.

En utilisant (6.4), le noyau semi-markovien de $(X_n)_{n\in\mathbb{N}}$ est :

$$q(k) = \begin{pmatrix} 0 & \mathcal{P}(\lambda_1, k) & 0 \\ \theta_1\, \mathcal{G}(\gamma_1, k) & 0 & (1-\theta_1)\, \mathcal{P}(\lambda_2, k) \\ \theta_2\, \mathcal{G}(\gamma_3, k) & (1-\theta_1)\, \mathcal{G}(\gamma_2, k) & 0 \end{pmatrix}, k \in \mathbb{N}.$$

Pour calculer la loi de transition semi-markovienne P selon (6.15), on a besoin

de calculer $(I-H)(k)$ pour tout $k \in \mathbb{N}$ avec :

$$\begin{aligned}
H_1(k) &= \sum_{l=1}^{k}\sum_{i=1}^{3} q_{1i}(l) = \sum_{l=1}^{k} \mathcal{P}(\lambda_1, l),\\
H_2(k) &= \sum_{l=1}^{k}\sum_{i=1}^{3} q_{2i}(l) = \sum_{l=1}^{k} \theta_1\, \mathcal{G}(\gamma_1, l) + (1-\theta_1)\, \mathcal{P}(\lambda_2, l),\\
H_3(k) &= \sum_{l=1}^{k}\sum_{i=1}^{3} q_{3i}(l) = \sum_{l=1}^{k} \theta_2\, \mathcal{G}(\gamma_2, l) + (1-\theta_1)\, \mathcal{G}(\gamma_3, l).
\end{aligned}$$

Nous obtenons :

$$(I-H)(k) = \operatorname{diag}\begin{pmatrix} 1-\sum_{l=1}^{k} \mathcal{P}(\lambda_1,k) \\ 1-\sum_{l=1}^{k}[\theta_1\, \mathcal{G}(\gamma_1,l)+(1-\theta_1)\, \mathcal{P}(\lambda_2,l)] \\ 1-\sum_{l=1}^{k}[\theta_2\, \mathcal{G}(\gamma_2,l)+(1-\theta_2)\, \mathcal{G}(\gamma_3,l)] \end{pmatrix}$$

pour tout $k \in \mathbb{N}$.

La moyenne $m = (m_1,\, m_2,\, m_3)$ des temps de séjour sur les états se calcule à l'aide de (6.10) :

$$\begin{aligned}
m_1 &= \sum_{k\geq 1}(1-H_1(k)) = \sum_{k\geq 1}[1-\sum_{l=1}^{k} \mathcal{P}(\lambda_1, l)],\\
m_2 &= \sum_{k\geq 1}(1-H_2(k)) = \sum_{k\geq 1}[1-\sum_{l=1}^{k}(\theta_1\, \mathcal{G}(\gamma_1, l) + (1-\theta_1)\, \mathcal{P}(\lambda_2, l)],\\
m_3 &= \sum_{k\geq 1}(1-H_3(k)) = \sum_{k\geq 1}[1-\sum_{l=1}^{k}(\theta_2\, \mathcal{G}(\gamma_2, l)] + (1-\theta_1)\, \mathcal{G}(\gamma_3, l)].
\end{aligned}$$

On trouve $m = (6.0465,\, 5.3001,\, 1.6666)$.

La loi de transition semi-markovienne s'obtient enfin par :

$$P(n) = \sum_{l=0}^{l} q^{(l)} * (I-H)(n),\ \forall n \in \mathbb{N}.$$

En prenant $\alpha = (0.5,\ 0.3,\ 0.2)$ et $\theta = (0.6,\ 0.3)$, la Figure 6.3 représente la loi de transition semi-markovienne P. Le processus converge rapidement vers sa loi limite, voir Figure 6.4, obtenue numériquement d'après (6.16) :

$$\Pi = (0.4218,\, 0.5136,\, 0.0646)\,.$$

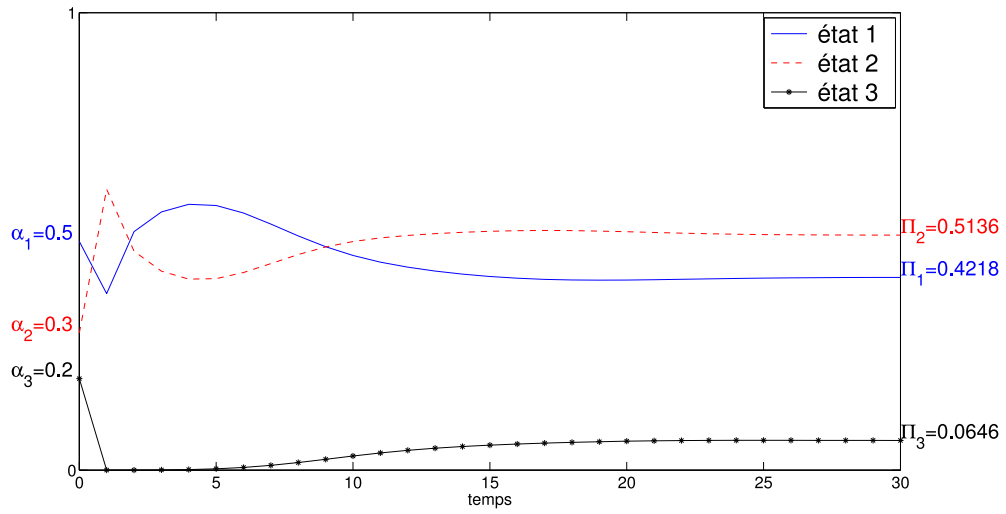

Figure 6.3: *Modèle semi-markovien à trois états :* Loi de transition de la chaîne semi-markovienne décrite par la Figure 6.2 . $\alpha = (\alpha_1,\ \alpha_2,\ \alpha_3)$ est la loi initiale et $\Pi = (\Pi_1,\ \Pi_2,\ \Pi_3)$ la loi limite.

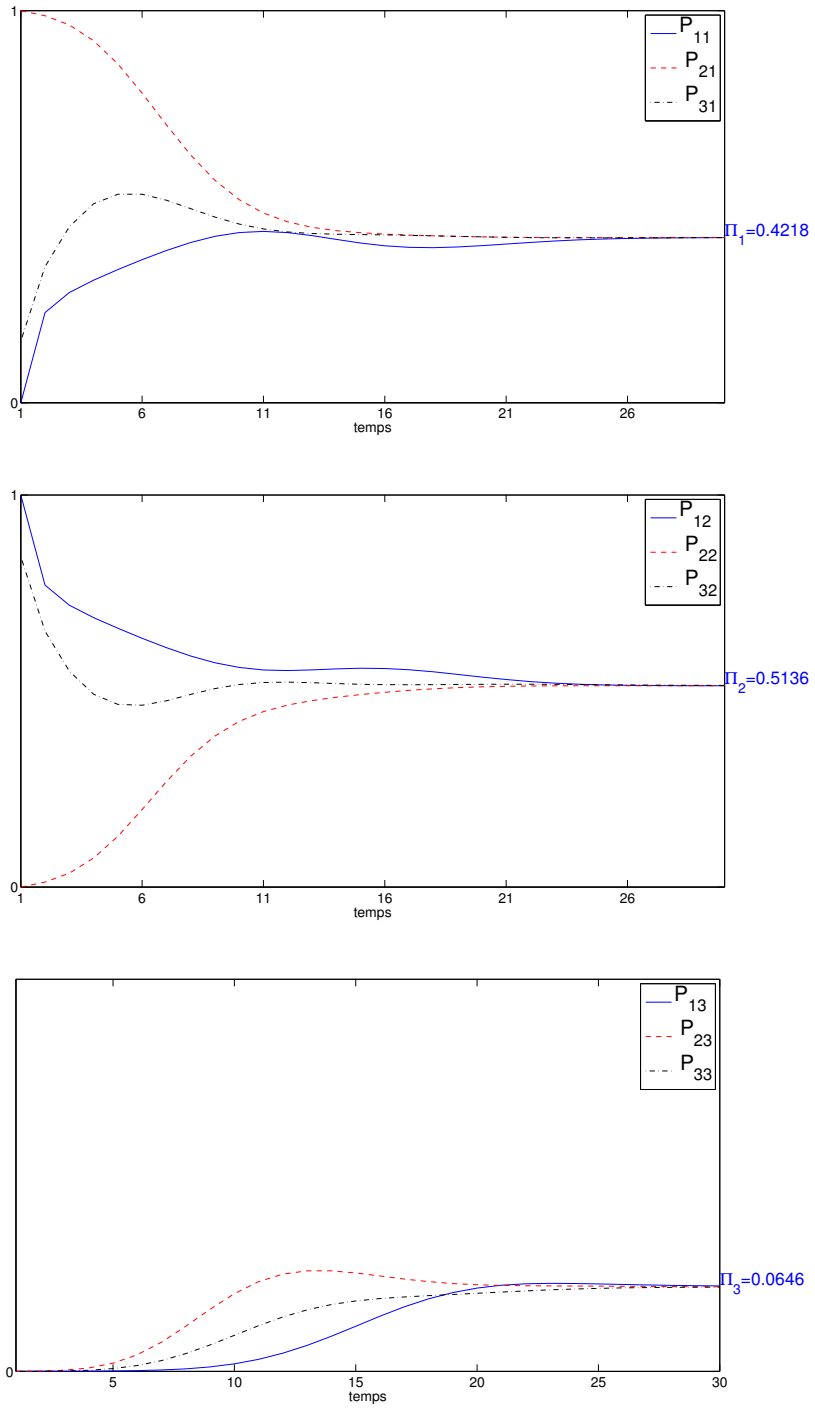

Figure 6.4: *Modèle semi-markovien à trois états :* Convergence de la loi semi-markovienne P vers la loi limite Π, en partant des différents états initiaux.

Chapitre 7

Modèles semi-markoviens

Comme les chaînes de Markov sont des outils relativement simples, elles imposent des contraintes assez fortes sur les modèles. L'hypothèse de Markov impose que les temps de séjour dans un état donné sont des lois géométriques qui en plus ne dépendent pas de l'état qui sera ensuite visité. Dans le premier jeu de données, cette hypothèse est acceptable mais, comme nous le verrons, elle ne l'est plus du tout dans le cas du second jeu de données. Les modèles semi-markoviens permettent justement de relâcher cette hypothèse : tout en gardant une facilité d'analyse des chaînes de Markov, ils permettent de prescrire les lois de temps de séjour et de faire dépendre ces lois de l'état visité à l'itération suivante. On trouve beaucoup d'applications de ces modèles dans la littérature [36, 28, 31]. Dans un modèle à temps continu, Monticino *et al.* [35] l'utilisent dans l'étude d'une dynamique des parcelles forestières.

Dans le cadre d'un modèle à temps discret, Barbu et Limnios [7] estiment la loi de transition d'un modèle semi-markovien par une méthode empirique (non paramétrique). Dans un article récent [6], ils montrent comment estimer la loi stationnaire d'une chaîne de semi-Markov à temps discret via cette méthode.

7.1 Motivation

Dans ce chapitre, nous utilisons les chaînes semi-markoviennes pour modéliser la dynamique d'occupation des sols dans le village d'Ambendrana (cf. figure 2.1) [43]. Nous travaillons sur le deuxième jeu de données concernant l'usage des 131 parcelles de 1956 à 2006 (cf. figure 2.3). En écartant les deux états "statiques" forêt naturelle et rizière, il reste $P = 112$ parcelles et 5 états possibles : culture annuelle (`Ca`), jachère (`Jc`), herbe (`He`), forêt secondaire (`Fs`) et culture

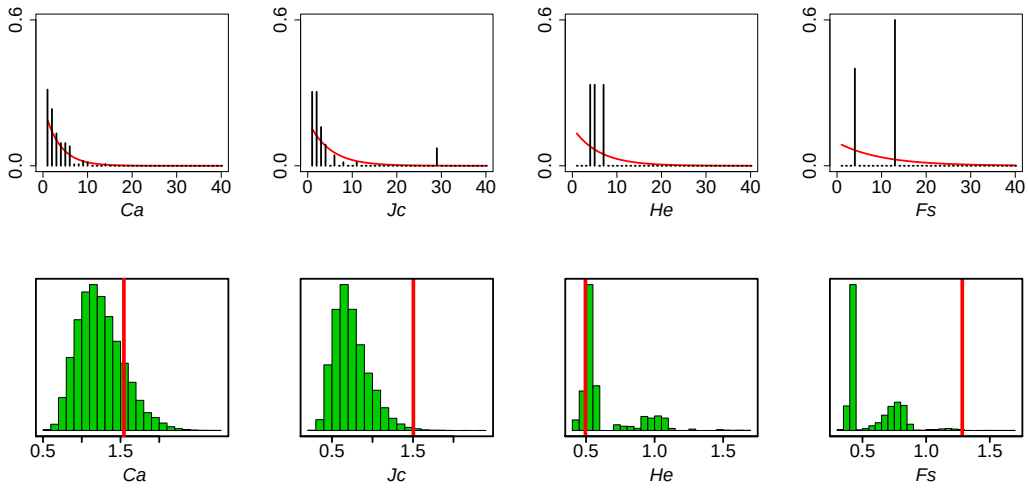

Figure 7.1: Test d'adéquation des lois des temps de séjour sur les états "culture annuelle" (`Ca`), "jachère (`Jc`), "herbe" (`He`) et "forêt secondaire" (`Fs`). Le temps de séjour sur l'état "culture pérenne" (`Cp`) n'est pas traité car il est infini d'après le jeu de données, voir figure 2.3.

pérenne (`Cp`).

Markov ou semi-Markov ?

Au chapitre 5, nous avons utilisé une technique de bootstrap paramétrique pour tester l'adéquation d'un modèle markovien à un jeu de données. Ce test porte sur la loi des temps séjour sur les états qui doit être nécessairement géométrique pour un modèle de Markov. Pour ce deuxième jeu de données, nous utilisons cette même technique afin de savoir si le modèle de Markov est le mieux adapté pour modéliser le processus.

Le résultat est représenté dans la figure 7.1. Le test est positif pour les états "culture annuelle" (`Ca`) et " herbe" (`He`). L'hypothèse de la loi géométrique est en revanche rejetée pour les temps de séjour en "jachère" (`Jc`) et "forêt secondaire" (`Fs`).

Nous avons refait le même test pour le temps de séjour sur l'état "jachère" (`Jc`) en séparant les temps de séjour en `Jc` suivis de l'état `Ca` de ceux suivis de l'état `Fs`. Le résultat est positif pour une transition `Jc` $\rightarrow$ `Ca` et négatif pour `Jc` $\rightarrow$ `Fs`, voir figure 7.2.

Donc d'après ce test :

(i) les lois des temps de séjour sur les états ne sont pas tous géométriques,

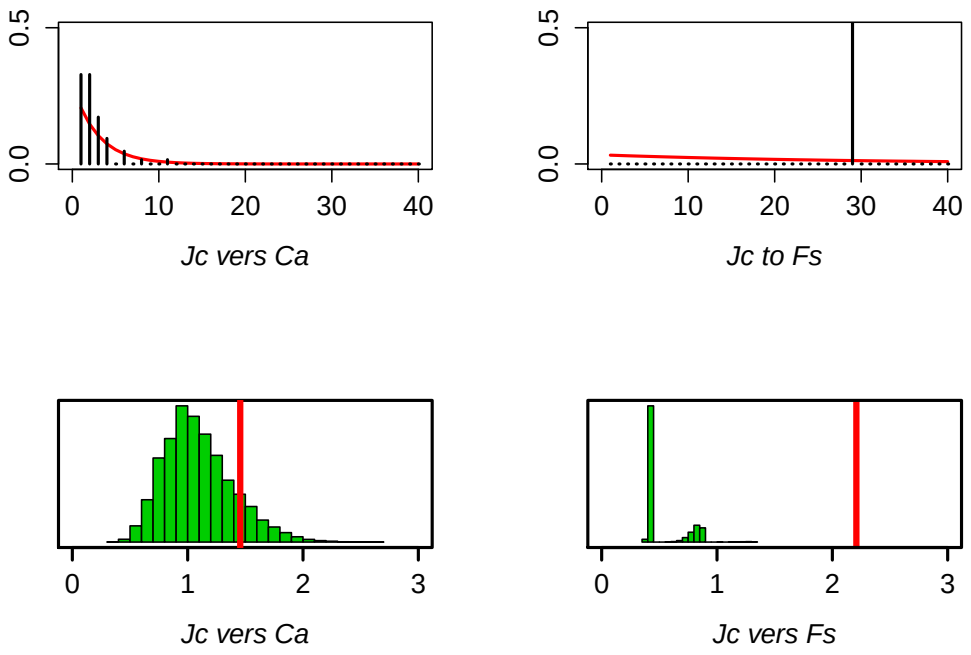

Figure 7.2: *Test d'adéquation des lois des temps de séjour en Jachère* (`Jc`)*, pour les transitions* `Jc` $\rightarrow$ `Ca`*, et* `Jc` $\rightarrow$ `Fs`*. L'hypothèse de la loi géométrique est rejetée pour le temps de séjour dans une transition* `Jc` $\rightarrow$ `Fs`*.*

(*ii*) pour une transition `Jc` $\rightarrow i \in E \setminus$ `Jc`, la loi du temps de séjour en `Jc` dépend de i.

