



HAL
open science

Modèles hiérarchiques et processus ponctuels spatio-temporels - Applications en épidémiologie et en sismologie

Larissa Valmy

► **To cite this version:**

Larissa Valmy. Modèles hiérarchiques et processus ponctuels spatio-temporels - Applications en épidémiologie et en sismologie. Applications [stat.AP]. Université des Antilles-Guyane, 2012. Français. NNT : 2012AGUY0555 . tel-00841146

HAL Id: tel-00841146

<https://theses.hal.science/tel-00841146>

Submitted on 3 Jul 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.



THÈSE

pour obtenir le grade de

Docteur de l'Université des Antilles et de la Guyane

SPÉCIALITÉ : MATHÉMATIQUES APPLIQUÉES

présentée par

Larissa Audrey VALMY

(larissa.valmy@univ-ag.fr)

**Modèles hiérarchiques et processus ponctuels
spatio-temporels - Applications en épidémiologie et en
sismologie**

soutenue publiquement le 05 Novembre 2012

Composition du jury :

Jean-Marc BARDET	Professeur	Université de Paris 1, France	<i>Rapporteur</i>
Jean-Heinrich DAUGROIS	Directeur de Recherche	CIRAD, FWI	<i>Examineur</i>
Jacques LAMINIE	Professeur	Université des Antilles-Guyane, FWI	<i>Examineur</i>
Andrey OLIFER	Professeur	Gwinnett College, Georgia, USA	<i>Rapporteur</i>
Alain PIÉTRUS	Professeur	Université des Antilles-Guyane, FWI	<i>Examineur</i>
Jean VAILLANT	Professeur	Université des Antilles-Guyane, FWI	<i>Directeur</i>

Résumé

Mots-clés : Processus ponctuels, processus de Hawkes, processus de Cox, processus de Dirichlet, processus de vraisemblance, échantillonnage, modèles bayésiens, méthodes MCMC

Les processus ponctuels sont souvent utilisés comme modèles de répartitions spatiales ou spatio-temporelles d'occurrences. Dans cette thèse, nous nous intéressons tout d'abord à des processus de Cox dirigés par un processus caché associé à un processus de Dirichlet. Ce modèle correspond à des occurrences cachées influençant l'intensité stochastique des occurrences observées. Nous généralisons la notion de "Shot noise Cox process" introduite par Møller et développons le traitement bayésien par un échantillonneur de Gibbs combiné à un algorithme de Metropolis-Hastings. Nous montrons que cette méthode MCMC est à sauts réversibles. Le modèle prend en compte, en effet, un nombre aléatoire de contributions cachées influençant l'intensité du processus ponctuel observé donc a un espace paramétrique de dimension variable. Nous focalisons l'inférence statistique sur l'estimation de la valeur espérée de chaque contribution cachée, le nombre espéré de contributions cachées, le degré d'influence spatiale de ces contributions et leur degré de corrélation. Le test d'égalité des contributions et celui de leur indépendance sont ainsi développés. L'utilité en épidémiologie et en écologie est alors démontrée à partir de données de *Rubus fruticosus*, *Ibicella lutea* et de mortalité dans les cantons de Georgia, USA. En termes de données observées, deux situations sont considérées : premièrement, les positions spatiales des occurrences sont observées entre plusieurs paires de dates consécutives ; deuxièmement, des comptages sont effectués, au cours d'une période fixée, dans des unités d'échantillonnage spatiales.

D'autre part, nous nous intéressons aux processus ponctuels à mémoire introduits par Kagan, Ogata et Vere-Jones, précurseurs de la statistique sismologique. En effet, les processus ponctuels spatio-temporels ont une place importante dans l'étude des catalogues sismiques puisque ces derniers sont généralement constitués d'événements sismiques datés et géo-référencés. Nous avons étudié un modèle ETAS (Epidemic Type Aftershock Sequence) avec une intensité d'arrière-plan indépendante du temps et plusieurs fonctions déclenchantes permettant d'intégrer les événements antérieurs récents. Cette approche est utilisée pour étudier la sismicité de l'arc des Petites Antilles. Une étude comparative des modèles Gamma, Weibull, Log-Normal et loi d'Omori modifiée pour les fonctions déclenchantes est menée. Nous montrons que la loi d'Omori modifiée ne s'ajuste pas aux données sismiques des Petites Antilles et la fonction déclenchante la plus adaptée est le modèle de Weibull. Cela implique que le temps d'attente entre répliques dans la zone des Petites Antilles est plus faible que celui des régions à sismicité décrite par la loi d'Omori modifiée. Autrement dit, l'agrégation des répliques après un événement majeur est plus prononcée dans la zone des Petites Antilles. La possibilité d'inclure une intensité d'arrière-plan suivant un processus de Dirichlet centré sur un processus spatial log-gaussien est discutée.

Abstract

Keywords : Point processes, Hawkes processes, Cox processes, Dirichlet process, likelihood process, sampling, bayesian model, MCMC methods

Point processes are often used as spatial or spatio-temporal distribution models of occurrences. In this PhD dissertation, we focus first on Cox processes driven by a hidden process associated with a Dirichlet process. This model corresponds to hidden occurrences influencing the stochastic intensity of observed occurrences. We generalize the notion of "Shot noise Cox process" introduced by Møller and develop its bayesian analysis by a Gibbs sampler combined with a Metropolis-Hastings algorithm. We show that our MCMC method is a reversible jump one. The model takes into account a random number of hidden contributions producing effects on the observed point process intensity. Therefore the parameter space has a variable dimension. We focus the statistical inference on the estimation of the hidden contribution expected value, the hidden contribution expected number, the spatial influence and correlation parameters. The contribution equality test and contribution independence test are proposed. Applications in epidemiology and ecology are shown from *Rubus fruticosus* data, *Ibicella lutea* data and death number data in counties of Georgia, USA. Two situations are considered with respect to available data : firstly, the spatial positions of occurrences are observed between several pairs of consecutive dates ; secondly, counts are carried out over a fixed time interval in several spatial sampling units.

Secondly, we focus on point processes with memory introduced by Kagan, Ogata and Vere-Jones, pioneers in statistical seismology. In fact, spatio-temporal point processes play an important role in the studies of earthquake catalogs since they consist of seismic

events with their dates and spatial locations. We studied an ETAS (Epidemic Type Aftershock Sequence) model with time independent background intensity and several triggering functions taking into account previous events. We illustrate our approach with a seismicity study of the Lesser Antilles arc. A comparison study of Gamma, Weibull, Log-Normal and modified Omori law triggering function models is also carried out. We show that the modified Omori law does not fit the Lesser Antilles seismic data and the best adjusted triggering function is the Weibull model. Consequently, the waiting time between aftershocks is weaker in the Lesser Antilles zone compared to the one in regions with seismicity described by the modified Omori law. In other words, aftershock aggregativity is higher in the Lesser Antilles region. The stochastic background intensity following a Dirichlet process centered on a spatial log-normal process is discussed.

Remerciements

Je souhaite en premier lieu exprimer ma gratitude au professeur Jean VAILLANT, mon directeur de thèse, pour sa disponibilité, ses conseils, ses encouragements, au cours de ces trois années, et également pour avoir cru en mon travail de recherche.

Je voudrais exprimer ma reconnaissance au professeur Alain PIÉTRUS qui m'a fait l'honneur de participer et présider ce jury de thèse, ainsi qu'aux professeurs Jean-Marc BARDET et Delphine BLANKE pour leurs remarques et suggestions qui m'ont permis d'améliorer cette thèse. Je remercie le professeur Jacques LAMINIE, pour son aide précieuse et le professeur Andrey OLIFER.