Il est évident que le temps de séjour en "jachère"(`Jc`) est beaucoup plus long dans le cas d'une transition `Jc` $\rightarrow$ `Fs` que dans le cas d'une transition `Jc` $\rightarrow$ `Ca`. Le test que l'on vient d'effectuer ne fait que confirmer cette évidence. Un modèle de Markov est donc inadapté à cette situation et le recours à des modèles de semi-Markov est naturel.

7.2 Description du modèle

7.2.1 Les observations

L'état d'une parcelle p au temps n est noté $e_n^{(p)} \in E \stackrel{\text{déf}}{=} \{\texttt{Ca}, \texttt{Jc}, \texttt{He}, \texttt{Fs}, \texttt{Cp}\}$. L'ensemble des observations pour toute les parcelles s'écrit alors :

$$\left(e_{n=1:N_p}^{(p)}\right)_{p=1:P},$$

où N_p est la longueur des observations de la parcelle p.

Soit $(t_k^{(p)})_{k\in\mathbb{N}}$ les instants de changement d'état pour une parcelle p et $(s_k^{(p)})_{k\in\mathbb{N}}$ la suite des temps de séjour sur les états avec :

$$t_k^{(p)} = \sum_{l=0}^{k} s_l^{(p)}.$$

Le nombre des changements d'état observés pour la parcelle p est :

$$M^{(p)} \stackrel{\text{déf}}{=} \max\{k \in \mathbb{N}, t_k^{(p)} \leq N_p\}\,.$$

Enfin, le temps de séjour censuré sur le dernier état visité $e_{M^{(p)}}^{(p)}$ sera noté

$$u_{N_p} \stackrel{\text{déf}}{=} N_p - T_{N_p}\,.$$

7.2.2 Hypothèses et présentation du modèle

D'après les tests statistiques effectués sur le jeu de données et des concertations avec les experts, nous faisons les hypothèses suivantes :

(h_1) *Les parcelles sont indépendantes les unes des autres et ont la même dynamique.*

(h_2) *Les changements d'état* $(Y_k)_{k\in\mathbb{N}}$ *forment une chaîne de Markov.*

(h_3) *Les transitions possibles sont* `Ca` $\to$ `Jc`, `Ca` $\to$ `He`, `Ca` $\to$ `Cp`, `Jc` $\to$ `Ca`, `Fa` $\to$ `Fs`, `He` $\to$ `Ca`, `He` $\to$ `Jc`, `Cp` $\to$ `Ca` *et* `Fs` $\to$ `Ca`. *Les autres transitions sont impossibles.*

(h_4) *Le temps de séjour sur chaque état dépend du prochain état qui sera visité. Pour les transitions* `Ca` $\to$ `Jc`, `Ca` $\to$ `He`, `Ca` $\to$ `Cp`, *et* `Jc` $\to$ `Ca`, *les lois des temps de séjour sont géométriques. Les temps de séjour pour les autres transitions sont poissonniens.*

Ainsi, nous supposons que les observations sont issues d'une chaîne semi-markovienne $(X_n)_{n\in\mathbb{N}}$, voir figure 7.3. Soit $(Y_k)_{k\in\mathbb{N}}$ la chaîne de Markov incluse, $(T_k)_{k\in\mathbb{N}}$ la suite des instants de changement d'état et $(S_k)_{k\in\mathbb{N}}$ la suite des temps de séjour.

La figure 7.4 représente la matrice de transition de la chaîne incluse $(Y_k)_{k\in\mathbb{N}}$ avec :

$$Q = \begin{pmatrix} 0 & \theta_1 & \theta_2 & 0 & 1-\theta_1-\theta_2 \\ \theta_3 & 0 & 0 & 1-\theta_3 & 0 \\ \theta_4 & 1-\theta_4 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{7.1}$$

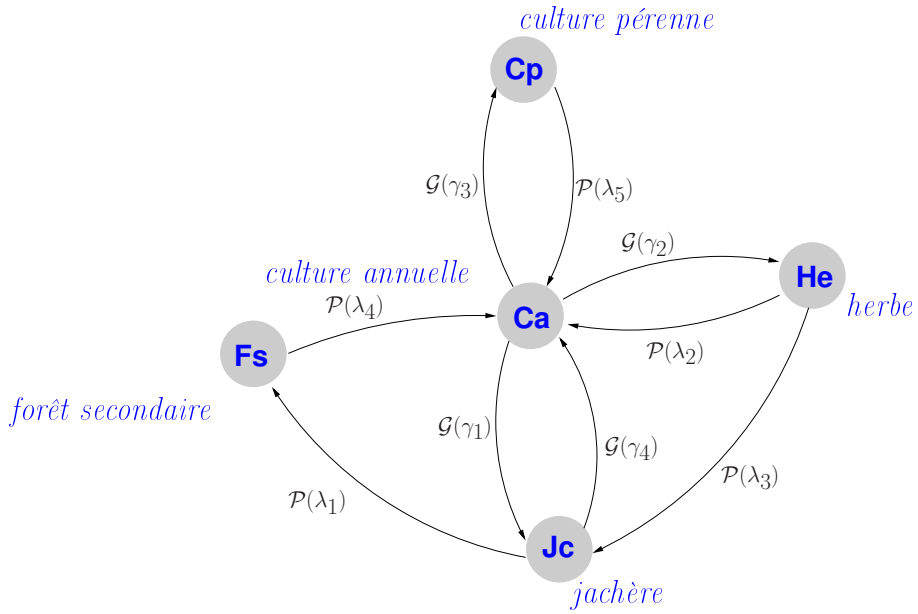

Figure 7.3: *Modèle semi-markovien à 5 états.* $\mathcal{G}(\gamma_i)_{1\leq i\leq 4}$ *est la loi géométrique de paramètre* γ_i *et* $\mathcal{P}(\lambda_i)_{1\leq i\leq 5}$ *la loi de Poisson de paramètre* λ_i

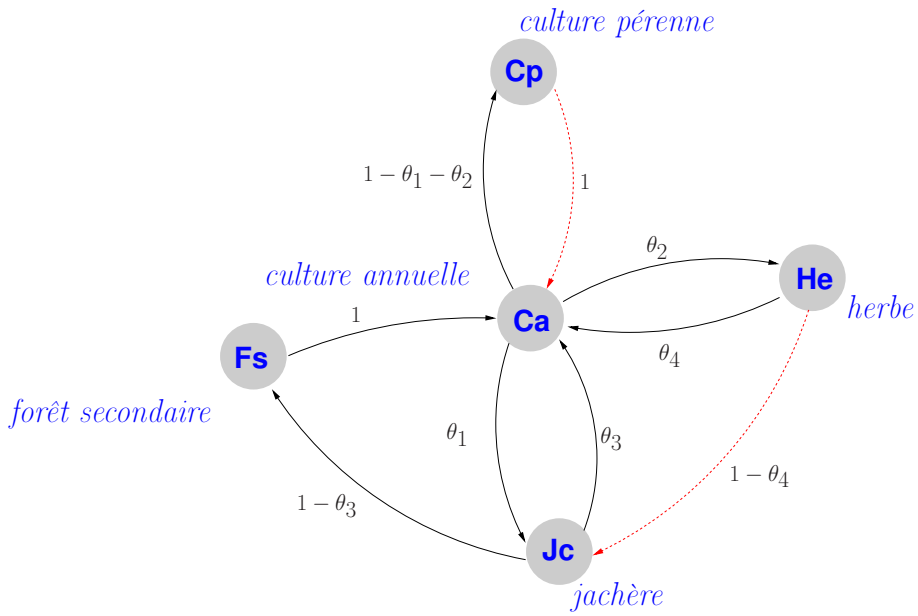

Figure 7.4: *Graphe de la matrice de transition de la chaîne de Makov incluse du processus* $(X_n)_{n\in\mathbb{N}}$.

où

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

La loi conditionnelle des temps de séjour est de la forme :

$$f(k) = \begin{pmatrix} 0 & \mathcal{G}(\gamma_1, k) & \mathcal{G}(\gamma_2, k) & 0 & \mathcal{G}(\gamma_3, k) \\ \mathcal{G}(\gamma_4, k) & 0 & 0 & \mathcal{P}(\lambda_1, k) & 0 \\ \mathcal{P}(\lambda_2, k) & \mathcal{P}(\lambda_3, k) & 0 & 0 & 0 \\ \mathcal{P}(\lambda_4, k) & 0 & 0 & 0 & 0 \\ \mathcal{P}(\lambda_5, k) & 0 & 0 & 0 & 0 \end{pmatrix} \tag{7.2}$$

où $\mathcal{G}(\gamma_i, .)$ est la loi géométrique de paramètre γ_i, , $1 \leq i \leq 4$ et $\mathcal{P}(\lambda_j, .)$ la loi de Poisson de paramètre λ_j, $1 \leq j \leq 5$.

En utilisant Q et f, la loi du processus s'écrit :

$$\mathbb{P}(Y^{(p)}_{0:M^{(p)}} = e^{(p)}_{0:M^{(p)}}, S^{(p)}_{1:M^{(p)}} = s^{(p)}_{1:M^{(p)}})\ \mathbb{P}(S^{(p)}_{M^{(p)}+1} \geq u_{N_p} | Y^{(p)}_{M^{(p)}} = e^{(p)}_{M^{(p)}})_{p=1:P}$$

$$= \prod_{p=1}^{P} \mathbb{P}(Y^{(p)}_{0:M^{(p)}} = e^{(p)}_{0:M^{(p)}}, S^{(p)}_{1:M^{(p)}} = s^{(p)}_{1:M^{(p)}})$$
$$\mathbb{P}(S^{(p)}_{M^{(p)}+1} \geq u_{N_p} | Y^{(p)}_{M^{(p)}} = e^{(p)}_{M^{(p)}})$$

$$= \prod_{p=1}^{P} \prod_{k=1}^{M^{(p)}} \delta_{\mathtt{Ca}}(e^{(p)}_0)\ Q_{e^{(p)}_{k-1} e^{(p)}_k}\ f_{e^{(p)}_{k-1} e^{(p)}_k}(s^{(p)}_k)\ \bar{H}_{e^{(p)}_{M^{(p)}}}(u_{N_p}) \tag{7.3}$$

avec $\delta_{\mathtt{Ca}}$ la loi initiale de $(Y_k)_{k\in\mathbb{N}}$ et :

$$\bar{H}^{(p)}_i(k) = \mathbb{P}(S^{(p)}_{M^{(p)}+1} > k | X_{M^{(p)}} = i),\ i \in E, k \in \mathbb{N},$$

est la fonction de survie définie dans (6.9).

7.3 Inférence des modèles

La loi du processus (7.3) est entièrement définie par la matrice de transition de la chaîne incluse Q, la loi conditionnelle des temps de séjour f et la fonction de survie des temps de séjour $\bar{H}$.

D'après (6.4), on peut calculer le noyau semi-markovien de $(X_n)_{n\geq 0}$ par :

$$q_{ij}(k) = Q_{ij}\ f_{ij}(k), \forall i, j \in E, k \in \mathbb{N}^*.$$

On en déduit, l'expression de la fonction de survie $\bar{H}$ avec (6.6), (6.7) et (6.9).

En notant que la loi initiale est $\delta_{\mathtt{Ca}}$, inférer le modèle revient alors à estimer la matrice de transition de la chaîne incluse Q et la loi conditionnelle des temps de séjour f à partir des observations.

Pour une parcelle p, pour tout $i, j \in E$, on définit :

(i) le nombre des transitions de i à j :

$$n_{ij}^{(p)} \overset{\text{déf}}{=} \sum_{k=0}^{M^{(p)}} 1_{\{Y_k=i, Y_{k+1}=j\}},$$

(ii) le nombre des transitions de i à j dont le temps de séjour sur l'état i est égal à $k \geq 1$:

$$n_{ij}^{(p)}(k) \overset{\text{déf}}{=} \sum_{n=0}^{M^{(p)}} 1_{\{Y_n=i, Y_{n+1}=j, S_n=k\}},$$

(iii) le nombre des visites de l'état i :

$$n_i^p \overset{\text{déf}}{=} \sum_{n=0}^{M^{(p)}-1} 1_{\{Y_n=i\}},$$

Pour l'ensemble des parcelles, soit :

$$n_{ij} \overset{\text{déf}}{=} \sum_{p=1}^{P} n_{ij}^{(p)}, \qquad n_{ij}(k) \overset{\text{déf}}{=} \sum_{p=1}^{P} n_{ij}^{(p)}(k), \qquad n_i \overset{\text{déf}}{=} \sum_{p=1}^{P} n_i^p.$$

Comme au chapitre 5, nous ne considérons pas les temps de séjour sur les états qui apparaissent à la fin des observations. Ainsi la valeur $\bar{H}_{e_{M^{(p)}}^{(p)}}(u_{N_p})$ qui correspond à la fonction de survie du temps de séjour sur l'état final $e_{M^{(p)}}^{(p)}$ sera négligée [7, p. 78].

Nous considérons alors la fonction de vraisemblance :

$$L(Q, f) = \prod_{p=1}^{P} \prod_{k=1}^{N_p} \prod_{i,j \in E} [Q_{ij}]^{n_{ij}^{(p)}} \, [f_{ij}(k)]^{n_{ij}^{(p)}(k)} \tag{7.4}$$

7.3.1 Méthode empirique

Ici, nous reprenons la technique utilisée par Barbu et Limnios [7] pour estimer Q et f.

Définition 7.3.1. *On définit l'estimateur empirique de la matrice de transition de la chaîne incluse Q et de la loi conditionnelle de temps de séjour f par :*

$$\bar{Q}_{ij} \overset{\text{déf}}{=} \frac{n_{ij}}{n_i}, \; i, j \in E, \tag{7.5}$$

$$\bar{f}_{ij}(k) \overset{\text{déf}}{=} \frac{n_{ij}(k)}{n_{ij}}, \; i, j \in E, \; k \in \mathbb{N}. \tag{7.6}$$

Si $n_i = 0$, on pose $\bar{Q}_{ij} = 0, i, j \in E$. De même, on pose $\bar{f}_{ij}(k) = 0$ si $n_{ij} = 0$.

Proposition 7.3.2. *Les estimateurs empiriques définis par* (7.5) *et* (7.6) *maximisent la fonction de vraisemblance* (7.4).