Je réserve également ces quelques lignes en signe de profonde gratitude à Florencia CHIMARD, Célia JEAN-ALEXIS, Jean-Marc MOUNSAMY, Jimmy NAGAU et Paul NUIRO pour leur soutien indéfectible et leurs remarques avisées. J'ai une pensée pour Cédric COPOL et Mathias PEROUMALNAÏK qui ont eu l'occasion de me conseiller au début de cette thèse ainsi que pour les membres de l'Association des Jeunes Chercheurs en Guadeloupe, entre autres Ludovic NAGESSEUR et Sabin THENARD.

Je remercie la Région Martinique pour son soutien financier, les membres du Département de Mathématiques et Informatique, avec une attention particulière pour Michel GEOFFROY, Marylène TROUPÉ, Maria MALÉAMA, ainsi que les membres du Laboratoire de Mathématiques Informatique et Applications.

Je tiens tout particulièrement à remercier vivement mes parents pour tout ce que je leur dois ainsi que mon entourage proche pour avoir contribué à un environnement favorable à mes activités de recherche. Je remercie toutes les personnes ayant contribué de près ou de loin à l'élaboration de cette thèse.

Table des matières

Résumé	3
Abstract	5
Remerciements	7
Notations	13
Abréviations et sigles	15
Liste des figures	19
Liste des algorithmes	21
Introduction	23
1 Processus Ponctuels	33
1.1 Quelques notions essentielles	33
1.2 Processus de vraisemblance d'un processus ponctuel	36
1.3 Analyse résiduelle d'un processus ponctuel	36
1.4 Analyse d'autocorrélation et de dispersion	37
1.5 Processus ponctuels de référence	39
1.5.1 Processus de Bernoulli et Binomial	39
1.5.2 Processus de Poisson	39
1.5.3 Processus de Poisson temporel	40
1.5.4 Processus de renouvellement	41

1.5.5	Processus de Cox	41
1.5.6	Processus de Hawkes	41
2	Échantillonnage de lois <i>A POSTERIORI</i>	43
2.1	Exemple d'algorithme stochastique de simulation	44
2.2	Méthodes MCMC	44
2.2.1	Définitions essentielles	44
2.2.2	Efficacité et vitesse de convergence	46
2.3	Algorithme de Metropolis-Hastings	46
2.3.1	Algorithme de Metropolis-Hastings indépendant	46
2.3.2	Algorithme de Metropolis-Hastings à marche aléatoire	47
2.3.3	Algorithme de Metropolis-Hastings dit <i>une-variable-à-la-fois</i>	48
2.4	Échantillonneur de Gibbs	49
2.5	Méthodes hybrides	50
2.6	MCMC à sauts réversibles	50
2.7	Applications aux processus ponctuels	51
3	Association spatio-temporelle de processus	53
3.1	Exemples de données observables	55
3.2	Statistique d'un processus de Cox spatio-temporel associé à un DP	57
3.2.1	Notations et contexte	59
3.2.2	Loi et processus de Dirichlet	62
3.3	Apport bayésien	63
3.3.1	Modélisation bayésienne hiérarchique semi paramétrique	63
3.3.2	Quelques propriétés distributionnelles	67
3.4	Génération de la loi <i>a posteriori</i>	70
3.5	Applications	72
3.5.1	Traitement de données artificielles	72
3.5.2	Traitement des données de <i>Rubus fruticosus</i>	77
3.6	Conclusion sur le modèle spatio-temporel	79
4	Association spatiale de processus	85
4.1	Statistique d'un processus de Cox spatial associé à un processus de Dirichlet	85
4.2	Quelques propriétés distributionnelles	87
4.2.1	Cas général	87
4.2.2	Cas particuliers	92

<i>TABLE DES MATIÈRES</i>	11
4.3 Applications	93
4.3.1 Analyse des données de comptages <i>Ibicella lutea</i>	93
4.3.2 Analyse de données de décès dus à la grippe/ pneumonie et aux maladies chroniques respiratoires	95
5 Processus ponctuels en sismologie	117
5.1 Introduction	118
5.2 Processus ponctuels en sismologie	118
5.2.1 Modèle temporel	119
5.2.2 Modèle spatio-temporel marqué	120
5.2.3 Techniques de vraisemblance	121
5.3 Étude des convergences	122
5.3.1 Propriétés asymptotiques de l'intensité stochastique	122
5.3.2 Propriétés asymptotiques de la log-vraisemblance	123
5.3.3 Propriétés asymptotiques des estimateurs de maximum de vrai- semblance	124
5.4 Modèle ETAS et processus de Dirichlet	126
5.5 Étude de cas : l'arc des Petites Antilles	127
5.5.1 Présentation des données	127
5.5.2 Méthode de thinning	131
5.5.3 Étude du modèle ETAS	131
5.5.4 Discussion	132
Conclusions et perspectives	137
A Vocabulaire rencontré en sismologie	141
B Programmes informatiques complémentaires sous <i>R</i>	143
B.1 Modèle de Cox associé à un Dirichlet	143
B.1.1 Cas de données de positions spatio-temporelles	143
B.1.2 Cas de données de comptages spatiales	148
B.2 Calcul du log-rapport de vraisemblance	152
B.2.1 Calcul de la probabilité sous un processus de Dirichlet	152
B.2.2 Cas α quelconque	152
B.2.3 Cas α tendant vers 0	153
B.2.4 Cas α tendant vers $+\infty$	154

B.3	Modèle ETAS	154
B.3.1	Estimations de paramètres	154
B.3.2	Calcul de vraisemblance	156
Bibliographie		169

Notations

mathématiques

$\#(A)$	cardinal de A
\mathbb{N}	ensemble des entiers naturels
\mathbb{R}	ensemble des réels
\mathbb{R}^+	ensemble des réels positifs
$f \propto g$	la fonction f est proportionnelle à la fonction g
$\Gamma(\alpha)$	fonction Gamma ($\alpha > 0$)
$\mathbb{1}_A(x)$	fonction indicatrice (1 si $x \in A$, 0 sinon)
$\delta_x(A)$	fonction Dirac (1 si $x \in A$, 0 sinon)
$\ \cdot\ _1$	norme \mathcal{L}_1
$\xrightarrow{p.s.}$	convergence presque sûre
\xrightarrow{p}	convergence en probabilité
$\xrightarrow{\mathcal{L}}$	convergence en loi

probabilistes

iid	indépendant et identiquement distribué (pour un échantillon)
$X \sim \mathcal{L}(\lambda, \alpha)$	X suit la loi \mathcal{L} de paramètres λ et α
$\mathcal{B}(\mathbb{R}^+)$	$\{A \in \mathcal{B}(\mathbb{R}), A \subset \mathbb{R}^+\}$
$\mathcal{B}_1 \otimes \mathcal{B}_2$	$\sigma(A_1 \times A_2, A_1 \in \mathcal{B}_1, A_2 \in \mathcal{B}_2)$
$\mathbb{E}(\cdot \mathcal{F})$	espérance conditionnelle à la tribu \mathcal{F}

distributionnelles

$\mathcal{B}(n, p)$	loi binomiale
$Beta(a, b)$	loi bêta
$Dir(a_1, \dots, a_k)$	loi de Dirichlet
$Exp(\lambda)$	loi exponentielle
$Gamma(\lambda, \alpha)$	loi gamma
$\mathcal{N}(m, \sigma^2)$	loi normale unidimensionnelle
$\mathcal{P}(\lambda)$	loi de Poisson
$\mathcal{U}(S)$	loi uniforme sur S

Abréviations et sigles

AIC	: Critère d'information d'Akaike (<i>Akaike Information Criterion</i>)
ANSS	: Advanced National Seismic System
BRGM	: Bureau de Recherches Géologiques et Minières
CDSA	: Centre de Données Sismologique des Antilles
CIRAD	: Centre de coopération Internationale en Recherche Agronomique pour le Développement
CLRD	: Maladies chroniques respiratoires (<i>Chronic Lower Respiratory Disease</i>)
DAG	: Graphe Acyclique Orienté (<i>Directed Acyclic Graph</i>)
DP	: Processus de Dirichlet (<i>Dirichlet process</i>)
DPM	: Mélange de processus de Dirichlet (<i>Dirichlet Process Mixture</i>)
EMV	: Estimateur du Maximum de Vraisemblance
ETAS	: Epidemic Type Aftershock Sequence
IP	: Grippe et Pneumonie (<i>Influenza/Pneumonia</i>)
MCMC	: Monte Carlo par Chaînes de Markov (<i>Markov Chain Monte Carlo</i>)
MH	: Modèle Hiérarchique
PP	: Processus Ponctuel
PPDS	: Processus de Poisson doublement stochastique
PPH	: Processus de Poisson Homogène
PPM	: Processus Ponctuel Marqué
PPNH	: Processus de Poisson Non Homogène
PPP	: Processus Ponctuel de Poisson

PPST	:	Processus ponctuel spatio-temporel
PPSTM	:	Processus Ponctuel Spatio-Temporel Marqué
p.s.	:	presque sûrement
RCA	:	Répartition Complètement Aléatoire
RJMCMC	:	Monte Carlo par Chaînes de Markov à sauts réversibles (<i>Reversible Jump Markov Chain Monte Carlo</i>)
RPM	:	Random Probability Measure
SB	:	Stick-Breaking

Table des figures

1.1	Illustration graphique d'un PP temporel	35
1.2	Illustration du processus de branchement	42
3.1	Cartes exhaustives en temps continu : \bullet représente les occurrences à la date $t = 0$ et \square représente les occurrences dans $]0, T]$	55
3.2	Cartes exhaustives en temps continu : \bullet représente les occurrences à la date t_1 , \square : les occurrences à la date t_2 , \triangle : les occurrences à la date t_3 et ∇ : les occurrences à la date t_k	56
3.3	Comptages dans le domaine d'étude à diverses dates	57
3.4	Graphe orienté acyclique du modèle hiérarchique	64
3.5	Graphe orienté acyclique du modèle hiérarchique (cas iid)	65
3.6	Exemples d'intensité cumulée à la date $t = 1$ pour $(\alpha, b, \sigma_1, \sigma_2) = (1; 1; 0, 05; 0, 1)$: (a) deux contributions aléatoires ; (b) six contributions aléatoires ; (c) 18 contributions aléatoires ; (d) 32 contributions aléatoires.	66
3.7	Exemples de cartes de données artificielles pour $(\alpha, b, \mu, \sigma_1, \sigma_2) = (2; 20; 40; 0, 1; 0, 1)$ à 3 dates d'observation : date 1 : \bullet , date 2 : \triangle , date 3 : ∇	73
3.8	Résultats MCMC obtenus par procédure hybride Gibbs-Métropolis-Hastings : (a) pour fig. 3.7a ; (b) pour fig. 3.7b	74
3.9	Lois <i>a posteriori</i> obtenues pour les paramètres pour 50000 itérations et une période de chauffe de 5000 : (a) pour fig. 3.7a ; (b) pour fig. 3.7b	75
3.10	Nuage de points 2 à 2 des paramètres et histogramme sur la diagonale : (a) pour fig. 3.7a ; (b) pour fig. 3.7b	76
3.11	Ronces <i>Rubus fruticosus</i>	78

3.12 Répartition spatiale des ronces <i>Rubus fruticosus</i> à 3 dates d'observation : date 1 : ●, date 2 : △, date 3 : ▽	80
3.13 Résultats MCMC obtenus par procédure hybride Gibbs-Métropolis-Hastings (Ronces <i>Rubus Fruticosus</i>)	81
3.14 Lois <i>a posteriori</i> obtenues pour les paramètres pour 50000 itérations et une période de chauffe de 5000 (Ronces <i>Rubus Fruticosus</i>)	82
3.15 Nuage de points 2 à 2 des paramètres et histogramme sur la diagonale pour les données de Ronces <i>Rubus Fruticosus</i>	83
4.1 Plante <i>Ibicella Lutea</i>	94
4.2 Grille régulière 8 × 8 superposée sur la distribution spatiale de <i>Ibicella lutea</i>	96
4.3 Grille régulière 16 × 16 superposée sur la distribution spatiale de <i>Ibicella lutea</i>	97
4.4 Grille régulière 32 × 32 superposée sur la distribution spatiale de <i>Ibicella lutea</i>	98
4.5 Résultats MCMC obtenus par procédure hybride Gibbs-Métropolis-Hastings pour la grille de taille 8 × 8 de <i>Ibicella lutea</i>	100
4.6 Résultats MCMC obtenus par procédure hybride Gibbs-Métropolis-Hastings pour la grille de taille 16 × 16 de <i>Ibicella lutea</i>	101
4.7 Résultats MCMC obtenus par procédure hybride Gibbs-Métropolis-Hastings pour la grille de taille 32 × 32 de <i>Ibicella lutea</i>	102
4.8 Lois <i>a posteriori</i> obtenues pour les paramètres pour 50000 itérations et une période de chauffe de, respectivement, 5000 pour (a) et 30000 pour (b) (grille de taille 8 × 8 et les données de <i>Ibicella lutea</i>)	103
4.9 Lois <i>a posteriori</i> obtenues pour les paramètres pour 50000 itérations et une période de chauffe de, respectivement, 5000 pour (a) et 10000 pour (b) (grille de taille 16 × 16 et les données de <i>Ibicella lutea</i>)	104
4.10 Lois <i>a posteriori</i> obtenues pour les paramètres pour 50000 itérations et une période de chauffe de 5000 (grille de taille 32 × 32 et les données de <i>Ibicella lutea</i>)	105
4.11 Nuage de points 2 à 2 des paramètres et histogramme sur la diagonale (grille de taille 8 × 8, données de <i>Ibicella lutea</i> , période de chauffe de, respectivement, 5000 pour (a) et 30000 pour (b))	106

4.12	Nuage de points 2 à 2 des paramètres et histogramme sur la diagonale (grille de taille 16×16 , données de <i>Ibicella lutea</i> , période de chauffe de, respectivement, 5000 pour (a) et 10000 pour (b))	107
4.13	Nuage de points 2 à 2 des paramètres et histogramme sur la diagonale (grille de taille 32×32 , données de <i>Ibicella lutea</i>)	108
4.14	Lien de voisinage et nombre de décès par canton pour l'état de Georgia, USA en 2007 pour : (a) IP ; (b) CLRD.	111
4.15	Résultats MCMC obtenus par procédure hybride Gibbs-Métropolis-Hastings pour les données de : (a) IP ; (b) CLRD.	113
4.16	Lois <i>a posteriori</i> obtenues pour les paramètres pour 50000 itérations et une période de chauffe de 5000 pour : (a) IP ; (b) CLRD.	114
4.17	Nuage de points 2 à 2 des paramètres et histogramme sur la diagonale pour : (a) IP ; (b) CLRD.	115
5.1	Répartition spatiale des séismes des Petites Antilles de 1999 à 2004. Les points rouges correspondent aux magnitudes supérieures à 4, 1 et les bleus ceux de magnitudes inférieures.	128
5.2	Distribution temporelle des séismes des Petites Antilles de 1999 à 2004 de magnitude supérieure à 3 et arrondie à l'entier supérieur. Le jour numéro 1 correspond au premier Janvier 1999.	129
5.3	Distribution temporelle des séismes par rapport à leur magnitude. Le jour numéro 1 correspond au premier Janvier 1999.	130
5.4	Densité des distributions Gamma (noir), Weibull (bleu), Lognormal (vert) et loi d'Omori modifiée (rouge) sur l'intervalle $[0, 1]$ (valeurs de c et p issues de la table 5.1).	133
5.5	Analyse résiduelle des événements sismiques avec comme fonction déclenchante la densité de Weibull : (a) Courbe de l'intensité cumulée en rouge, droite d'équation $y=x$ en noir ; (b) Histogramme des temps d'attente rééchelonnés, densité exponentielle de paramètre 1 en bleu.	134