Démonstration. L'estimateur du maximum de vraisemblance est défini par :

$$L(\hat{Q}, \hat{f}) = \max_{Q \in \mathcal{M}_E,\ f \in \mathcal{M}_E(\mathbb{N})} L(Q, f),$$

où $\mathcal{M}_E$ est l'ensemble de matrices carrées $E \times E$ et $\mathcal{M}_E(\mathbb{N})$ l'ensemble des applications de $\mathbb{N}$ dans $\mathcal{M}_E$. Cela revient à résoudre l'équation de la vraisemblance :

$$\begin{aligned} &\frac{\partial L(Q,f)}{\partial Q_{ij}}|_{Q=\hat{Q},\ f=\hat{f}} = 0 \quad \forall i,j \in E, \\ &\frac{\partial L(Q,f)}{\partial f_{ij}(k)}|_{Q=\hat{Q},\ f=\hat{f}} = 0 \quad \forall i,j \in E, \forall k \in \mathbb{N}. \end{aligned}$$

Pour simplifier les calculs, considérons la fonction log–vraisemblance :

$$l(Q,f) = \sum_{p=1}^{P} \sum_{k=1}^{N_p} \sum_{i,j \in E} [n_{ij}^{(p)} \log(Q_{ij}) + n_{ij}^{(p)}(k) \log[f_{ij}(k)]]. \tag{7.7}$$

En utilisant l'égalité :

$$\sum_{j \in E} Q_{ij} = 1, \tag{7.8}$$

l'équation (7.7) peut s'écrire :

$$l(Q,f) = \sum_{p=1}^{P} \sum_{k=1}^{N_p} \sum_{i,j \in E} [n_{ij}^{p} \log(Q_{ij}) + n_{ij}^{p}(k) \log[f_{ij}(k)] + c_i(1 - \sum_{j \in E} Q_{ij})],$$

où c_i est une constante appartenant à $\mathbb{N}^*$.

Si on note :

$$M \overset{\text{déf}}{=} \max_{p=1:P} N_p,$$

alors :

$$l(Q,f) = \sum_{k=1}^{M} \sum_{i,j \in E} [n_{ij} \log(Q_{ij}) + n_{ij}(k) \log[f_{ij}(k)] + c_i(1 - \sum_{j \in E} Q_{ij})]. \tag{7.9}$$

Pour tout $i, j \in E$:

$$\frac{\partial l}{\partial Q_{ij}} = 0 \qquad \text{donne} \qquad \hat{Q}_{ij} = \frac{n_{ij}}{c_i}.$$

L'équation (7.8) devient :

$$1 = \sum_{j \in E} \hat{Q}_{ij} = \sum_{j \in E} \frac{n_{ij}}{c_i} = \frac{n_i}{c_i}.$$

On obtient $c_i = n_i$ et :

$$\hat{Q}_{ij} = \frac{n_{ij}}{n_i},$$

qui n'est autre que l'estimateur empirique de Q_{ij} dans (7.5).

Pour prouver que $\bar{f}$ maximise aussi (7.7), on refait le même calcul en utilisant l'égalité :

$$\sum_{k=1}^{\infty} f_{ij}(k) = 1. \tag{7.10}$$

On a alors :

$$l(Q, f) = \sum_{k=1}^{M} \sum_{i,j \in E} \left[n_{ij} \log(Q_{ij}) + n_{ij}(k) \log[f_{ij}(k)] + c_{ij}(1 - \sum_{k=1}^{\infty} f_{ij}(k) = 1) \right], \tag{7.11}$$

avec c_{ij} une constante dans $\mathbb{N}^*$.

La maximisation de (7.11) suivant $f_{ij}(k),\ i, j \in E, k > 0$ donne :

$$\hat{f}_{ij}(k) = \frac{n_{ij}(k)}{c_{ij}}.$$

En appliquant (7.10), on obtient $c_{ij} = n_{ij}$ et on retrouve :

$$\hat{f}_{ij}(k) = \frac{n_{ij}(k)}{n_{ij}}\, i, j \in E, k > 0.$$

□

Les estimateurs que nous avons calculés sont des estimateurs non paramétriques : $\hat{Q}$ et $\hat{f}$ sont calculés directement à partir du jeu de données sans tenir compte des hypothèses sur les lois de paramètres θ, γ et λ énoncées dans (h_3) et (h_4). Les estimateurs que nous allons voir dans les sections suivantes sont des estimateurs paramétriques. Les estimations de Q et f se feront par les estimations des paramètres qui les caractérisent.

7.3.2 Méthode du maximum du vraisemblance

En tenant compte les expressions de Q et f dans (7.1) et (7.2), l'expression de la fonction de vraisemblance définie dans (7.4) peut s'écrire :

$$\begin{aligned} L(\theta,\gamma,\lambda) = \prod_{k=1}^{M} & \theta_1^{n_{\texttt{CaJc}}}\,\theta_2^{n_{\texttt{CaHe}}}\,(1-\theta_1-\theta_2)^{n_{\texttt{CaCp}}}\,\theta_3^{n_{\texttt{JcCa}}}\,(1-\theta_3)^{n_{\texttt{JcFs}}}\,\theta_4^{n_{\texttt{HeCa}}}\,(1-\theta_4)^{n_{\texttt{HeJc}}} \\ & \Bigg[\Big(\gamma_1\,(1-\gamma_1)^{k-1}\Big)^{n_{\texttt{CaJc}}(k)}\,\Big(\gamma_2\,(1-\gamma_2)^{k-1}\Big)^{n_{\texttt{CaHe}}(k)}\,\Big(\gamma_3\,(1-\gamma_3)^{k-1}\Big)^{n_{\texttt{CaCp}}(k)} \\ & \Big(\gamma_4\,(1-\gamma_4)^{k-1}\Big)^{n_{\texttt{JcCa}}(k)}\,\Big(\frac{e^{-\lambda_1}\,\lambda_1^k}{k!}\Big)^{n_{\texttt{JcFs}}(k)}\,\Big(\frac{e^{-\lambda_2}\,\lambda_2^k}{k!}\Big)^{n_{\texttt{HeCa}}(k)} \\ & \Big(\frac{e^{-\lambda_3}\,\lambda_3^k}{k!}\Big)^{n_{\texttt{HeJc}}(k)}\,\Big(\frac{e^{-\lambda_4}\,\lambda_4^k}{k!}\Big)^{n_{\texttt{FsCa}}(k)}\,\Big(\frac{e^{-\lambda_5}\,\lambda_5^k}{k!}\Big)^{n_{\texttt{CpCa}}(k)}\Bigg] \end{aligned} \tag{7.12}$$

On peut la décomposer de la façon suivante :

$$L(\theta,\gamma,\lambda) = L_1(\theta)\,L_2(\gamma)\,L_2(\lambda)$$

où :

$$L_1(\theta) = \theta_1^{n_{\texttt{CaJc}}}\,\theta_2^{n_{\texttt{CaHe}}}\,(1-\theta_1-\theta_2)^{n_{\texttt{CaCp}}}\,\theta_3^{n_{\texttt{JcCa}}}\,(1-\theta_3)^{n_{\texttt{JcFs}}}\,\theta_4^{n_{\texttt{HeCa}}}\,(1-\theta_4)^{n_{\texttt{HeJc}}},$$

$$\begin{aligned} L_2(\gamma) = \prod_{k=1}^{M} & \Bigg[\Big(\gamma_1\,(1-\gamma_1)^{k-1}\Big)^{n_{\texttt{CaJc}}(k)}\,\Big(\gamma_2\,(1-\gamma_2)^{k-1}\Big)^{n_{\texttt{CaHe}}(k)} \\ & \Big(\gamma_3\,(1-\gamma_3)^{k-1}\Big)^{n_{\texttt{CaCp}}(k)}\,(\gamma_4\,(1-\gamma_4)^{k-1})^{n_{\texttt{JcCa}}(k)}\Bigg] \end{aligned}$$

et

$$\begin{aligned} L_3(\lambda) = \prod_{k=1}^{M} & \Bigg[\Big(\frac{e^{-\lambda_1}\,\lambda_1^k}{k!}\Big)^{n_{\texttt{JcFs}}(k)}\,\Big(\frac{e^{-\lambda_2}\,\lambda_2^k}{k!}\Big)^{n_{\texttt{HeCa}}(k)} \\ & \Big(\frac{e^{-\lambda_3}\,\lambda_3^k}{k!}\Big)^{n_{\texttt{HeJc}}(k)}\,\Big(\frac{e^{-\lambda_4}\,\lambda_4^k}{k!}\Big)^{n_{\texttt{FsCa}}(k)}\,\Big(\frac{e^{-\lambda_5}\,\lambda_5^k}{k!}\Big)^{n_{\texttt{CpCa}}(k)}\Bigg]. \end{aligned}$$

Soit :

$$l(\theta,\gamma,\lambda) = \log L(\theta,\gamma,\lambda) = l_1(\theta) + l_2(\gamma) + l_3(\lambda)$$

la fonction log–vraisemblance, avec :

$$
\begin{aligned}
l_1(\theta) &= n_{\texttt{CaJc}} \log(\theta_1) + n_{\texttt{CaHe}} \log(\theta_2) + n_{\texttt{CaCp}} \log(1-\theta_1-\theta_2) \\
&\quad + n_{\texttt{JcCa}} \log(\theta_3) + n_{\texttt{JcFs}} \log(1-\theta_3) + n_{\texttt{HeCa}} \log(\theta_4) + n_{\texttt{HeJc}} \log(1-\theta_4), \\
l_2(\gamma) &= \sum_{k=1}^{M} \big(n_{\texttt{CaJc}}(k)\ \log(\gamma_1) + (k-1)\, n_{\texttt{CaJc}}(k)\ \log(1-\gamma_1)\big) \\
&\quad + \sum_{k=1}^{M} \big(n_{\texttt{CaHe}}(k)\ \log(\gamma_2) + (k-1)\, n_{\texttt{CaHe}}(k) \log(1-\gamma_2)\big) \\
&\quad + \sum_{k=1}^{M} \big(n_{\texttt{CaCp}}(k)\ \log(\gamma_3) + (k-1)\, n_{\texttt{CaCp}}(k)\ \log(1-\gamma_3)\big) \\
&\quad + \sum_{k=1}^{M} \big(n_{\texttt{JcCa}}(k)\ \log(\gamma_4) + (k-1)\, n_{\texttt{JcCa}}(k) \log(1-\gamma_4)\big)
\end{aligned}
$$

et

$$
\begin{aligned}
l_3(\lambda) &= \sum_{k=1}^{M} \big(-\lambda_1\, n_{\texttt{JcFs}}(k) + k\, n_{\texttt{JcFs}}(k) \log(\lambda_1) - \log(k!)\big) \\
&\quad + \sum_{k=1}^{M} \big(-\lambda_2\, n_{\texttt{HeCa}}(k) + k\, n_{\texttt{HeCa}}(k) \log(\lambda_2) - \log(k!)\big) \\
&\quad + \sum_{k=1}^{M} \big(-\lambda_3\, n_{\texttt{HeJc}}(k) + k\, n_{\texttt{HeJc}}(k) \log(\lambda_3) - \log(k!)\big). \\
&\quad + \sum_{k=1}^{M} \big(-\lambda_4\, n_{\texttt{FsCa}}(k) + k\, n_{\texttt{FsCa}}(k) \log(\lambda_4) - \log(k!)\big). \\
&\quad + \sum_{k=1}^{M} \big(-\lambda_5\, n_{\texttt{CpCa}}(k) + k\, n_{\texttt{CpCa}}(k) \log(\lambda_5) - \log(k!)\big).
\end{aligned}
$$

Matrice de transtion de la chaîne incluse

L'estimateur du maximum de vraisemblance (EMV) $\hat{\theta}$ du paramètre θ est obtenu en maximisant $l_1(\theta)$, c'est-à-dire :

$$
l_1(\hat{\theta}) = \max_{\theta \in \Theta} l_1(\theta).
$$

On obtient $\hat{\theta}$ par la résolution de :

$$
\frac{\partial l_1(\theta)}{\partial \theta}\Big|_{\theta=\hat{\theta}} = 0,
$$

avec :

$$\begin{aligned}
\frac{\partial l_1(\theta)}{\partial \theta_1} &= -\frac{n_{\texttt{CaCp}}}{(1-\theta_1-\theta_2)} + \frac{n_{\texttt{CaJc}}}{\theta_1}, \\
\frac{\partial l_1(\theta)}{\partial \theta_2} &= \frac{n_{\texttt{CaCp}}}{(1-\theta_1-\theta_2)} + \frac{n_{\texttt{CaHe}}}{\theta_2}, \\
\frac{\partial l_1(\theta)}{\partial \theta_3} &= \frac{n_{\texttt{JcFs}}}{(1-\theta_3)} + \frac{n_{\texttt{JcCa}}}{\theta_3}, \\
\frac{\partial l_1(\theta)}{\partial \theta_4} &= \frac{n_{\texttt{HeJc}}}{(1-\theta_4)} + \frac{n_{\texttt{HeCa}}}{\theta_4}.
\end{aligned}$$

On obtient :

$$\begin{aligned}
\hat{\theta}_1 &= \frac{n_{\texttt{CaJc}}}{n_{\texttt{CaJc}} + n_{\texttt{CaHe}} + n_{\texttt{CaCp}}}, \\
\hat{\theta}_2 &= \frac{n_{\texttt{CaHe}}}{n_{\texttt{CaJc}} + n_{\texttt{CaHe}} + n_{\texttt{CaCp}}}, \\
\hat{\theta}_3 &= \frac{n_{\texttt{JcCa}}}{n_{\texttt{JcCa}} + n_{\texttt{JcFs}}}, \\
\hat{\theta}_4 &= \frac{n_{\texttt{HeCa}}}{n_{\texttt{HeCa}} + n_{\texttt{HeJc}}}.
\end{aligned} \tag{7.13}$$

Lois des temps de séjour

La loi conditionnelle des temps de séjour f dépend des paramètres $\gamma = (\gamma_i)_{1\leq i\leq 4}$ et $\lambda = (\lambda_j)_{1\leq j\leq 5}$, cf. équation (7.2). Les EMVs $\hat{\gamma}$ et $\hat{\lambda}$ sont définis par :

$$l_2(\hat{\gamma}) = \max_{\gamma\in\Gamma} l_2(\gamma), \qquad l_3(\hat{\lambda}) = \max_{\lambda\in\Lambda} l_3(\lambda).$$

On calcule $\hat{\gamma}$ en résolvant :

$$\frac{\partial l_2(\gamma)}{\partial \gamma}|_{\gamma=\hat{\gamma}} = 0$$

avec :

$$\begin{aligned}
\frac{\partial l_2(\theta)}{\partial \gamma_1} &= \sum_{k=1}^{M} \frac{n_{\texttt{CaJc}}(k)}{\gamma_1} - \frac{(k-1)\, n_{\texttt{CaJc}}(k)}{(1-\gamma_1)}, \\
\frac{\partial l_2(\theta)}{\partial \gamma_2} &= \sum_{k=1}^{M} \frac{n_{\texttt{CaHe}}(k)}{\gamma_2} - \frac{(k-1)\, n_{\texttt{CaHe}}(k)}{(1-\gamma_2)}, \\
\frac{\partial l_2(\theta)}{\partial \gamma_3} &= \sum_{k=1}^{M} \frac{n_{\texttt{CaCp}}(k)}{\gamma_3} - \frac{(k-1)\, n_{\texttt{CaCp}}(k)}{(1-\gamma_3)}, \\
\frac{\partial l_2(\theta)}{\partial \gamma_4} &= \sum_{k=1}^{M} \frac{n_{\texttt{JcCa}}(k)}{\gamma_4} - \frac{(k-1)\, n_{\texttt{JcCa}}(k)}{(1-\gamma_4)}.
\end{aligned}$$

Après calcul, on trouve :

$$\hat{\gamma_1} = \frac{\sum_{k=1}^{M} n_{\texttt{CaJc}}(k)}{\sum_{k=1}^{M} k\, n_{\texttt{CaJc}}(k)}, \qquad \hat{\gamma_2} = \frac{\sum_{k=1}^{M} n_{\texttt{CaHe}}(k)}{\sum_{k=1}^{M} k\, n_{\texttt{CaCp}}(k)},$$

$$\hat{\gamma_3} = \frac{\sum_{k=1}^{M} n_{\texttt{CaCp}}(k)}{\sum_{k=1}^{M} k\, n_{\texttt{CaCp}}(k)}, \qquad \hat{\gamma_4} = \frac{\sum_{k=1}^{M} n_{\texttt{JcCa}}(k)}{\sum_{k=1}^{M} k\, n_{\texttt{JcCa}}(k)}.$$