Liste des Algorithmes

1	Algorithme d'acceptation-rejet	44
2	Algorithme de Metropolis-Hastings	47
3	Algorithme de Gibbs	49
4	RJMCMC pour l'inférence sur $(\alpha, b, \mu, \sigma_1, \sigma_2)$ du modèle (3.11)	71

Introduction

Contexte de la thèse

Dans de nombreuses disciplines, les scientifiques sont confrontés à l'étude de phénomènes aléatoires. Quand tel est le cas, des techniques statistiques appropriées doivent être employées. Dans cette thèse, nous nous intéressons à l'étude statistique d'occurrences d'événements se répartissant spatialement et/ou temporellement. Les processus ponctuels sont des outils de modélisation des répartitions d'occurrences d'événements souvent utilisés dans divers domaines tels la sismologie (Ogata [65, 66], 1988, 1998 ; Zhuang et al. [114], 2005), l'écologie (Vaillant [93], 1992), l'épidémiologie (Vaillant [91], 1991 ; Vaillant et al. [96], 2011), la neurophysiologie (Truccolo et al. [90], 2005 ; Beneš et al. [6], 2008). Ainsi, de nombreux auteurs se sont penchés par exemple sur l'étude des phénomènes sismiques d'un point de vue mathématique afin d'analyser leur répartition en fonction de leur localisation, leur date d'occurrence et/ou leur magnitude (Ogata [65, 66], 1988, 1998 ; Kagan [48], 2004 ; Zhuang et al. [114], 2005). Une contribution plus théorique sur les processus ponctuels est fournie par Karr ([49], 1991) ainsi que Daley et Vere-Jones ([22], 2003). En réalité, la réalisation d'un processus ponctuel spatio-temporel est un ensemble aléatoire de points, où chaque point représente la date et/ou la localisation d'un événement (Karr [49], 1991). Notons les exemples suivants :

- les instants d'arrivée de parasites sur une culture constituent la réalisation d'un processus ponctuel temporel,
- les positions de plantules dans une région constituent la réalisation d'un processus ponctuel spatial,
- la localisation d'épicentres et la date d'occurrence du séisme dans une région

au cours d'une période constituent la réalisation d'un processus ponctuel spatio-temporel.

Un processus ponctuel est en fait une mesure stochastique de dénombrement sur un espace métrique qui modélise des localisations de points dans un espace donné (Daley et Veres-Jones [22], 2003). Sous des conditions de régularité usuelle, la loi d'un processus ponctuel est complètement caractérisée par son processus intensité conditionnelle $\lambda(\cdot)$ (Cox et Isham [19], 1980 ; Daley et Vere-Jones [22], 2003 ; Karr [49], 1991 ; Tanner [87], 1996). La modélisation se fait alors uniquement à travers ce dernier. Un exemple trivial est le processus de Poisson pour lequel l'intensité conditionnelle $\lambda(\cdot)$ a une valeur constante égale au nombre d'événements espéré par unité de temps et/ou d'espace. Toutefois, la plupart des phénomènes ont une intensité variant au cours du temps et/ou dans l'espace et présentant de l'autocorrélation (Vaillant [92], 1992). Il faut par conséquent prendre en compte les caractéristiques du phénomène considéré et s'intéresser à un modèle adéquat. Par exemple, certains auteurs s'intéressent aux processus de Cox (Brémaud [9], 1981 ; Brix et Diggle [13], 2001) et d'autres aux processus à mémoire tels que les processus de Hawkes (Hawkes [40], 1971 ; Ogata [65], 1988 ; Puri et Tuan [74], 1986).

Il faut noter qu'en général, $\lambda(\cdot)$ dépend également des événements antérieurs. Soit une filtration $(\mathcal{H}_t)_{t \geq 0}$, on pourra définir un processus ponctuel spatio-temporel par son intensité stochastique conditionnelle $\lambda(x, t | H_{t-})$ qui désigne de façon heuristique le taux espéré d'occurrences d'événements dans le voisinage de x conditionnellement à l'histoire du processus jusqu'à la date t exclue, c'est-à-dire H_{t-} . Le processus intensité dépend en général de covariables pouvant être observées au moyen d'un plan d'échantillonnage spatial. Cependant, conditionnellement aux covariables observées, une surdispersion peut être malgré tout présente, ce qui signifie la présence d'effets aléatoires dus à des covariables inconnues ou insoupçonnées. Suite aux travaux de Brix ([11], 1999), Brix et Chadœuf ([12], 2002), Møller ([58], 2003) sur les processus de Cox shot noise et ceux de Brix et Diggle ([13], 2001) sur les processus de Cox log-gaussien, il était intéressant d'étudier l'apport d'un processus de Dirichlet (Ferguson [32], 1973) dans un modèle shot noise généralisé (Møller et Torrisi [61], 2005 ; Duan et al. [29], 2007) afin de décrire les éventuelles corrélations entre facteurs environnementaux influençant l'intensité du processus ponctuel observé. Les travaux présentés au chapitre 3 concernent donc la situation où un processus ponctuel caché a des répercussions sur l'intensité du processus observé. L'intégration dans notre modèle de processus ponctuel des connaissances *a priori* sur

des variables de l'environnement, conduit à une approche bayésienne (Droesbeke et al. [28], 2002). La méthodologie inférentielle bayésienne est appliquée pour prendre en compte des effets présumés des variables environnementales à partir d'un modèle hiérarchique englobant les informations disponibles. Nous développons une méthode MCMC à sauts réversibles pour tenir compte du fait que le nombre de contributions cachées est variable dans le modèle, comme recommandé par Green et Richardson ([38], 2002).

Un deuxième volet de la thèse concerne l'étude des phénomènes sismiques et plus précisément les probabilités d'occurrences de tremblements de terre. En statistique sismologique, l'outil principal de modélisation est la théorie des processus ponctuels qui permet, pour une date donnée, d'intégrer l'information jusqu'à cette date (exclue) dans l'expression du risque conditionnel (Vere-Jones [107], 1995; Ogata [65, 66], 1988, 1998; Zhuang et al. [113], 2008; Schœnberg [82], 2003). Les techniques d'inférence statistique sont basées sur la vraisemblance des observations (Ogata [64], 1978; Rathbun [76], 1996; Puri et Tuan [74], 1986). Notons que l'analyse résiduelle des processus ponctuels, qui consiste à transformer le processus observé afin de le comparer à un processus de Poisson d'intensité unité, est souvent utilisée en statistique sismologique. Cet outil a été développé par Ogata ([65], 1988) pour les processus temporels, par Baddeley et al. ([4], 2005) pour les processus spatiaux, par Schœnberg ([82], 2003) pour les processus multidimensionnels et par Zhuang ([112], 2006) et par Ogata et al. ([69], 2003) pour les processus spatio-temporels. Nous avons étudié un modèle ETAS en introduisant une intensité d'arrière-plan qui suit un processus de Dirichlet centré sur un processus spatial log-normal. Des données sismiques de l'arc des Petites Antilles sont traitées et une comparaison de différentes fonctions déclenchantes est effectuée.