Enfin pour calculer l' EMV de λ, on résout :

$$\frac{\partial l_3(\lambda)}{\partial \lambda}|_{\lambda=\hat{\lambda}} = 0$$

avec :

$$\frac{\partial l_3(\theta)}{\partial \lambda_1} = \sum_{k=1}^{M} -\lambda_1\, n_{\texttt{JcFs}}(k) + \frac{k\, n_{\texttt{JcFs}}(k)}{\lambda_1},$$

$$\frac{\partial l_3(\theta)}{\partial \lambda_2} = \sum_{k=1}^{M} -\lambda_2\, n_{\texttt{HeCa}}(k) + \frac{k\, n_{\texttt{HeCa}}(k)}{\lambda_2},$$

$$\frac{\partial l_3(\theta)}{\partial \lambda_3} = \sum_{k=1}^{M} -\lambda_3\, n_{\texttt{HeJc}}(k) + \frac{k\, n_{\texttt{HeJc}}(k)}{\lambda_3},$$

$$\frac{\partial l_3(\theta)}{\partial \lambda_4} = \sum_{k=1}^{M} -\lambda_4\, n_{\texttt{FsCa}}(k) + \frac{k\, n_{\texttt{FsCa}}(k)}{\lambda_4},$$

$$\frac{\partial l_3(\theta)}{\partial \lambda_5} = \sum_{k=1}^{M} -\lambda_5\, n_{\texttt{CpCa}}(k) + \frac{k\, n_{\texttt{CpCa}}(k)}{\lambda_5}.$$

On obtient :

$$\hat{\lambda}_1 = \frac{\sum_{k=1}^{M} k\, n_{\texttt{JcFs}}(k)}{\sum_{k=1}^{M} n_{\texttt{JcFs}}(k)}, \qquad \hat{\lambda}_2 = \frac{\sum_{k=1}^{M} k\, n_{\texttt{HeCa}}(k)}{\sum_{k=1}^{M} n_{\texttt{HeCa}}(k)},$$

$$\hat{\lambda}_3 = \frac{\sum_{k=1}^{M} k\, n_{\texttt{HeJc}}(k)}{\sum_{k=1}^{M} n_{\texttt{HeJc}}(k)}, \qquad \hat{\lambda}_4 = \frac{\sum_{k=1}^{M} k\, n_{\texttt{FsCa}}(k)}{\sum_{k=1}^{M} n_{\texttt{FsCa}}(k)},$$

$$\hat{\lambda}_5 = \frac{\sum_{k=1}^{M} k\, n_{\texttt{CpCa}}(k)}{\sum_{k=1}^{M} n_{\texttt{CpCa}}(k)}.$$

7.3.3 Méthode bayésienne

Cette partie concerne l'estimation de la matrice de transition Q de la chaîne incluse $(Y_k)_{k\in\mathbb{N}}$ ainsi que de la loi conditionnelle des temps de séjour f par l'approche bayésienne. Dans le chapitre 4 nous avons déjà introduit cette méthode pour estimer la matrice de transition d'un modèle markovien. En premier lieu, nous reprenons cette technique pour estimer la matrice de transition de la chaîne incluse $(Y_k)_{k\in\mathbb{N}}$. Puis, nous calculons les estimateurs bayésiens correspondants aux paramètres de la loi des temps de séjour. Dans les deux cas, nous calculons les estimateurs avec la méthode de Monte Carlo par chaîne de Markov (MCMC).

Estimation de la matrice de transition de la chaîne incluse

On suppose que θ dans la matrice de transition Q dans (7.1) est une variable aléatoire selon une loi a priori de densité $\pi_{\text{prior}}(\theta)$.

Loi a priori $\pi_{\text{prior}}(\theta)$

Nous considérons la loi a priori non informative de Jeffreys [32] :

$$\pi_{\text{prior}}(\theta) \propto \sqrt{\det[\mathcal{I}(\theta)]}, \tag{7.14}$$

où

$$\mathcal{I}(\theta) = \left[\mathbb{E}_\theta\Big(-\frac{\partial^2 l_1(\theta)}{\partial\theta_k\,\partial\theta_l}\Big)\right]_{1\le k,l\le 4}$$

est la matrice d'information de Fisher.

On peut décomposer $\mathcal{I}(\theta)$:

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

avec :

$$A_{1,2} \overset{\text{déf}}{=} -\mathbb{E}_\theta \begin{pmatrix} \frac{\partial^2 l_1(\theta)}{\partial^2\theta_1} & \frac{\partial^2 l_1(\theta)}{\partial\theta_1\partial\theta_2} \\ \frac{\partial^2 l_1(\theta)}{\partial\theta_2\partial\theta_1} & \frac{\partial^2 l_1(\theta)}{\partial^2\theta_2} \end{pmatrix}$$

et

$$a_k \overset{\text{déf}}{=} -\mathbb{E}_\theta\left(\frac{\partial^2 l_1(\theta)}{\partial^2\theta_k}\right),\ k = 3, 4.$$

Le calcul du déterminant donne :

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

avec :

$$\begin{aligned} \det A_{1,2} &= \left(\tfrac{\mathbb{E}_\theta[n_{\texttt{CaCp}}]}{(1-\theta_1-\theta_2)^2} + \tfrac{\mathbb{E}_\theta[n_{\texttt{CaJc}}]}{\theta_1^2}\right)\left(\tfrac{\mathbb{E}_\theta[n_{\texttt{CaCp}}]}{(1-\theta_1-\theta_2)^2} + \tfrac{\mathbb{E}_\theta[n_{\texttt{CaHe}}]}{\theta_2^2}\right) - \left(\tfrac{\mathbb{E}_\theta[n_{\texttt{CaCp}}]}{(1-\theta_1-\theta_2)^2}\right)^2, \\ a_3 &= \tfrac{\mathbb{E}_\theta[n_{\texttt{JcFs}}]}{(1-\theta_3)^2} + \tfrac{\mathbb{E}_\theta[n_{\texttt{JcCa}}]}{\theta_3^2}, \\ a_4 &= \tfrac{\mathbb{E}_\theta[n_{\texttt{HeJc}}]}{(1-\theta_4)^2} + \tfrac{\mathbb{E}_\theta[n_{\texttt{HeCa}}]}{\theta_4^2}. \end{aligned}$$

Pour deux états i et j de E, l'espérance du nombre des transitions $\mathbb{E}_\theta[n_{ij}]$ se calcule par :

$$\begin{aligned}
\mathbb{E}_\theta[n_{ij}] &= \sum_{p=1}^{P} \sum_{k=2}^{M^{(p)}} \mathbb{P}_\theta(Y_k^{(p)} = j, Y_{k-1}^{(p)} = i) \\
&= \sum_{p=1}^{P} \sum_{k=2}^{M^{(p)}} \mathbb{P}_\theta(Y_k^{(p)} = j | Y_{k-1}^{(p)} = i)\, \mathbb{P}_\theta(Y_{k-1} = i) \\
&= Q_{ij} \sum_{p=1}^{P} \sum_{k=2}^{M^{(p)}} \mathbb{P}_\theta(Y_{k-1}^{(p)} = i) \\
&= Q_{ij} \sum_{p=1}^{P} \sum_{k=2}^{M^{(p)}} \left(\delta_{\mathtt{Ca}} Q^{k-1}\right)_i = Q_{ij} \sum_{p=1}^{P} \sum_{k=2}^{M^{(p)}} [Q^{k-1}]_{\mathtt{Ca}\, j}. \qquad (7.15)
\end{aligned}$$

Loi a posteriori

La loi a posteriori $\pi_{\text{post}}(\theta)$ s'obtient avec la formule de Bayes :

$$\pi_{\text{post}}(\theta) \propto L_1(\theta)\, \pi_{\text{prior}}(\theta), \qquad (7.16)$$

où

$$L_1(\theta) = \theta_1^{n_{\mathtt{CaJc}}}\, \theta_2^{n_{\mathtt{CaHe}}}\, (1-\theta_1-\theta_2)^{n_{\mathtt{CaCp}}}\, \theta_3^{n_{\mathtt{JcCa}}}\, (1-\theta_3)^{n_{\mathtt{JcFs}}}\, \theta_4^{n_{\mathtt{HeCa}}}\, (1-\theta_4)^{n_{\mathtt{HeJc}}}$$

est la fonction de vraisemblance associée à θ. L'estimateur bayésien $\tilde{\theta}$ est la moyenne de la loi a posteriori :

$$\tilde{\theta} \overset{\text{déf}}{=} \int_\Theta \theta\, \pi_{\text{post}}(\theta)\, \mathrm{d}\theta = \frac{\int_\Theta \theta\, L_1(\theta)\, \pi_{\text{prior}}(\theta)\, \mathrm{d}\theta}{\int_\Theta L_1(\theta)\, \pi_{\text{prior}}(\theta)\, \mathrm{d}\theta}\,. \qquad (7.17)$$

Le calcul de la loi a posteriori (7.17) ne peut pas se faire de façon explicite et nous faisons appel à une méthode de Monte Carlo par chaîne de Markov (MCMC). Nous utilisons l'algorithme de Metropolis-Hastings à noyau gaussien décrit par l'algorithme 7.1.

Estimation des lois de temps de séjour

La loi conditionnelle des temps de séjour (7.2) dépend des paramètres $\gamma = (\gamma_i)_{1\leq i\leq 4}$ des lois géométriques et $\lambda = (\gamma_i)_{1\leq i\leq 5}$ des lois de Poisson. Nous faisons l'hypothèse que les lois de γ et λ admettent comme densités $\pi^g_{\text{prior}}(\gamma)$ et $\pi^p_{\text{prior}}(\lambda)$.

Lois a priori

Comme dans la section précédente, nous utilisons la loi a priori non informative de Jeffreys.

---

choisir θ
$\tilde{\theta} \leftarrow \theta$
pour $k = 2, 3, \ldots$ **faire**
  $\varepsilon \sim \mathcal{N}(0, \sigma^2)$ % loi gaussienne
  $\theta^{\text{prop}} \leftarrow \theta + \varepsilon$
  $u \sim U[0, 1]$
  $\alpha \leftarrow \min\left\{1, \dfrac{\pi_{\text{post}}(\theta^{\text{prop}})\ g(\theta - \theta^{\text{prop}})}{\pi_{\text{post}}(\theta)\ g(\theta^{\text{prop}} - \theta)}\right\}$
  si $u \leq \alpha$ **alors**
    $\theta \leftarrow \theta^{\text{prop}}$
  fin si
  $\tilde{\theta} \leftarrow \dfrac{k-1}{k}\,\tilde{\theta} + \dfrac{1}{k}\,\theta$
fin pour

Alg. 7.1: *Algorithme de Metropolis-Hastings à tirages indépendants pour échantillonner la loi a posteriori du paramètre de la matrice de transition Q de la chaîne incluse $(Y_k)_{k\in\mathbb{N}}$.*

Proposition 7.3.3. (i) *Considérons un échantillon de taille n d'une loi géométrique de paramètre γ ; la loi a priori de Jeffreys s'obtient par :*

$$\pi^g_{\text{prior}}(\gamma) \propto \sqrt{\frac{n}{\gamma^2}\left(\frac{1}{(1-\gamma)}\right)}. \tag{7.18}$$

(ii) *Considérons un échantillon de taille n d'une loi de Poisson de paramètre λ ; la loi a priori de Jeffreys est :*

$$\pi^p_{\text{prior}}(\lambda) \propto \sqrt{\frac{n}{\lambda}}. \tag{7.19}$$

Démonstration. (i) Si $S_1, \ldots, S_n$ est un échantillon d'une loi géométrique de paramètre γ, la fonction de vraisemblance s'écrit :

$$L_g(\gamma) = (1-\gamma)^{S_1-1}\gamma \ldots (1-\gamma)^{S_n-1}\gamma = (1-\gamma)^{\sum_{k=1}^n S_k - n}\ \gamma^n$$

et la fonction log-vraisemblance est :

$$l_g(\gamma) = \sum_{k=1}^n S_k - n\ \log(1-\gamma) + n \log \gamma\,.$$

Alors, la matrice d'information de Fisher est :

$$\mathcal{I}(\gamma) = \mathbb{E}_\gamma\left(-\frac{\partial^2 l_g(\gamma)}{\partial^2 \gamma}\right) = -\left(-\frac{\sum_{k=1}^n \mathbb{E}_\gamma(S_k) - n}{(1-\gamma)^2} + \frac{n}{\gamma^2}\right).$$

Comme

$$\mathbb{E}_\gamma(S_k) = \frac{1}{\gamma},\ k = 1, \ldots, n,$$

alors $\mathcal{I}(\gamma)$ devient :

$$\mathcal{I}(\gamma) = \frac{n(1-\gamma)}{\gamma(1-\gamma)^2} + \frac{n}{\gamma^2} = \frac{n}{\gamma^2}\left(\frac{1}{(1-\gamma)}\right).$$

On obtient (7.18) par la définition de la loi a priori de Jeffreys :

$$\pi^g_{\text{prior}}(\gamma) \propto \sqrt{\det[\mathcal{I}(\gamma)]}\,.$$

(*ii*) La fonction de vraisemblance associée à l'échantillon $S_1, \ldots, S_k$ d'une loi de Poisson s'écrit :

$$L_p(\lambda) = \prod_{k=1}^{n} e^{-\lambda}\frac{\lambda^{S_k}}{S_k!} = e^{-\lambda n}\prod_{k=1}^{n}\frac{\lambda^{S_k}}{S_k!}\,.$$

Sa fonction log-vraisemblance est

$$l_p = -\lambda n + \log(\lambda)\sum_{k=1}^{n} S_k - \sum_{k=1}^{n}\log(S_k)$$

et la matrice de Fisher est :

$$\mathcal{I}(\lambda) = \mathbb{E}_\lambda\left(-\frac{\partial^2 l_p(\lambda)}{\partial^2\lambda}\right) = \frac{1}{\lambda^2}\sum_{k=1}^{n}\mathbb{E}_\lambda(S_k) = \frac{\lambda\, n}{\lambda^2} = \frac{n}{\lambda}$$

car $\mathbb{E}_\lambda(S_k) = \lambda,\ k = 1, \ldots, n$. On trouve (7.19) à l'aide de

$$\pi^p_{\text{prior}}(\lambda) \propto \sqrt{\det[\mathcal{I}(\lambda)]}.$$

Lois a posteriori

Les lois a posteriori de γ et λ se calculent à l'aide des lois a priori et des fonctions de vraisemblance :

$$\pi^g_{\text{post}}(\gamma) \propto L_g(\gamma)\,\pi^g_{\text{prior}}(\gamma)\,, \qquad \pi^p_{\text{post}}(\lambda) \propto L_p(\lambda)\,\pi^p_{\text{prior}}(\lambda)\,.$$

Les estimateurs bayésiens sont :

$$\tilde{\gamma} \overset{\text{déf}}{=} \int \gamma\,\pi^g_{\text{post}}(\gamma)\,\mathrm{d}\gamma\,, \qquad \tilde{\lambda} \overset{\text{déf}}{=} \int \lambda\,\pi^p_{\text{post}}(\lambda)\,\mathrm{d}\lambda\,. \tag{7.20}$$

Les algorithmes 7.2 et 7.3 permettent de calculer des approximations de $\tilde{\gamma}$ et $\tilde{\lambda}$ par une méthode MCMC.

---

choisir γ
$\hat{\gamma} \leftarrow \gamma$
pour $k = 2, 3, 4, \ldots$ **faire**
  $\gamma_{\text{prop}} \sim U[0,1]$ % loi uniforme sur $[0,1]$
  $u \sim U[0,1]$
  $\alpha \leftarrow \min\left\{1, \dfrac{\pi^g_{\text{post}}(\gamma_{\text{prop}})}{\pi^g_{\text{post}}(\gamma)}\right\}$
  si $u \leq \alpha$ **alors**
    $\theta \leftarrow \gamma_{\text{prop}}$ % acceptation
  fin si
  $\hat{\gamma} \leftarrow \dfrac{k-1}{k}\,\hat{\gamma} + \dfrac{1}{k}\,\gamma$
fin pour

Alg. 7.2: *Algorithme de Metropolis-Hastings pour échantillonner selon* $\pi^g_{\text{post}}(\gamma)$ *et calculer une approximation de* $\tilde{\gamma}$ *par la moyenne de l'échantillon.*

---

choisir λ
$\hat{\lambda} \leftarrow \lambda$
pour $k = 2, 3, 4, \ldots$ **faire**
  $\lambda_{\text{prop}} \sim U([1,n])$ % loi discrète uniforme sur $[1,n]$
  , % où n est la longueur d'observation.
  $u \sim U[0,1]$
  $\alpha \leftarrow \min\left\{1, \dfrac{\pi^p_{\text{post}}(\lambda_{\text{prop}})}{\pi^p_{\text{post}}(\lambda)}\right\}$
  si $u \leq \alpha$ **alors**
    $\lambda \leftarrow \lambda_{\text{prop}}$ % acceptation
  fin si
  $\hat{\lambda} \leftarrow \dfrac{k-1}{k}\,\hat{\lambda} + \dfrac{1}{k}\,\lambda$
fin pour

Alg. 7.3: *Algorithme de Metropolis-Hastings pour échantillonner selon* $\pi^p_{\text{post}}(\lambda)$ *et calculer une approximation de* $\tilde{\lambda}$.

Remarque 7.3.4. *Pour la loi de Poisson, le calcul de* (7.20) *peut se faire de manière explicite à l'aide de la loi Gamma. En utilisant l'expression de la loi a priori de Jeffreys dans* (7.19)*, la loi a posteriori s'écrit :*

$$\begin{aligned}\pi^p_{\text{post}}(\lambda) &\propto \sqrt{\frac{n}{\lambda}}\,\lambda^{\sum_{k=1}^n S_k} e^{-n\lambda}\\ &\propto \lambda^{(\sum_{k=1}^n S_k - \frac{1}{2})}\; e^{-n\lambda}\\ &\propto g(\lambda, \sum_{k=1}^n S_k + \frac{1}{2}, n),\end{aligned}$$

où

$$g(\lambda, \alpha, \beta) \overset{\text{déf}}{=} \beta^\alpha \frac{1}{\Gamma(\alpha)} \lambda^{\alpha-1}\; e^{-\beta\lambda}$$

est la densité d'une loi gamma de paramètre (α, β)*. Donc,* $\lambda \overset{iid}{\sim} \Gamma(\sum_{k=1}^n S_k + \frac{1}{2}, n)$ *et la moyenne est :*

$$\hat{\lambda} = \frac{\sum_{k=1}^n S_k + \frac{1}{2}}{n}. \tag{7.21}$$

Dans la pratique, les deux méthodes coïncident car les résultats donnés par l'algorithme 7.3 convergent vers la valeur calculée par (7.21)*.*

7.4 Résultats et discussions

Les trois méthodes d'estimation proposées dans la section précédente conduisent à des modèles semi-markoviens différents. L'estimateur empirique et l'estimateur du maximum de vraisemblance conduisent à des modèles absorbants dont l'état absorbant est l'état "culture pérenne" (`Cp`), voir figure 7.5. Avec la méthode bayésienne, le modèle est une chaîne semi-markovienne régulière. Nous étudions le comportement de ces modèles dans la suite.

7.4.1 Modèles absorbants

Lois de transition semi-markovienne

L'estimateur empirique (7.5) et l'estimateur du maximum de vraisemblance (7.13) de la matrice de transition Q coïncident :

$$\bar{Q} = \hat{Q} = \begin{pmatrix} 0 & 0.8267 & 0.0400 & 0 & 0.1333 \\ 0.9275 & 0 & 0 & 0.0725 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{7.22}$$

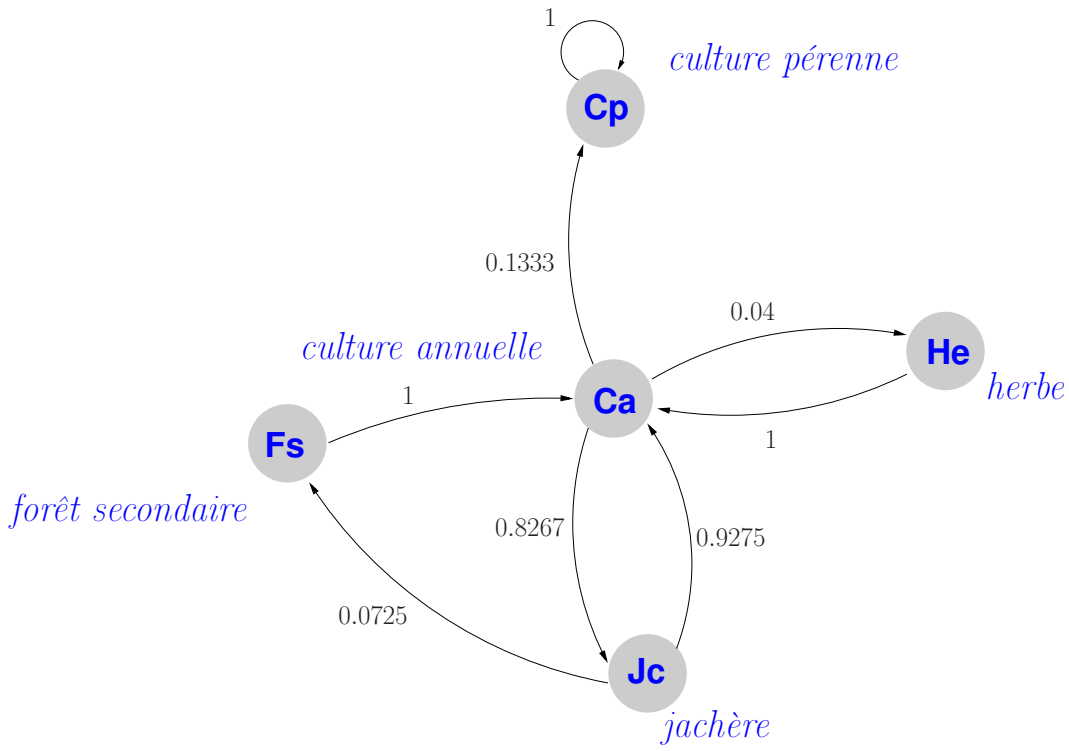

Figure 7.5: *Chaîne de Markov incluse associée aux modèles déduits de la méthode empirique et de la méthode du maximum de vraisemblance.*

En effet ces deux estimateurs se basent sur le nombre des transitions observées entre les états. Avec ce modèle, l'état (`Cp`) est un état absorbant car il n y a aucune transition partant de cet état dans le jeu de données.

Il y a une légère différence entre les deux méthodes au niveau de l'estimation de la loi conditionnelle des temps de séjour, voir la figure 7.6. L'estimateur empirique $\bar{f}$ s'obtient directement par les fréquences d'occurrence des temps de séjours observés (7.6). L'estimateur du maximum de vraisemblance $\hat{f}$ repose sur l'hypothèse (h_4) et estime les paramètres des lois à l'aide des observations. Toutefois, les résultats sont cohérents car :

(i) si $(S_1, S_2, \ldots, S_n)$ est un échantillon d'une loi géométrique de paramètre γ, l'estimateur du maximum de vraisemblance de γ est

$$\frac{1}{m_s}$$

où

$$m_s = \frac{\sum_{i=0}^{n} S_i}{n}$$

est la moyenne empirique correspondant à l'échantillon. Ceci est vérifié pour $\hat{\gamma}_i,\ i = 1, 2, 3, 4$.

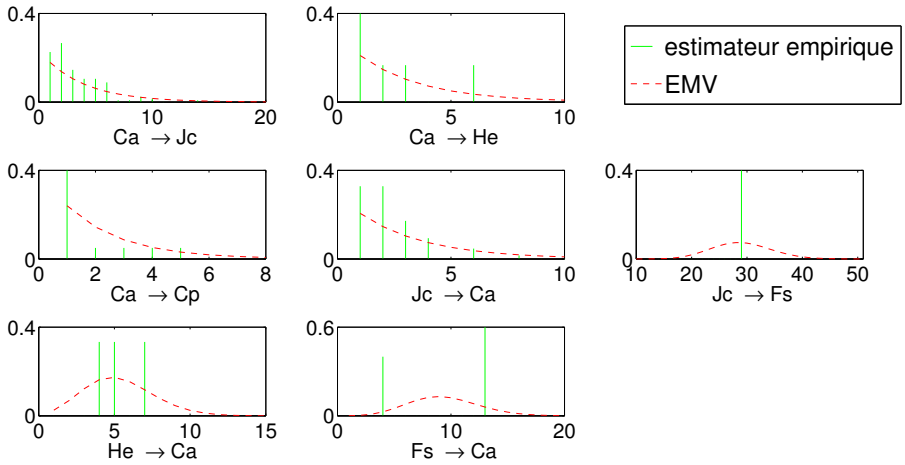

Figure 7.6: *Lois des temps de séjour : estimateur empirique et estimateur du maximum de vraisemblance des lois conditionnelles des temps de séjour* $f_{ij}(k)$, $i, j \in E$, $k \in \mathbb{N}$. *L'estimateur du maximum de vraisemblance estime les paramètres des lois et l'estimateur empirique compte les fréquences d'occurrence des temps de séjour selon les observations.*

(*ii*) si $(S_1, S_2, \ldots, S_n)$ est un échantillon d'une loi de Poisson de paramètre λ, l'estimateur du maximum de vraisemblance de λ n'est autre que la moyenne empirique, ce qui est le cas pour $\hat{\lambda}_i$, $i = 1, 2, 3$.

Dans les deux cas, le temps de séjour sur l'état absorbant "culture pérenne" (`Cp`) est infini. Ainsi, les deux lois supposées (géométrique et Poisson) modélisent correctement les fréquences empiriques des temps de séjour.

En utilisant les EMVs de Q et f, nous obtenons les estimations du noyau semi-markovien :

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

Puis, pour tout $k \in \mathbb{N}$, $i, j \in E$, les estimateurs de $h(k)$ et $H(k)$ se calcule facilement par les relations (6.6) et (6.7) :

$$\bar{h}_i(k) = \sum_{j \in E} \bar{q}_{ij}(k), \qquad \hat{h}_i(k) = \sum_{j \in E} \hat{q}_{ij}(k),$$

$$\bar{H}_i(k) = \sum_{l=1}^{k} \bar{h}_i(l), \qquad \hat{H}_i(k) = \sum_{l=1}^{k} \hat{h}_i(l).$$

On en déduit les estimateurs de la loi de transition semi-markovienne P par

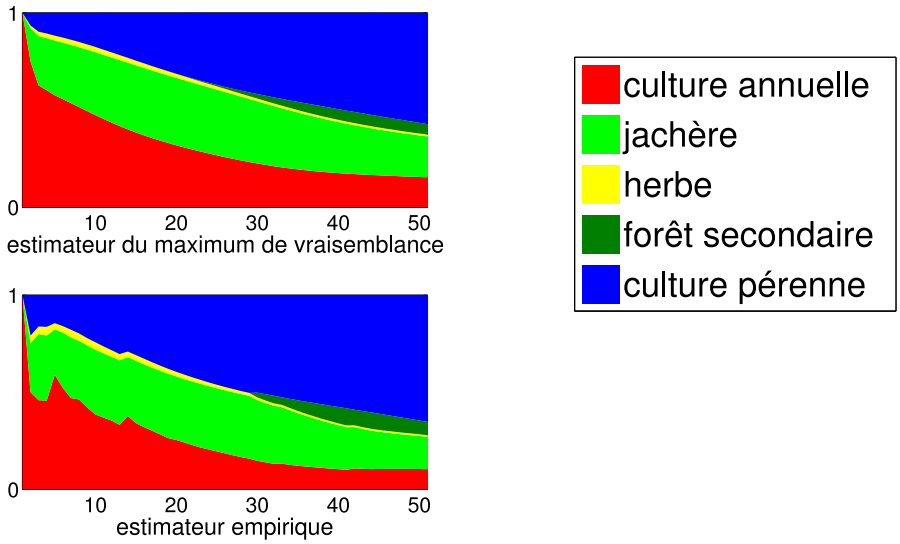

Figure 7.7: *Estimateur empirique et estimateur du maximum de vraisemblance des proportions des parcelles occupées par les états* `Ca`, `Jc`, `He`, `Fs` *et* `Cp` *pendant les 50 années d'observations. On observe une légère différence entre les deux estimateurs.*

l'équation (4.5) :

$$\bar{P}(n) = \sum_{k=0}^{n} \bar{q}^{(k)} * (I - \bar{H})(n), n \in \mathbb{N}, \tag{7.23}$$

$$\hat{P}(n) = \sum_{k=0}^{n} \hat{q}^{(k)} * (I - \hat{H})(n), n \in \mathbb{N}. \tag{7.24}$$

La figure 7.7 représente la loi de transition semi-markovienne

$$P_j(n) = \mathbb{P}(X_n = j)$$

en terme de proportions. Les deux modèles déduits de l'estimateur empirique et de l'estimateur du maximum de vraisemblance sont quasiment identiques.

Temps d'absorption

Les chaînes semi-markoviennes déduites de l'estimateur empirique et de l'estimateur du maximum de vraisemblance admettent "culture pérenne" `Cp` comme étant absorbant, voir la figure 7.8. Nous nous intéressons au temps mis par le système avant d'atteindre cet état.

Soit $T_i^{\mathtt{Cp}}$ le temps d'atteinte de $\mathtt{Cp}$ partant d'un état $i \in E \setminus \{\mathtt{Cp}\}$, c'est-à-dire :

$$T_i^{\mathtt{Cp}} \overset{\text{déf}}{=} T_m, \text{ avec } m = \min\{l \in \mathbb{N},\ Y_l = \mathtt{Cp},\ Y_0 = i\}$$

où $(Y_k)_{k\in\mathbb{N}}$, la chaîne de Markov incluse de $(X_n)_{n\in\mathbb{N}}$.

Notons :

$$g_i^{\mathtt{Cp}}(n) \overset{\text{déf}}{=} \mathbb{P}_i(T_i^{\mathtt{Cp}} = n), \quad n \in \mathbb{N}$$

la probabilité d'atteindre $\mathtt{Cp}$ au temps n partant de l'état i, avec :

$$g_i^{\mathtt{Cp}}(0) = 0, \forall\, i \in E \setminus \{\mathtt{Cp}\}.$$

Proposition 7.4.1. *La moyenne $m_{i\mathtt{Cp}}$ du temps d'atteinte de l'état absorbant vérifie :*

$$m_{i\mathtt{Cp}} = \nu_i + \sum_{k\neq\mathtt{Cp}} Q_{ik}\, m_{k\mathtt{Cp}} \tag{7.25}$$

où ν_i est la moyenne des temps de séjour dans l'état $i \in E \setminus \{\mathtt{Cp}\}$.