Contributions et résultats

Un premier volet de cette thèse consiste en l'association d'un processus de Cox shot noise généralisé avec un processus de Dirichlet. Ce processus de Dirichlet modélise les effets aléatoires sur l'intensité spatio-temporelle dont l'expression est :

$$\forall (x, t) \in X \times \mathbb{R}_+, \quad \lambda(x, t) = \sum_{j=1}^{M_t-} a_j K_{(\sigma_1, \sigma_2)}((x, t), (y_j, s_j)) \quad (1)$$

où les (y_j, s_j) forment la réalisation d'un processus de Poisson stationnaire d'intensité μ sur $X \times \mathbb{R}_+$ et les a_j celle d'un processus de Dirichlet centré sur une loi à support inclus

dans \mathbb{R}_+ , d'espérance b et de paramètre de concentration α . D'autre part, $K_{(\sigma_1, \sigma_2)}$ est un noyau spatio-temporel de bande passante spatiale σ_1 associé à une loi gaussienne et de bande passante temporelle σ_2 associé à une loi exponentielle. Les éléments inconnus intervenant dans (1) sont : $\theta^* = (\sigma_1, \sigma_2, M_{t-}, (y_i, s_i, a_i)_{i \in \llbracket 1, M_{t-} \rrbracket})$ où M_{t-} est le nombre de contributions jusqu'à la date t exclue. On a les résultats distributionnels suivants :

$\forall (x, t) \in X \times \mathbb{R}_+$, l'intensité cumulée en x à la date t est

$$\int_0^t \lambda(x, s) ds = \sum_{j=1}^{M_t} a_j \frac{1}{(2\pi\sigma_1^2)^{d/2}} \exp\left(-\frac{\|x - y_j\|^2}{2\sigma_1^2}\right) \left(1 - \exp\left(-\frac{t - s_j}{\sigma_2}\right)\right)$$

où $M_t = \sum_j \delta_{s_j}([0, t])$ suit une loi de Poisson de paramètre $\mu t \nu(X)$.

$\forall (B, t) \in \mathcal{B}_X \times \mathbb{R}_+$, $\Lambda(B \times [0, t]) = \int_B \int_0^t \lambda(x, s) dx ds$ est une variable aléatoire correspondant à l'intensité cumulée sur B à la date t , d'espérance

$$E(\Lambda(B \times [0, t])) = \mu b t \left(1 - \frac{\sigma_2}{t} \left(1 - \exp\left(-\frac{t}{\sigma_2}\right)\right)\right) \int_X \int_B \frac{\exp\left(-\frac{\|x-y\|^2}{2\sigma_1^2}\right)}{(2\pi\sigma_1^2)^{d/2}} \nu(dx) \nu(dy)$$

et de variance

$$\begin{aligned} V(\Lambda(B \times [0, t])) &= \mu b (1 + b) \int_0^t \left(1 - \exp\left(-\frac{t-s}{\sigma_2}\right)\right)^2 ds \int_X \left(\int_B \frac{\exp\left(-\frac{\|x-y\|^2}{2\sigma_1^2}\right)}{(2\pi\sigma_1^2)^{d/2}} \nu(dx)\right)^2 \nu(dy) \\ &\quad + \frac{1}{b(\alpha + 1)} (E(\Lambda_\theta(B \times [0, t])))^2. \end{aligned}$$

$$\forall(x, t) \in X \times \mathbb{R}_+^*,$$

$$\begin{aligned} \lambda(x, t) &\xrightarrow[\sigma_1 \rightarrow 0]{\mathcal{L}} \sum_{j=1}^{M_{t^-}} a_j \delta_{y_j}(\{x\}) \times \frac{1}{\sigma_2} \exp\left(-\frac{t-s_j}{\sigma_2}\right) \mathbb{1}_{[0,t]}(s_j) \\ \lambda(x, t) &\xrightarrow[\sigma_2 \rightarrow 0]{\mathcal{L}} \sum_{j=1}^{M_{t^-}} a_j \frac{\exp\left(-\frac{\|x-y_j\|^2}{2\sigma_1^2}\right)}{(2\pi\sigma_1^2)^{d/2}} \times \delta_{s_j}(\{t\}) \\ \lambda(x, t) &\xrightarrow[(\sigma_1, \sigma_2) \rightarrow (0,0)]{\mathcal{L}} \sum_{j=1}^{M_{t^-}} a_j \times \delta_{(y_j, s_j)}(\{(x, t)\}) \end{aligned}$$

Pour le traitement statistique de ce modèle, nous considérons la situation où nous disposons de k cartes exhaustives à des dates d'observations échelonnées dans le temps. En d'autres termes, nous supposons que toutes les positions spatiales des occurrences sont disponibles mais que l'information sur les dates d'occurrence est sous la forme d'appartenance à des intervalles dont les bornes sont des dates d'observation consécutives. Soient t_1, t_2, \dots, t_k les k dates d'observations d'un processus de Cox N de loi définie par (1), posons $N_i = N(X \times [0, t_i])$ alors la vraisemblance des k cartes d'occurrence est :

$$\mathcal{L}(\theta^*, (x_j)) = \prod_{i=1}^k \prod_{N_{i-1}+1}^{N_i} \int_{t_{i-1}}^{t_i} e^{-\Lambda(X \times [0, t_k])} \lambda(x_j, t) dt$$

et, pour des lois *a priori* de μ, α, σ_1 et σ_2 indépendantes, la loi *a posteriori* de $\theta = (\mu, \alpha, b, \sigma_1, \sigma_2)$ est proportionnelle à :

$$p(\mu)p(\alpha)p(\sigma_1)p(\sigma_2)p(M_{t_k}|\mu)p(a_1, \dots, a_{M_{t_k}}|\alpha, b, M_{t_k})\mathcal{L}(\theta^*, (x_j))$$

où $p(a_1, \dots, a_{M_{t_k}}|\alpha, b, M_{t_k})$ est donné par le théorème de Finetti.

L'application à la répartition spatio-temporelle des données de *Rubus fruticosus* (ronces communes), pour laquelle les événements cachés sont les absorptions et déjections de semence par certains mammifères et oiseaux, nous a fourni les résultats suivants :

l'estimateur bayésien du nombre espéré d'événements cachés est proche de 26 avec une contribution estimée de 46,7 et un paramètre de confiance proche de 4. Le paramètre d'influence spatial est estimé à 0,074, ce qui correspond à une agrégation spatiale modérée mais plus prononcée que l'agrégation temporelle qui est estimée à

1, 170. Nous avons effectué un test du rapport de vraisemblance pour modèles emboîtés (Spiegelhalter et al. [84], 2002) appliqué à l'hypothèse H_0 selon laquelle les contributions cachées sont indépendantes de même loi Gamma contre l'hypothèse selon laquelle elles forment une réalisation d'un processus de Dirichlet centré sur une loi Gamma. Le rapport de vraisemblance obtenu conduit au non rejet de H_0 , ceci signifiant que les événements cachés ont des impacts indépendants sur l'intensité des ronces *Rubus fruticosus*.

D'autre part, nous considérons également la situation où nous avons des données de comptages d'occurrences dans des unités spatiales. Ces situations se retrouvent fréquemment dans des domaines tels que la sismologie, l'écologie, l'épidémiologie et la neurophysiologie.