Démonstration. Nous nous inspirons de la démonstration de Kulkarni [28] :

$$\begin{aligned}
m_{i\mathtt{Cp}} &= \mathbb{E}_i(T_i^{\mathtt{Cp}} 1_{\{Y_1=\mathtt{Cp}\}}) + \mathbb{E}_i(T_i^{\mathtt{Cp}} 1_{\{Y_1\neq\mathtt{Cp}\}}) \\
&= Q_{i\mathtt{Cp}}\, \nu_i + \sum_{r\neq\mathtt{Cp}} \mathbb{E}_i(T_i^{\mathtt{Cp}} | 1_{\{Y_1=k\}})\mathbb{P}_i(Y_1 = r) \\
&= Q_{i\mathtt{Cp}}\, \nu_i + \sum_{r\neq\mathtt{Cp}} \left(\mathbb{E}_r(T_i^{\mathtt{Cp}}) + \mathbb{E}_i(T_1)\right) \\
&= Q_{i\mathtt{Cp}}\, \nu_i + \sum_{r\neq\mathtt{Cp}} Q_{ir}\, m_{r\mathtt{Cp}} + (1 - Q_{i\mathtt{Cp}})\, \nu_i.
\end{aligned}$$

□

Étant donné que l'état initial du processus est $\mathtt{Ca}$, nous nous intéressons à $m_{\mathtt{CaCp}}$. En utilisant l'équation (7.25), nous avons le système d'équations :

$$\begin{cases}
m_{\mathtt{CaCp}} = \nu_{\mathtt{Ca}} + Q_{12}\, m_{\mathtt{JaCp}} + Q_{13}\, m_{\mathtt{HeCp}}, \\
m_{\mathtt{JcCp}} = \nu_{\mathtt{Jc}} + Q_{21}\, m_{\mathtt{CaCp}} + Q_{24}\, m_{\mathtt{FsCp}}, \\
m_{\mathtt{HeCp}} = \nu_{\mathtt{He}} + Q_{31}\, m_{\mathtt{CaCp}}, \\
m_{\mathtt{FsCp}} = \nu_{\mathtt{Fs}} + Q_{41}\, m_{\mathtt{CaCp}}
\end{cases} \tag{7.26}$$

où $\nu_{\mathtt{Ca}}$, $\nu_{\mathtt{Jc}}$, $\nu_{\mathtt{He}}$ et $\nu_{\mathtt{Fs}}$ sont les moyennes des temps de séjour sur les quatre états non absorbants. La résolution de (7.26) permet d'écrire que :

$$m_{\mathtt{CaCp}} = \frac{\nu_{\mathtt{Ca}} + Q_{12}\nu_{\mathtt{Jc}} + Q_{13}\nu_{\mathtt{He}} + Q_{12}Q_{24}\nu_{\mathtt{Fs}}}{1 - (Q_{13}Q_{31} + Q_{12}Q_{21} + Q_{12}Q_{24}Q_{41})}. \tag{7.27}$$

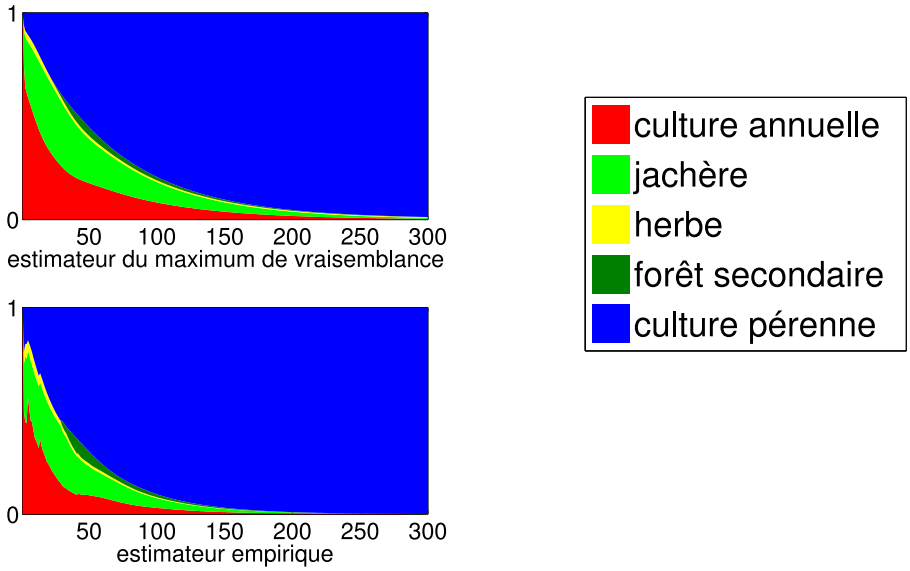

Figure 7.8: *Modèles semi-markoviens absorbants : Absorption du processus par l'état culture pérenne* (`Cp`). *Le temps moyenne d'absorption est* 65 *ans pour l'estimateur du maximum de vraisemblance et 39 ans pour l'estimateur empirique.*

En calculant cette valeur avec les estimateurs empiriques $\bar{Q}$ et $\bar{f}$, on trouve :

$$\bar{m}_{\texttt{CaCp}} = 39 \text{ ans}\,.$$

Pour l'estimateur du maximum de vraisemblance :

$$\hat{m}_{\texttt{CaCp}} = 65 \text{ ans}.$$

7.4.2 Modèle régulier

Nous analysons maintenant le modèle semi-markovien déduit de l'estimateur bayésien de la matrice de transition $\tilde{Q}$ et de la loi conditionnelle des temps de séjour $\tilde{f}$.

Loi de transition

L'estimateur bayésien de la matrice de transition $\tilde{Q}$ est une matrice régulière :

$$\tilde{Q} = \begin{pmatrix} 0 & 0.8164 & 0.0470 & 0 & 0.1366 \\ 0.9242 & 0 & 0 & 0.0758 & 0 \\ 0.8070 & 0.1930 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{7.28}$$

Les figures 7.9, 7.10 et 7.11 montrent les estimations des paramètres des lois des temps de séjour par la procédure de Monte Carlo par chaîne de Markov. L'estimateur du noyau semi-markovien et de la fonction de répartition des temps de séjour sont :

$$\tilde{q}_{ij}(k) = \tilde{Q}_{ij}\ \tilde{f}_{ij}(k), i, j \in E, k \in \mathbb{N}, \tag{7.29}$$

$$\tilde{H}_i(k) = \sum_{l=1}^{k} \sum_{j \in E} \tilde{q}_{ij}(l),\ i \in E,\ k \geq 0. \tag{7.30}$$

La loi de transition semi-markovienne correspondant au modèle est

$$\tilde{P}(n) = \sum_{k=0}^{n} \tilde{q}^{(k)} * (I - \tilde{H})(n), n \in \mathbb{N}. \tag{7.31}$$

Les figures 7.12 et 7.14 représentent les lois de transition semi-markoviennes pour les modèles déduites de l'estimateur empirique, de l'estimateur du maximum de vraisemblance et de l'estimateur bayésien calculé avec la méthode MCMC.

Loi limite

Avec les estimateurs bayésiens de $\tilde{Q}$ et $\tilde{f}$, la chaîne de Markov incluse est irréductible et les temps de séjour sur les états ont tous une moyenne finie. D'après la proposition 6.2.5, la chaîne semi-markovienne $(X_n)_{n \in \mathbb{N}}$ admet une loi limite

$$\Pi_j = \frac{\pi_j\, \nu_j}{\sum_{i \in E} \pi_i\, \nu_i},\ j \in E \tag{7.32}$$

où π est la loi invariante de la chaîne de Markov incluse $(Y_k)_{k \in \mathbb{N}}$ et ν le vecteur contenant les moyennes des temps de séjour sur les états (6.10).

Après calcul, on trouve :

$$\Pi = (0.1490,\ 0.2134,\ 0.0524,\ 0.0379,\ 0.5473). \tag{7.33}$$

La loi limite Π correspond aux proportions des parcelles occupées par chaque état à l'équilibre, voir la figure 7.15. C'est-à-dire : à l'équilibre, 14.9% des parcelles sont en état "culture annuelle" (`Ca`), 21.34% en "jachère" (`Jc`), 5.24% en "herbe" (`He`), 3.79% en "forêt secondaire" (`Fs`) et 54.73% en "culture pérenne" (`Cp`), voir la figure 7.13.

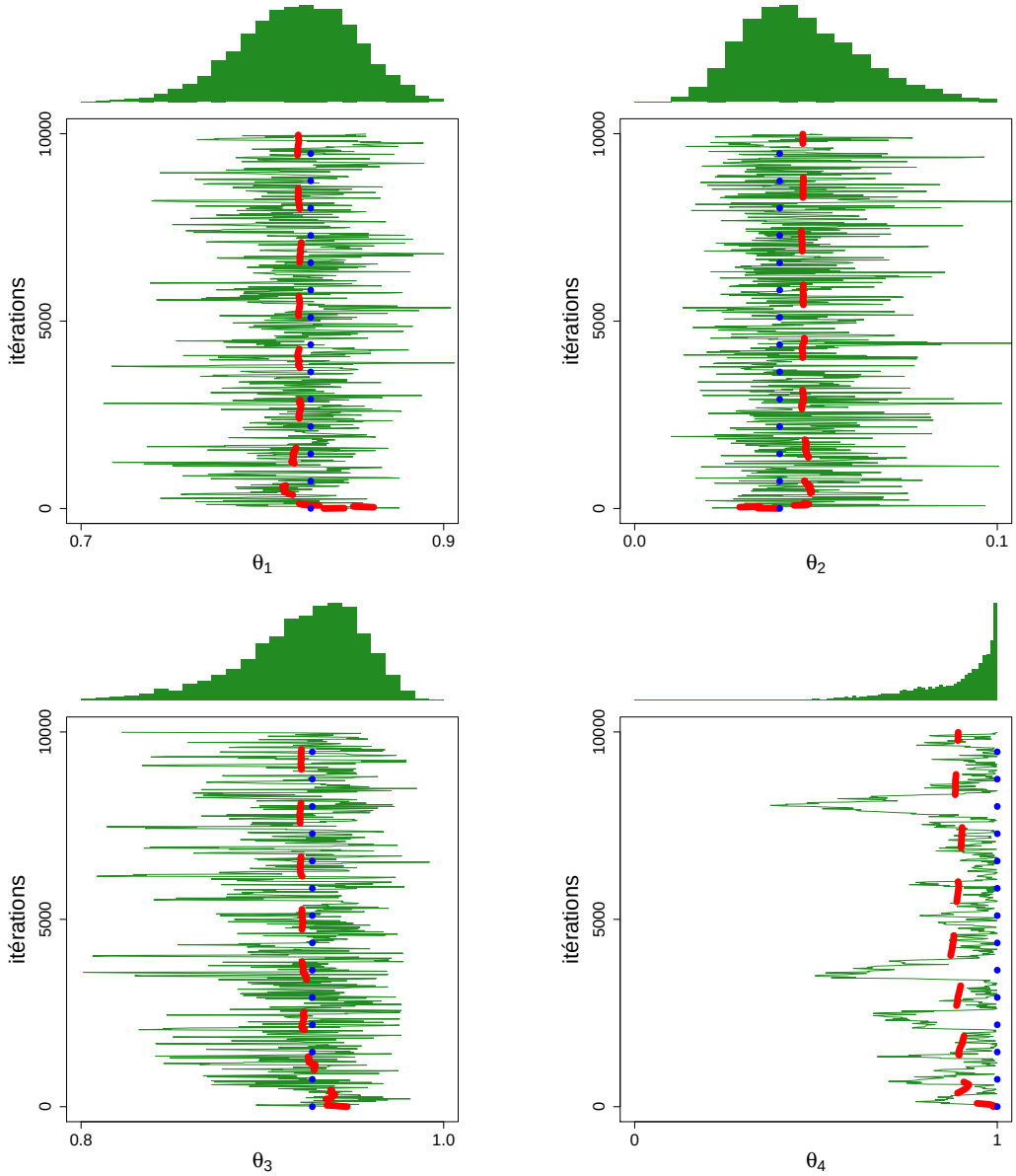

Figure 7.9: *Estimateur bayésien et estimateur du maximum de vraisemblance du paramètre θ de la matrice de transition de la chaîne incluse : échantillon de la loi a posteriori obtenu par la méthode MCMC en "—/vert", la moyenne associée en "- - -/rouge" et l'estimateur du maximum de vraisemblance en "$\cdots$/bleu".*

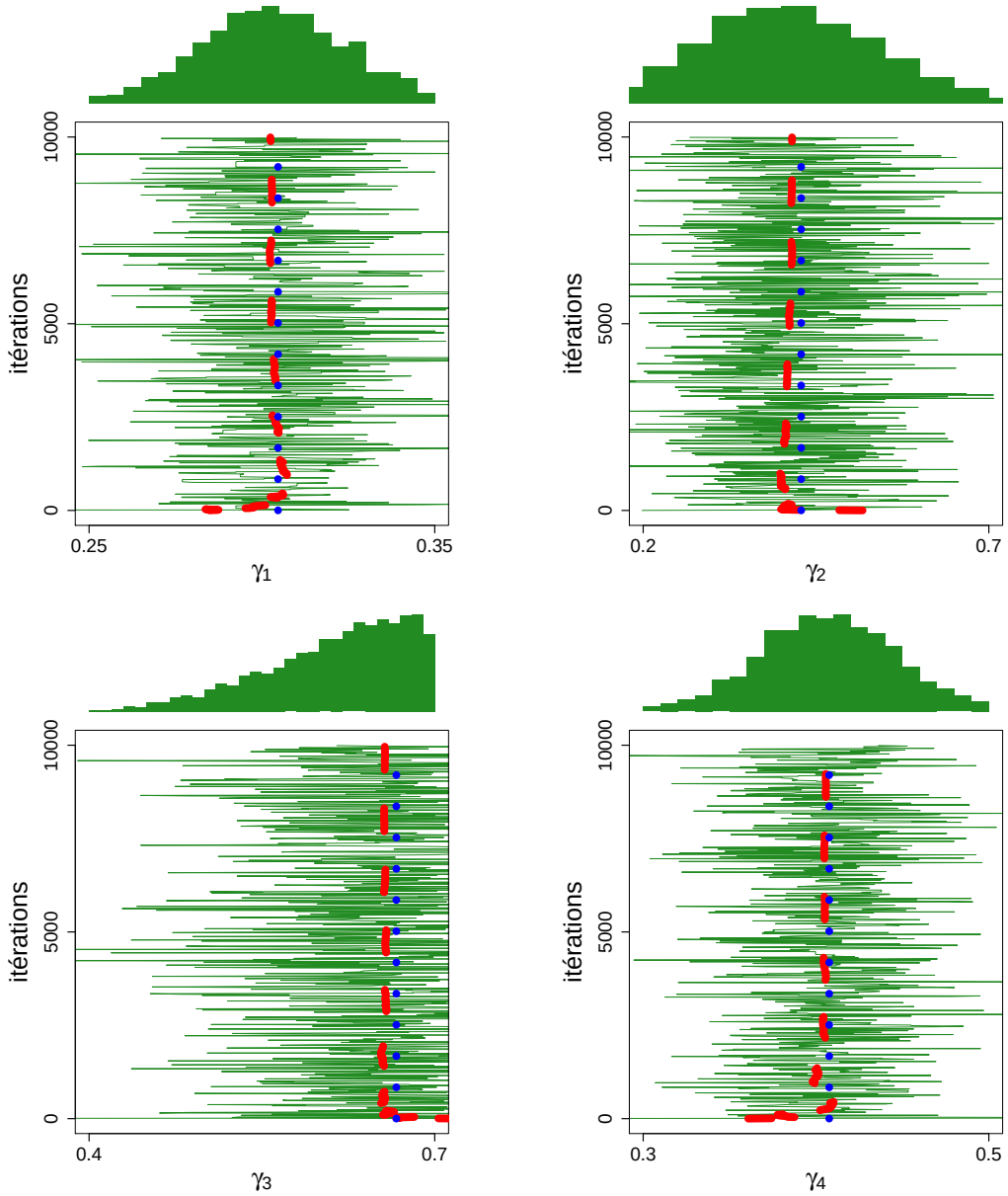

Figure 7.10: *Estimation bayésienne des paramètres des lois géométriques correspondants aux temps de séjour pour les transitions* `CaJc`, `CaHe`, `CaCp` *et* `JcCa` *: échantillon de la loi a posteriori obtenu par la méthode MCMC en "—/vert", la moyenne associée en "- - -/rouge" et l'estimateur du maximum de vraisemblance en "· · · /bleu".*

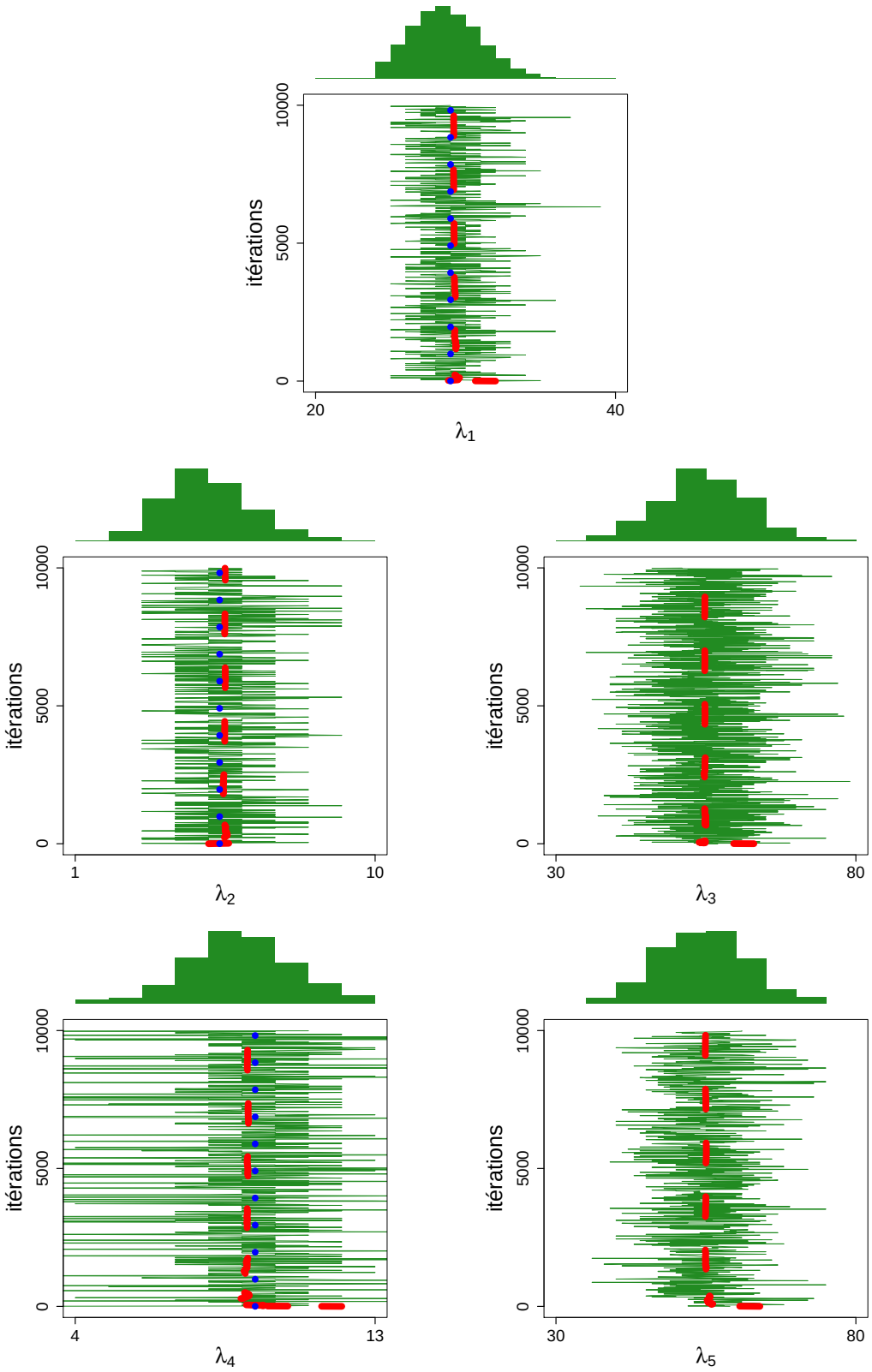

Figure 7.11: *Estimation bayésienne des paramètres des lois de Poisson correspondants aux temps de séjour pour les transitions* `FaFs`, `HeCa`, `HeJc`, `FsCa` *et* `CpCa` *: échantillon de la loi a posteriori obtenu par la méthode MCMC en "—/vert", la moyenne associée en "- - -/rouge" et l'estimateur du maximum de vraisemblance en "$\cdots$/bleu".*

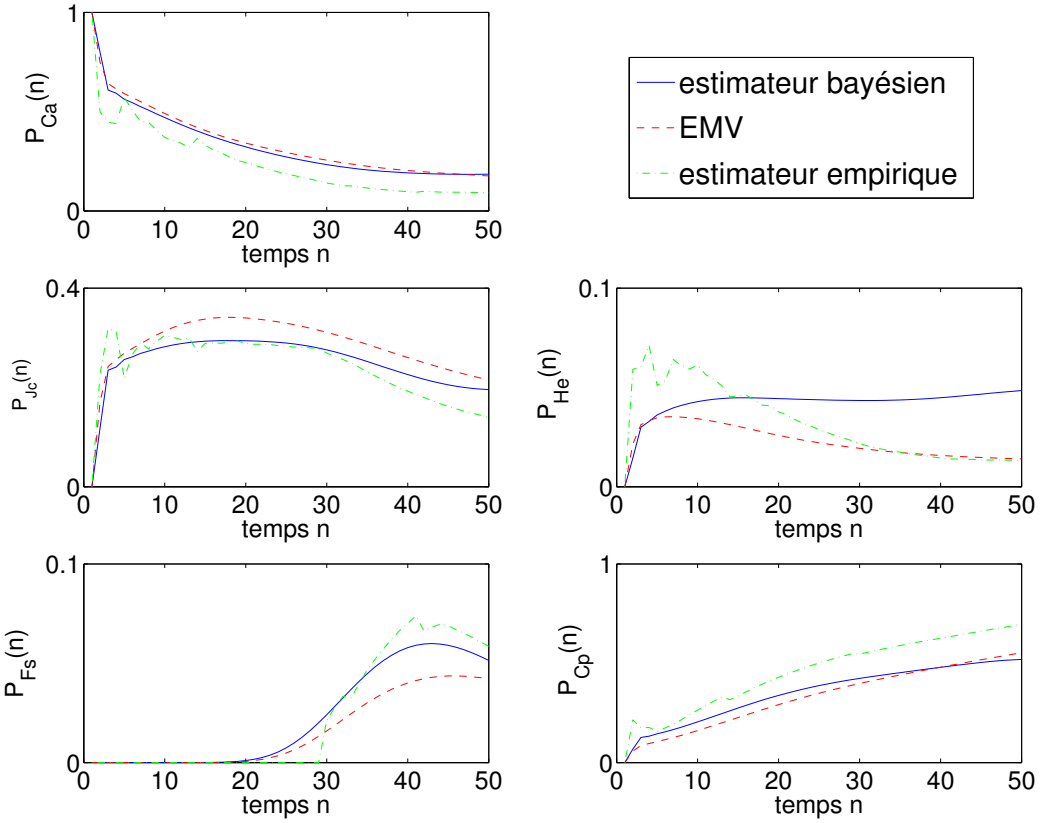

Figure 7.12: *Estimation de la loi de transition semi-markovienne de 1956 à 2006.*

7.4.3 Discussion

Selon les trois méthodes d'estimation utilisées, on obtient trois modèles semi-markoviens différents. Les modèles déduits de l'estimateur empirique et de l'estimateur du maximum de vraisemblance sont absorbants. Le modèle proposé par la méthode bayésienne est régulier et admet une loi limite qui correspond aux proportions des occupations des parcelles à l'équilibre.

À court terme, le modèle absorbant et le modèle régulier donnent des résultats semblables. Les différences sont nettes à partir de la $50^{\text{ième}}$ année d'observation, voir la figure 7.15.

Dans la figure 7.16, on représente l'évolution des usages des parcelles sans l'état absorbant (`Cp`). Pour le modèle déduit de l'estimateur du maximum de vraisemblance, la loi du processus semi-markovien converge vers une proportion stable entre "culture annuelle" et "jachère" alors que les états "forêt secondaire"

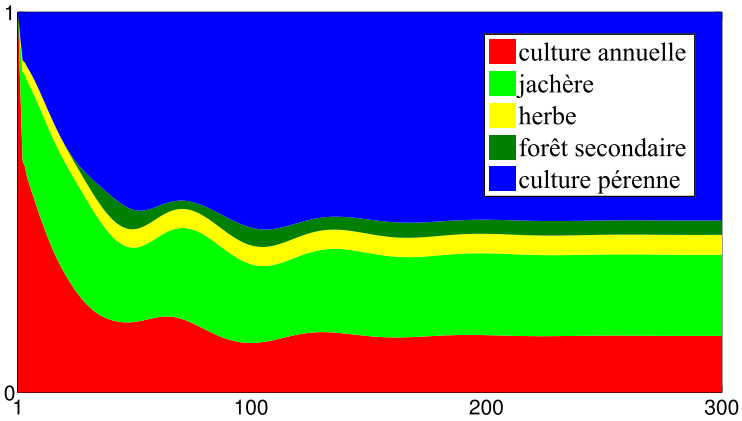

Figure 7.13: *Loi limite du modèle semi-markovien régulier déduit de l'estimateur bayésien*

et "herbe" disparaissent. Pour le modèle régulier, les proportions d'usages des parcelles se stabilisent pour tous les états.

Pour les deux modèles, les proportions des parcelles en "jachère" sont presque identiques : les sommes des proportions des parcelles sur les états "culture annuelle", "herbe" et "forêt secondaire" sont les mêmes, voir la figure 7.17.

Dans tous les cas, la proportion des parcelles en "jachère" atteint son équilibre vers la $50^{\text{ième}}$ année d'observations.

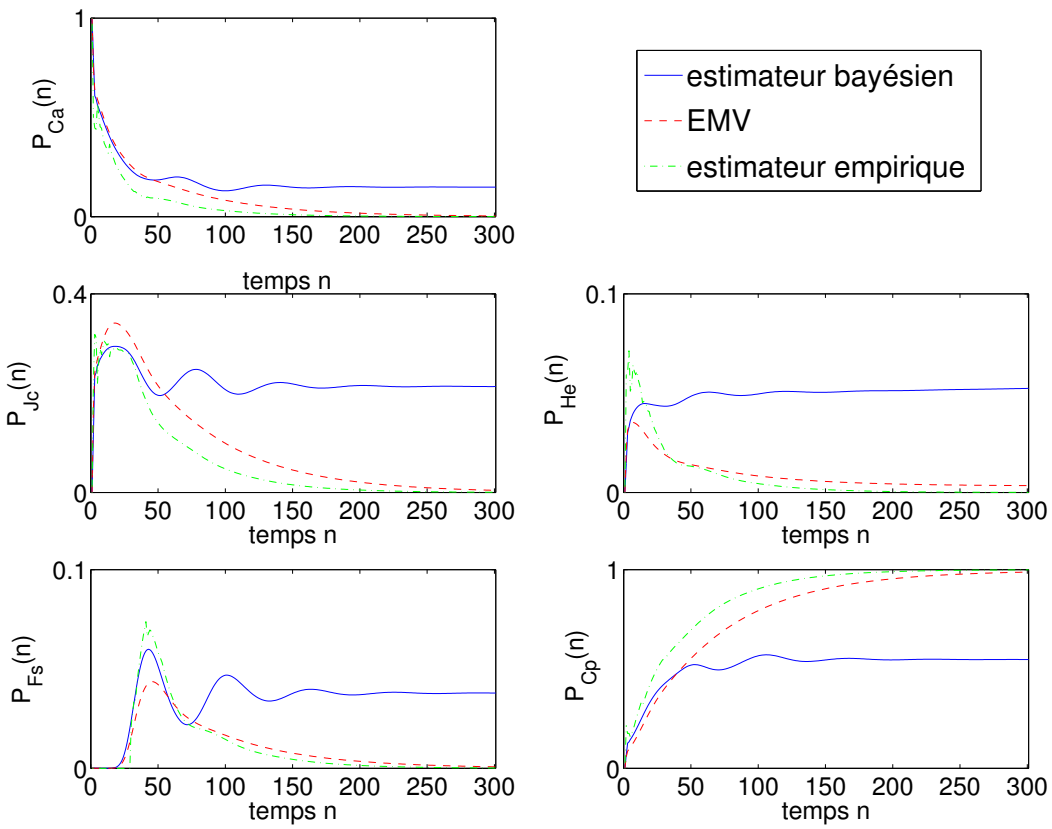

Figure 7.14: *Comparaison des lois de transition semi-markoviennes estimées selon les trois modèles déduits de l'estimateur empirique, de l'estimateur du maximum de vraisemblance et de l'estimateur bayésien.*

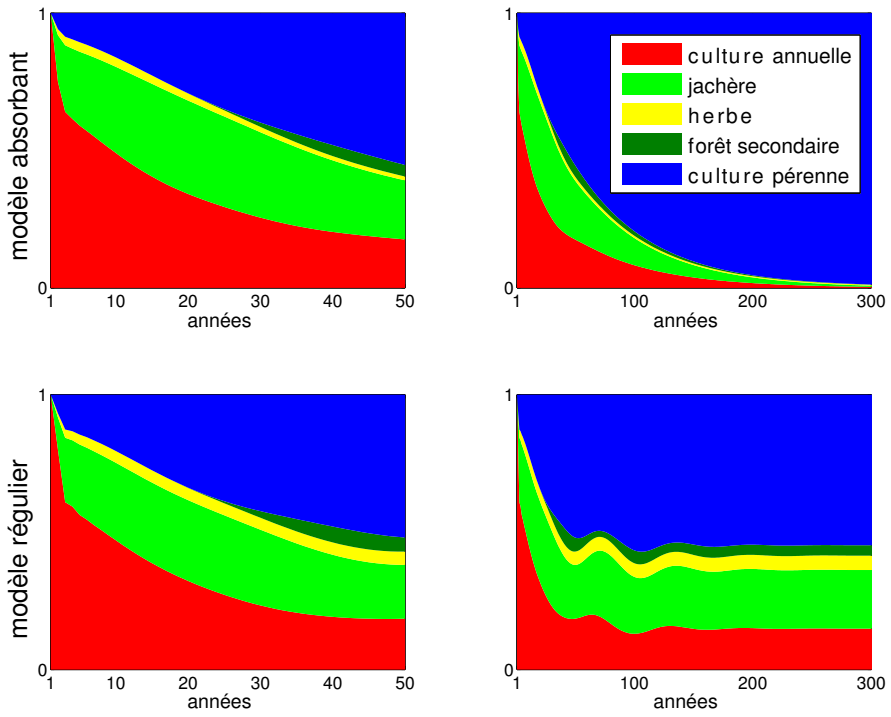

Figure 7.15: *Proportions d'occupations des parcelles selon les modèles déduits de l'EMV et de l'estimateur bayésien.*

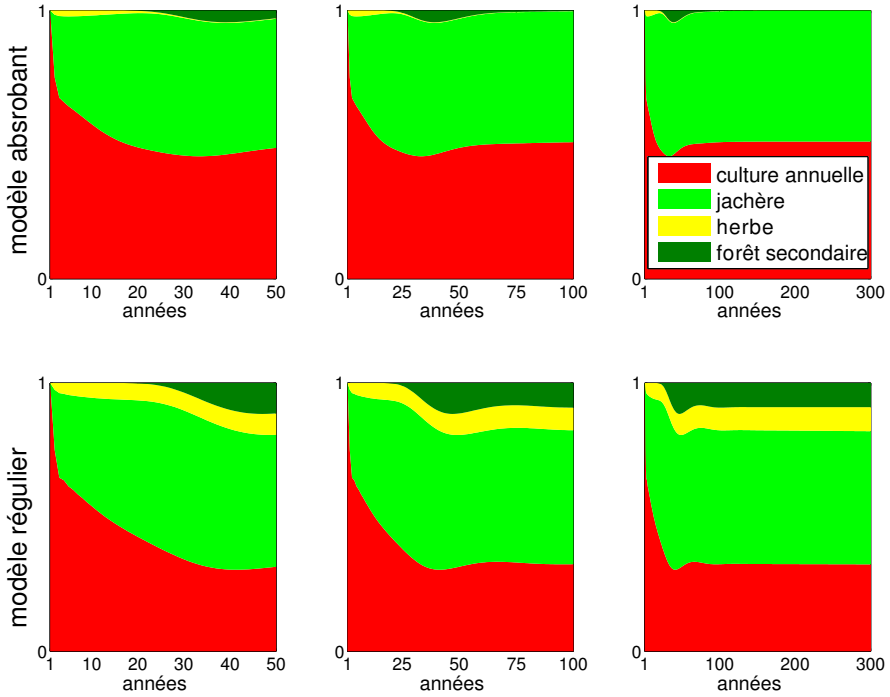

Figure 7.16: *Proportions d'usages des parcelles sans l'état* (`Cp`) *pour les modèles déduites de l'EMV et de l'estimateur bayésien.*