L'intensité spatiale considérée est :

$$\forall x \in X, \quad \lambda(x) = \sum_{j=1}^M a_j K_\sigma(x, y_j) \quad (2)$$

avec

$$\begin{aligned} (a_j) \mid G &\stackrel{i.i.d}{\sim} G \\ (y_j) &\sim HPP(\mu) \\ G &\sim DP(\alpha, G_0). \end{aligned} \quad (3)$$

L'expression (2) dépend des inconnus $\sigma, M, (a_i)_{i \in [1, M]}$.

Quelques propriétés distributionnelles pour un processus de Cox N sur un espace mesuré (X, \mathcal{B}, ν) et dont l'intensité est définie par les équations (2) et (3) sont alors

$$\forall B \in \mathcal{B},$$

$$E(N(B)) = \mu E(a_1) \int_X \int_B K_\sigma(x, y) \nu(dx) \nu(dy)$$

et

$$V(N(B)) = E(N(B)) \left(1 - \frac{1}{\alpha + 1} E(N(B)) \right) + \mu E(a_1^2) \int_X \left(\int_B K_\sigma(x, y) \nu(dx) \right)^2 \nu(dy).$$

$$\forall (B_1, B_2) \in \mathcal{B}^2 \text{ avec } B_1 \cap B_2 = \emptyset,$$

$$\begin{aligned} Cov(N(B_1), N(B_2)) &= \mu E(a_1^2) \int_X \left(\int_{B_1} K_\sigma(z, y) \nu(dz) \int_{B_2} K_\sigma(x, y) \nu(dx) \right) \nu(dy) \\ &\quad - \frac{1}{\alpha + 1} E(N(B_1)) E(N(B_2)). \end{aligned}$$

On démontre que le rapport variance-espérance $\frac{V(N(B))}{E(N(B))}$ est une fonction croissante de α vérifiant $\lim_{\sigma \rightarrow 0} \frac{V(N(B))}{E(N(B))} = 1 - \frac{\mu E(a_1)}{\alpha + 1} \nu(B) + \frac{E(a_1^2)}{E(a_1)}$ et que pour B et B' éléments disjoints de \mathcal{B} , nous avons $\lim_{\sigma \rightarrow 0} Cov(N(B), N(B')) = \mu^2 (E(a_1))^2 \frac{\alpha}{\alpha + 1} \nu(B) \nu(B')$.

Les données de comptage d'*Ibicella lutea* sur une grille régulière 32×32 ont été traitées avec notre méthode. Les événements cachés concernent la chute et la dispersion par des animaux de graines issues de fruits pourris. Les résultats d'inférence bayésienne effectuée sont les suivants : le nombre moyen d'événements cachés varie entre 26 et 27, celui de la contribution à l'intensité, entre 1 et 2. Les estimations du paramètre d'influence sont entre 0,06 et 0,071 et du paramètre de concentration, entre 17 et 29.

Nous avons également traité des données de mortalité dus à la grippe/pneumonie et aux maladies respiratoires au cours de l'année 2007 dans l'état de Georgia, USA. Ces données sont sous forme de comptages pour chacun des 159 cantons de cet état. Les événements cachés concernent le tabagisme, les changements brutaux de température. Les résultats bayésiens ont permis de montrer que les contributions espérées sont significativement différentes pour les deux maladies ainsi que pour le paramètre d'influence spatiale.

Un échantillonneur de Gibbs et un algorithme à sauts réversibles ont été mis au point pour mener à bien cette inférence bayésienne.

Dans un second temps, nous nous sommes intéressés à l'analyse statistique de processus ponctuels appliqués à la sismologie (Ogata [65], 1988 ; Vere-Jones [107, ?], 1995, 2010 ; Vere-Jones et al. [108], 2001) et au modèle ETAS (Epidemic Type Aftershock Sequence) (Ogata [66], 1998 ; Zhuang et al. [114], 2005). Nous proposons de considérer le modèle d'intensité conditionnelle spatio-temporelle telle que proposée par Ogata ([68], 2011) en y ajoutant une intensité d'arrière-plan, non plus déterministe, mais de loi un processus de Dirichlet centré sur un processus spatial log-gaussien. Ensuite l'analyse de la distribution des épacentres et des magnitudes des séismes ayant eu lieu dans les Petites Antilles de 1999 à 2004 a été effectuée, en utilisant notre modèle ETAS généralisé. D'autre part, une comparaison de différentes fonctions déclenchantes est menée en programmant dans l'environnement **R**.

Organisation du manuscrit

Ce document est organisé en six chapitres : dans les deux premiers chapitres, nous présentons quelques notions essentielles à la compréhension des modèles et procédures proposés aux chapitres 3 et 4.

Au chapitre 1, nous exposons quelques éléments de la théorie des processus ponctuels ainsi que quelques outils permettant leur analyse statistique.

Le chapitre 2 fournit quelques concepts de la démarche bayésienne et des méthodes de Monte Carlo. Les procédures MCMC et d'échantillonnage dont le but est de calculer ou d'approcher des lois *a posteriori* sont présentées et discutées.

Le chapitre 3 est consacré au développement d'un modèle spatio-temporel de processus de Cox "shot noise généralisé". Nous considérons des situations pour lesquelles il existe un processus caché contribuant à des effets aléatoires sur l'intensité spatio-temporelle du processus de Cox observé. Le processus caché est un processus de Poisson associé à un processus de Dirichlet. L'écriture de la vraisemblance pour des données du type carte complète des occurrences spatiales à des dates d'observations nous permet de développer l'inférence bayésienne. Une application en écologie est présentée.

Le chapitre 4 concerne l'approche purement spatiale du modèle de processus spatio-temporel présenté au chapitre 3. Les propriétés distributionnelles de comptages dans des unités d'échantillonnage spatiales sont présentées et une inférence bayésienne est proposée pour estimer les paramètres d'intérêt du modèle et tester l'indépendance des contributions ou leur égalité.

Le chapitre 5 est dédié à l'application des processus ponctuels en sismologie et à la modélisation stochastique décrivant au mieux l'activité sismique. Le modèle ETAS spatio-temporel est étendu en y intégrant une intensité d'arrière-plan suivant un processus de Dirichlet centré sur un processus spatial log-gaussien. Une étude comparative de fonctions déclenchantes est effectuée pour l'arc des Petites Antilles en utilisant des données sismiques datant de la période allant de 1999 à 2004.

Ce document se termine par un chapitre consacré aux conclusions et perspectives de

ce travail de thèse et est complété par différentes annexes. Ce dernier chapitre présente différentes perspectives, dont celle concernant une étude approfondie d'un modèle ETAS étendu du fait de considérer une intensité d'arrière-plan appartenant à une classe de processus spatio-temporels semi-paramétriques.

ÉLÉMENTS SUR LES PROCESSUS PONCTUELS

1.1 Quelques notions essentielles

La théorie des processus ponctuels fait appel tout d'abord à des notions de théorie de la mesure. En effet, un processus ponctuel peut être considéré comme une mesure stochastique à valeurs entières. Il sert à décrire la répartition d'occurrences sur un espace métrique (X, d) . L'appel à un espace métrique est dû à la nécessité de considérer les proximités entre occurrences.

Définition 1.1.1 *Un processus ponctuel N sur un espace probabilisé (Ω, \mathcal{A}, P) est une application mesurable à valeurs dans \mathcal{M}_0 , l'ensemble des mesures sur X à valeurs entières, définie par :*

$$\begin{aligned} N : \quad \Omega &\rightarrow \mathcal{M}_0 \\ \omega &\mapsto \sum_i \delta_{X_i(\omega)} \end{aligned} \tag{1.1}$$

On peut écrire :

$$N = \sum_i \delta_{X_i}$$

où $(X_i)_{i \in I}$ est une famille de variables aléatoires à valeurs dans (X, d) .