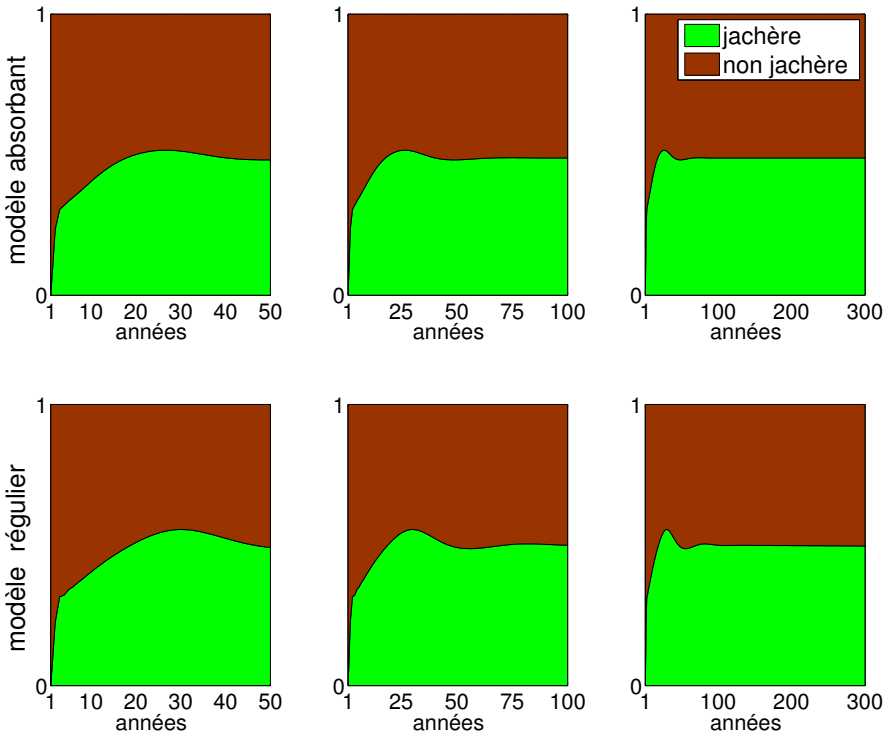

Figure 7.17: *Proportions des parcelles en état "jachère" et "non jachère" pour le modèle absorbant déduit de l'EMV et le modèle régulier déduit de l'estimateur bayésien.*

Conclusions et perspectives

Nous avons développé une démarche de modélisation de dynamique d'usage de parcelles agricoles fondée sur des chaînes de Markov à espace d'état fini. Nous avons adopté le point de vue pragmatique suivant : partant de relevés annuels d'usages de ces parcelles, il est communément admis de construire la matrice de transition empirique afin d'inférer la dynamique sous-jacente. Considérer une telle matrice revient implicitement à supposer que ces parcelles sont toutes indépendantes les unes des autres et qu'elles suivent tout la dynamique de Markov homogène en temps. Cette hypothèse est naturellement fausse mais, comme nous l'avons fait remarquer, il s'agit de l'étape initiale dans l'étude de telles dynamiques. De plus cette matrice de transition empirique est en fait l'estimateur du maximum de vraisemblance de la matrice de transition de la chaîne de Markov. Cette approche consiste à exploiter uniquement les données et ne permet pas d'intégrer les connaissances a priori des experts. Afin de prendre en compte certaines transitions non-présentes dans les données mais pourtant jugées comme crédibles par les experts nous avons fait appel à une approche de statistique bayésienne. La loi a priori retenue est la loi de Jeffreys et nous avons fait appel à une méthode d'approximation de Monte Carlo par chaîne de Markov afin d'identifier la matrice de transition. La matrice de transition empirique n'est pas régulière car la chaîne correspondante admet un état absorbant. La matrice de transition obtenue par la méthode bayésienne est elle régulière. Les deux dynamiques coïncident sur 20 ans. Au-delà ce n'est plus le cas, mais pour les états non absorbant les deux dynamiques coïncident. On constate notamment que sur ces états, la dynamique converge en moins de 5 ans vers une loi limite de 58% de culture annuelle et 48% de jachère. Nous avons effectué plusieurs tests statistiques portant sur l'homogénéité et sur la loi des temps de séjour. Ces tests sont positifs.

En revanche, pour un second jeu de données étendu, l'hypothèse sur les temps de séjour de loi géométrique a été rejetée. Pour certains états, la loi du temps de séjour dépend de l'état suivant et de plus cette loi n'est parfois pas géométrique.

En d'autres termes, l'hypothèse markovienne est rejetée. En fait cette situation correspond parfaitement au cadre des modèles semi-markoviens. Dans ce cadre il faut identifier la matrice de transition de la chaîne incluse ainsi que la loi des temps de séjour conditionnés par l'état suivant. Nous avons effectué une analyse similaire au cas précédent. Dans le cadre semi-markovien l'estimateur empirique ne correspond plus exactement à l'estimateur du maximum de vraisemblance mais les deux modèles ainsi identifiés sont très proches. À nouveau l'approche bayésienne nous a permis d'intégrer des transitons non présentes dans les données mais jugées pertinentes par les spécialistes dans les échelles de temps considérées. Ce second cas comprend notamment un état de forêt secondaire et notre approche nous a permis de quantifier les temps de retour à cet état et l'équilibre qui s'instaure entre "culture/jachère" d'une part et forêt secondaire d'autre part. L'équilibre concernant la forêt secondaire est atteint en environ 50 ans, celui entre culture et jachère en environ 25 ans.

Comme nous l'avons à plusieurs fois précisé au cours de ce travail, les chaînes de Markov sont des instruments simples qui impliquent donc des hypothèses certainement simplistes du point de vue de l'application. La principale faiblesse est de supposer que les parcelles doivent être indépendantes les unes des autres. Mais avant d'étudier les extensions de ces modèles, nous pouvons explorer deux voies.

D'abord nous pouvons appliquer les outils déjà développés dans cette thèse à des domaines d'applications connexes : au cours de ce travail de thèse nous avons par exemple étudié des applications de dynamiques agraires dans un contexte de séquestration du carbone. Ces travaux très préliminaires n'ont pas été exposés dans cette thèse mais pourront être développés dans un avenir proche.

Ensuite on peut étendre l'application de ces mêmes modèles au contrôle optimal dans le but de proposer des stratégies d'exploitation "optimale" ; le contrôle se ferait donc au travers de l'exploitant agricole. En fait dans le cadre de cette application, le contrôle optimal n'est peut être pas l'approche la plus adaptée. On peut plutôt s'orienter vers des outils de la théorie de la viabilité : dans ce cadre il s'agit d'amener la dynamique vers un régime de fonctionnement "viable" selon des critères de développement durable et de protection des ressources naturelles. Cet axe ne peut naturellement se faire que dans une étude plus globale sur le plan socio-agronomique menée par des spécialistes du domaine par exemple en collaboration avec l'IRD ou le CIRAD.

Enfin, en termes d'extension du modèle on peut considérer deux possibilités. La première serait la prise en compte de dynamiques cachées au travers de modèles de Markov cachés ou de modèles de semi-Markov cachés. La seconde serait de proposer des modèles de dynamiques spatio-temporelles. La composante spa-

tiale pourrait se limiter à quelques compartiments ou bien à une représentation plus fine sous forme de grille spatiale. Ce genre de modèles permettrait d'intégrer des éléments de connaissances spatiales comme le coût d'exploitation d'une parcelle mesuré en termes d'accessibilité de la parcelle. À une échelle plus macroscopique on pourrait également intégrer des observations satellitaires. On ne doit néanmoins pas perdre de vue que s'il est relativement facile et intéressant de complexifier les modèles, les données et mesures disponibles sont relativement limitées dans le contexte malgache ce qui nous oblige à contenir la complexité des modèles proposés.

Bibliographie

[1] C. Agarwal, G. M. Green, J. Morgan Grove, T. P. Evans et C. M. Schweik : A review and assessment of land-use change models : dynamics of space, time, and human choice. Rapport technique, Gen. Tech. Rep. NE-297. Newton Square, PA : U.S. Department of Agriculture, Forest Service, Northeastern Research Station, 2002.

[2] D. K. Agarwal, J. A. Silander Jr., A. E. Gelfand, R. E. Deward et 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 et L. A. Goodman : Statistical inference about Markov chains. *Annals of Mathematical Statistics*, 28:89–109, 1957.

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

[5] S. Aubert, S. Razafiarison et A. Bertrand : *Déforestation et systèmes agraires à Madagascar : les dynamiques des tavy sur la côte orientale*. Repères. Cirad, 2003.

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

[7] Vlad Stefan Barbu et Nikolaos Limnios : Reliability of semi-markov systems in discrete time : Modeling and estimation. *In* Krishna B. Misra, éditeur : *Handbook of Performability Engineering*, pages 369–380. Springer London, 2008.

[8] B. O. Burnham : Markov intertemporal land use simulation model. *Southern Journal of Agricultural Economics*, 5:253–258, 1973.

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

[10] F. Campillo, A. Raherinirina et R. Rakotozafy : Un modèle markovien de transition agraire. *In Proceedings of the 8th International African Conference on Research in Computer Science and Applied mathematics (CARI'08), October 27-30, Rabat, Maroc*, 2008.

[11] S. Carrière, Ph. Roche, M. Viano, E. Iticène, M. Picot et Th. Tatoni : Hétérogénéité des paysages, dispersion des graines et biodiversité : le cas du terroir d'Ambendrana. *In Transitions agraires, dynamiques écologiques et conservation : le "corridor" Ranomafana-Andringitra*, 2007.

[12] T. Casse, A. Milhoj, S. Ranaivoson et J. R. Randriamanarivo : Causes of deforestation in southwestern Madagascar : what do we know ? *Forest Policy and Economics*, 6(1):33–48, 2004.

[13] H. Caswell : *Matrix Population Models : Construction, Analysis, and Interpretation.* Sinauer Associates, seconde édition, 2001.

[14] E. Cinlar : Markov renewal theory : A survey. Discussion Papers 53, Northwestern University, Center for Mathematical Studies in Economics and Management Science, septembre 1973.

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

[16] R. L. Dobrushin : Central limit theorem for non-stationary Markov chains I, II. *Theory of Probability and its Applications*, 1, 1956.

[17] O. Feldman, V. N. Korotkov et D. O. Logofet : The monoculture vs. rotation strategies in forestry : formalization and prediction by means of Markov-chain modelling. *Journal of Environmental Management*, 77:111–121, 2005.

[18] A. E. Gelfand et A. F. M. Smith : Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85 (410):398–409, juin 1990.

[19] C. Genest et 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.

[20] G.M. Green et R.W. Sussman : Deforestation history of the eastern rain forests of Madagascar from satellite images. *Science*, 248:212–215, 1990.

[21] K. W. Hastings : Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, avril 1970.

[22] 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.

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

[24] H. S. Horn : Markovian properties of forest succession. *In Ecology and Evolution of Communities*, pages 196–211. Harvard University Press, 1975.

[25] D. M. Howard, P. J. A. Howard et D. C. Howard : A markov model projection of soil organic carbon stores following land use changes. *Journal of Environmental Management*, 45:287–302, 1995.

[26] J. Janssen et R. Manca : *Applied semi-Markov processes.* Springer, 2006.

[27] V. Krishnaprasad, J. Cardina, B. Stinner, K. Badarinath, R. Moore, R. Stinner et C. Hoy : Modeling carbon flows from land use change and forestry datasets using transition probabilities - a case study from India. *In 34th COSPAR Scientific Assembly*, volume 34 de *COSPAR, Plenary Meeting*, 2002. Abstract.

[28] V. G. Kulkarni : *Introduction to Modeling and Analysis of Stochastic Systems.* Springer, 2011.

[29] S. W. Lagakos, C. J. SOMMER et M. ZELEN : Semi-markov models for partially censored data. *Biometrika*, 65(2):311–317, 1978.

[30] P. Lévy : Processus semi-markoviens. *In Proceedings of the International Congress on Mathematics, Amsterdam*, volume 3, pages 416–426, 1954.

[31] N. Limnios et G. Oprisan : Ch. 14. an introduction to semi-markov processes with application to reliability. *In* D.N. Shanbhag et C.R. Rao, éditeurs : *Stochastic Processes : Modelling and Simulation*, volume 21, pages 515 – 556. Elsevier, 2003.

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

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

[34] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller et E. Teller : Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1091, 1953.

[35] M.G. Monticino, T. Cogdill et M.F. Acevedo : Cell interaction in semi-markov forest landscape models. *In* A.E. Rizzoli et A.J. Jakeman, éditeurs : *Integrated assessment and decision support, Proceedings of the First Biennial Meeting of the IEMSS. Lugano, Switzerland*, pages 227–232, 2002.

[36] A.D. Moore : The semi-Markov model : a useful tool in the analysis of vegetation dynamics for management. *Journal of Environmental Management*, 30:111–30, 1990.

[37] J. R. Norris : *Markov chains*. Cambridge University Press, 1998.

[38] Odile Pons : Estimation of semi-markov models with right-censored data. *In* N. Balakrishnan et C.R. Rao, éditeurs : *Advances in Survival Analysis*, volume 23 de *Handbook of Statistics*, pages 175 – 194. Elsevier, 2003.

[39] V.K. Prasad, K.V.S. Badarinath, H. Tsuruta, S. Sudo, S. Yonemura, J. Cardina, B. Stinner, R. Moore, D. Stinner et C. Hoy : Implications of land use changes on carbon dynamics and sequestration-evaluation from forestry datasets, india. *The Environmentalist*, 23:175–187, 2003.

[40] R. Pyke : Markov Renewal Processes with Finitely Many States. *The Annals of Mathematical Statistics*, 32:1243–1259, 1961.

[41] H. Rabetaliana, A. Bertrand, N. Razafimamonjy et E. Rabemananjara : Dynamiques des forêts naturelles de montagne à Madagascar. *Bois et Forêts des Tropiques*, 276:59–72, 2003.

[42] A. Raherinirina, R. Rakotozafy, Hervé D. et F. Campillo : A markov model of land use dynamics. *Rapport de recherche INRIA n° 7670*, 2011.

[43] A. Raherinirina, R. Rakotozafy, D. Hervé et F. Campillo : Inférence d'une dynamique d'usage des sols par des modèles semi-markovie. *In Statistique Appliquée pour le Développement de l'Afrique (SADA)*, 2013.

[44] N. Ramankutty, H. K. Gibbs, F. Achard, R. Defries, J. A. Foley et R. A. Houghton : Challenges to estimating carbon emissions from tropical deforestation. *Global Change Biology*, 13(1):51–66, 2007.

[45] A. Satake, H.M. Leslie, Y. Iwasa et S.A. Levin : Coupled ecological-social dynamics in a forested landscape : Spatial interactions and information flow. *Journal of Theoretical Biology*, 246:695–707, 2007.

[46] U. A. Schneider : Soil organic carbon changes in dynamic land use decision models. *Agriculture, Ecosystems and Environment*, 119, 2007.

[47] E. Seneta : *Nonnegative matrices and Markov chains.* Springer-Verlag, New York, second édition, 1981.

[48] G. Serpantié, Rasolofoharinoro et S. Carrière : *Transitions agraires, dynamiques écologiques et conservations. Le corridor Ranomafana-Andringitra.* Acte du séminaire GEREM, IRD CITE, 2007.

[49] W. L. Smith : Regenerative stochastic processes. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 232(1188): 6–31, 1955.

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

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

[52] L. Takás : Some investigation concerning recurrent stochastic processes of a certain type. *Magyar Tud. Akad. Mat. Kutató Int. Kozl*, 3:115–128, 1954.

[53] P. K. Thornton et P. G. Jones : A conceptual approach to dynamic agricultural land-use modelling. *Agricultural Systems*, 57(4):505–521, August 1998.

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

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

[56] M. B. Usher : Modelling ecological succession, with particular reference to Markovian models. *Plant Ecology*, 46:11–18, 1981.

[57] R. van Hulst : On the dynamics of vegetation : Markov chains as models of succession. *Vegetatio*, 40:3–14, 1979.

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

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

[60] Q. Weng : Land use change analysis in the Zhujiang Delta of China using satellite remote sensing, GIS and stochastic modelling. *Journal of Environmental Management*, 64:273ñ284, 2002.

[61] Christopher K. Wikle, L. Mark Berliner et Noel Cressie : Hierarchical bayesian space-time models. *Environmental and Ecological Statistics*, 5(2): 117–154, 1998.