La littérature porte souvent sur des occurrences réparties dans le temps (Andersen et al., [3], 1993 ; Daley et Vere-Jones, [22], 2003). L'évolution chronologique impose

l'utilisation d'un espace filtré, c'est-à-dire un quadruplet $(\Omega, \mathcal{A}, P, (\mathcal{F}_t)_{t \geq 0})$ où (Ω, \mathcal{A}, P) est un espace probabilisé et $(\mathcal{F}_t)_{t \geq 0}$ est une famille de sous-tribus de \mathcal{A} vérifiant :

$$\forall t, s < t, \mathcal{F}_s \subset \mathcal{F}_t.$$

On note :

$$\mathcal{F}_{t-} = \bigcap_{s < t} \mathcal{F}_s$$

pour traduire l'information (ou l'histoire) jusqu'à la date t exclue. Ceci permet d'introduire la notion de processus prévisible.

Définition 1.1.2 *Étant donné une filtration $\mathbf{F}=(\mathcal{F}_t)$ sur (Ω, \mathcal{A}, P) , un processus (X_t) de fonctions mesurables est \mathbf{F} -prévisible si, pour tout $t \in \mathbb{R}_+$, X_t est \mathcal{F}_{t-} -mesurable.*

Définition 1.1.3 *Un processus ponctuel N est dit simple si au plus une occurrence se produit à un instant donné :*

$$\forall x \in X, N(\{x\}) \in \{0; 1\}.$$

Définition 1.1.4 *Soit le processus ponctuel temporel*

$$N = \sum_{i=1}^{\infty} \delta_{X_i}$$

où les X_i sont des variables aléatoires positives telles que $X_1 \leq \dots \leq X_i \leq \dots$.

Le processus de comptage $(N_t)_{t \geq 0}$ associé à N est défini par :

$$N_t = \sum_{i=1}^{\infty} 1_{\{X_i \leq t\}}.$$

Soit $\mathbf{F}=(\mathcal{F}_t)_{t \geq 0}$ une filtration telle que (N_t) soit \mathbf{F} -adapté.

Le compensateur de N relatif à \mathbf{F} est l'unique processus $A = (A_t)$ vérifiant :

1. A est \mathbf{F} -prévisible
2. $M_t = N_t - A_t$ est une martingale centrée.

S'il existe un processus positif prévisible (λ_t) tel que

$$dA_t = E(dN_t | \mathcal{F}_{t-}) = \lambda_t dt$$

alors

$$A_t = \int_0^t \lambda_s ds$$

et $\lambda = (\lambda_t)$ est par définition la \mathbf{F} -intensité stochastique de N .

Notons qu'on peut associer des marques à un processus ponctuel (Holden et al., [41], 2003). Ces marques peuvent être qualitatives ou quantitatives. Par exemple, on peut décrire la succession de dates de séismes dans une zone géographique par un processus ponctuel temporel dont les marques seraient quantitatives (magnitude, coordonnées de l'épicentre) ou qualitatives (cause). Une des particularités d'un processus ponctuel temporel est qu'il peut être décrit par des dates d'occurrence X_i mais aussi par des durées d'attente $U_i = X_i - X_{i-1}$. Quand il est simple, la fonction en escalier croissante de saut 1, représentant le processus de comptage observé, décrit parfaitement la réalisation de ce processus ponctuel temporel (figure 1.1).

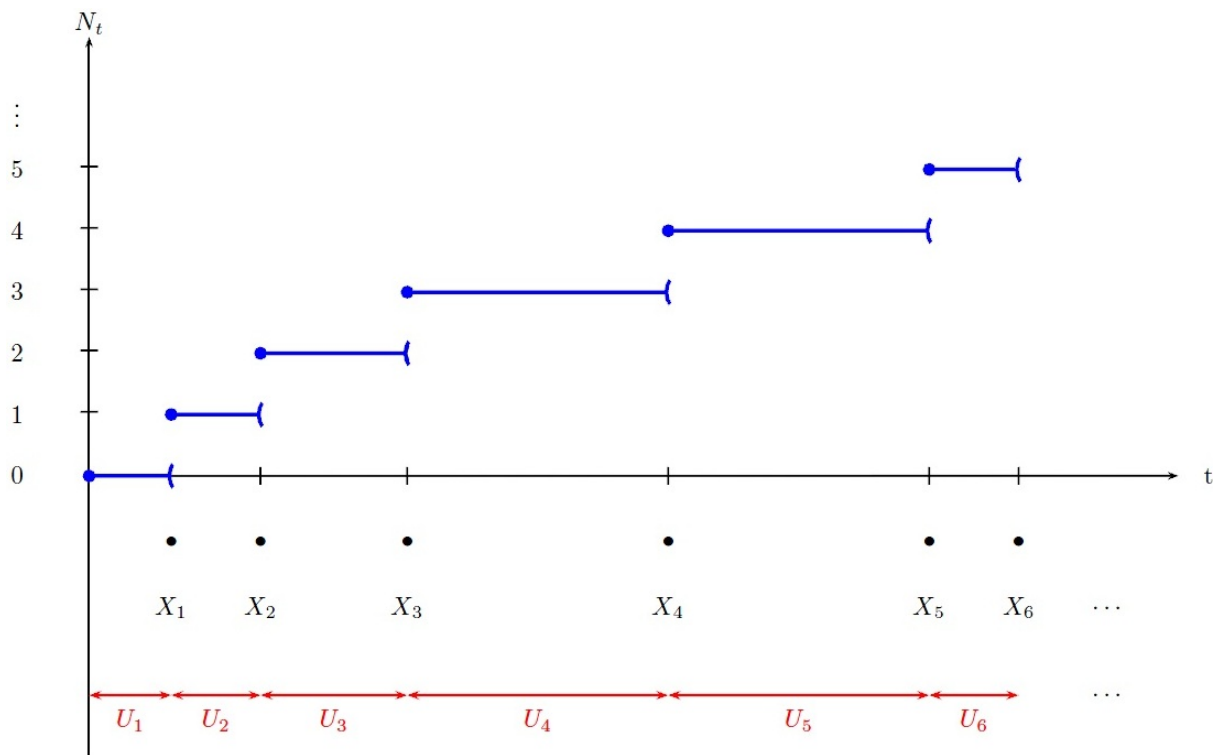


FIGURE 1.1 – Illustration graphique d'un PP temporel

1.2 Processus de vraisemblance d'un processus ponctuel

La vraisemblance d'un processus ponctuel est un processus stochastique discuté dans des ouvrages tels que Andersen et al. ([3], 1993), Daley et Vere-Jones ([21], 1988, chapitre 13).

Dans un espace mesuré (X, \mathcal{A}, ν) , pour des données observées de la forme $(x_i, t_i)_{i \in [1, n]}$ correspondant à la réalisation d'un processus ponctuel spatio-temporel simple N dans X au cours d'une période $[0, T]$, la log-vraisemblance est :

$$\mathcal{L}_T(\theta; (x_i, t_i)) = \sum_{0 < t_j \leq T} \log \lambda_\theta(x_j, t_j | H_{t_j^-}) - \int_0^T \int_X \lambda_\theta(x, t | H_{t^-}) \nu(dx) dt \quad (1.2)$$

où $\lambda_\theta(x, t | H_{t^-})$ est l'intensité du processus en (x, t) et $(H_t)_{t \geq 0}$ est une filtration fournissant l'historique des observations à toute date, θ est le paramètre, souvent un vecteur de $\mathbb{R}^k (k \in \mathbb{N}^*)$, sur lequel porte l'inférence statistique .

Les estimateurs du maximum de vraisemblance pour les éléments de θ sont obtenus en maximisant l'expression (1.2). Ils possèdent généralement certaines propriétés intéressantes, comme la consistance, le non-biais, la normalité asymptotique et l'efficacité (Ogata, [64], 1978; Rathbun, [76], 1996; Jensen et Petersen, [44], 1999).

Notons que l'expression pour la log-vraisemblance du processus n'est pas la même si l'on n'observe pas précisément les dates d'occurrences t_i comme nous le verrons au chapitre 3.

L'utilité des techniques de vraisemblance est également dans la possibilité de tester l'adéquation de certains modèles. Ainsi dans le cas de modèles emboîtés, la méthode du rapport de vraisemblance aboutit à la construction d'un test permettant d'accepter un modèle particulier dans une classe plus générale de modèles (Vaillant, [92], 1992; Paninski, [72], 2004; Møller et Diaz-Avalos, [59], 2010).

1.3 Analyse résiduelle d'un processus ponctuel

Les résidus, définis comme étant la différence entre les observations et les valeurs attendues sous un modèle, sont utilisés pour déterminer la qualité de l'ajustement au modèle en question. Dans le contexte des processus ponctuels, la notion de résidus d'un processus ponctuel peut-être consultée dans Baddeley et al. ([4], 2005). Dans cet

article, il fournit une généralisation aux processus ponctuels spatiaux de la notion de résidu pour les processus ponctuels temporels. En effet, cette dernière est largement appliquée en analyse de survie alors que l'absence d'ordre naturel dans l'espace a gêné sa généralisation.

Soient $(t_i)_{i \in [1, n]}$ les dates d'occurrences d'un processus ponctuel temporel, le principe de l'analyse résiduelle pour un processus ponctuel temporel consiste à calculer l'estimation de l'intensité cumulée en t_i , notée $\widehat{\Lambda}(t_i)$. En posant $\tau_i = \widehat{\Lambda}(t_i)$, on a la réalisation d'un processus de Poisson d'intensité 1 sous l'hypothèse que le modèle de processus ponctuel considéré est d'intensité cumulée $\widehat{\Lambda}$.

Si $(N_t)_{t \geq 0}$ désigne le processus de comptage associé au processus ponctuel temporel défini par un processus d'intensité conditionnelle $(\lambda(t))$ alors le processus résiduel peut être construit à partir du fait que le processus innovation $(I(t))_{t \geq 0}$ défini comme suit $I(t) = N_t - \int_0^t \lambda(s) ds$ est une martingale d'espérance zéro dès lors que le modèle est vrai. Le processus résiduel est alors :

$$R(t) = N_t - \widehat{\Lambda}(t).$$

La qualité de l'adéquation des données au modèle est vérifiée en étudiant le degré de proximité des résidus avec zéro.

Notons que l'analyse résiduelle des processus ponctuels a été étudiée également dans le cas des processus multidimensionnels (Schœnberg, [82], 2003 ; Zhuang, [112], 2006) et pour les processus spatio-temporels (Ogata et al. ([69], 2003).

1.4 Analyse d'autocorrélation et de dispersion

L'analyse statistique d'une réalisation d'un processus ponctuel N sur un espace métrique (X, d) dépend de la méthode d'observation. On peut en effet être confronté à plusieurs situations selon les contraintes techniques et le coût d'échantillonnage : disposer des positions des occurrences dans X de façon exhaustive, disposer d'un échantillon de comptage des occurrences ou des mesures de distance entre occurrences et points d'échantillonnage (Diggle [24, 25], 1983 et 2005 ; Diggle et Milne [27], 1983 ; Vaillant et Derridj [95] ; Melville et Welsh [57], 2001 ; Fewster et al. [34], 2005). Une présentation synthétique de ces méthodes a été effectuée par Perry et al. ([73], 2006) et une étude comparative est fournie dans cet article pour des applications en écologie quanti-

tative. Notons que ces différentes procédures d'analyse d'une répartition de points sont généralement basées sur les propriétés de premier ordre et second ordre d'un processus ponctuel stationnaire et parfois isotropique (Cliff et Ord, [16], 1981). Ces derniers auteurs présentent différentes mesures d'autocorrélation spatiale. Le principe est, par analogie avec les séries chronologiques (Gouriéroux et Monfort, [36], 1995), de calculer le degré de similarité pour une variable statistique donnée pour des unités spatiales proches. Un indice bien connu est l'indice de Moran dont la formule est la suivante :

$$I_m = \frac{n \sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{S_0 \sum_i (x_i - \bar{x})^2}$$

où w_{ij} est la pondération associée aux unités spatiales i et j , $S_0 = \sum_i \sum_j w_{ij}$ est la somme de ces pondérations et x_i est la valeur observée sur l'unité i .

Cet indice permet de vérifier si des données géo-référencées sont significativement autocorrélées, le test correspondant peut être consulté dans Cliff et Ord ([16], 1981). Notons que cet indice de Moran peut être généralisé par un indice d'autocorrélation spatio-temporelle (Vaillant et al., [96], 2011).

Un autre aspect des configurations spatiales de points est le degré de variabilité des comptages de points dans des unités spatiales de même taille et de même forme par rapport à celles attendues sous une répartition complètement aléatoire (RCA). Cette dernière répartition correspond à la réalisation d'un processus de Poisson homogène, dont nous précisons la définition au paragraphe 1.5.2. La loi de dénombrement associée à une RCA est la loi de Poisson dont l'espérance est égale à la variance. Ceci explique l'utilisation du rapport variance moyenne comme outil de détection de l'écart à une RCA. Soit N_1, N_2, \dots, N_n les comptages de points de la réalisation effectuée dans n unités spatiales, l'indice de dispersion observée est :

$$\hat{I}_d = \frac{\sum_{i=1}^n (N_i - \bar{N})^2}{(n-1)\bar{N}}$$

où \bar{N} est la moyenne arithmétique des N_i .

Taylor ([88], 1984) a souligné l'intérêt théorique et pratique de cet indice, en écologie en particulier, en tant que mesure du degré de dispersion entre individus écologiques.

Sous l'hypothèse de processus de Poisson homogène, l'indice de dispersion théorique I_d vaut 1 alors que pour des répartitions de type agrégatif ou des répartitions en milieu hétérogène, I_d est strictement supérieure à 1. Quand tel est le cas, on parle de surdispersion des points ou de points surdispersés. Ceci se produit par exemple en écologie quand des individus ont des affinités mutuelles ou des affinités communes avec des variables environnementales. Par contre quand I_d est strictement inférieure à 1, on parle de sous-dispersion ou d'individus sous-dispersés. En écologie, ce phénomène se produit généralement quand il y a compétition territoriale et comportement individualiste, ce qui fait que les individus sont rarement regroupés. Notons à ce sujet qu'il existe d'autres indices mesurant le degré de dispersion, tel que l'indice d'attroupement (Taylor, [88], 1984). \widehat{I}_d est utilisée pour bâtir des tests de l'hypothèse de RCA contre l'hypothèse de surdispersion ou de sous-dispersion en se basant sur la loi asymptotique de $(n-1)\widehat{I}_d$ qui est la loi du Khi-deux à $n-1$ degrés de liberté pour une RCA (Rogers, [80], 1974).

1.5 Processus ponctuels de référence

Nous considérons l'espace mesuré (S, \mathcal{A}, ν) dans ce qui suit.

1.5.1 Processus de Bernoulli et Binomial

Un processus ponctuel N contenant un unique point réparti uniformément sur S est appelé processus de Bernoulli. Il s'écrit : $N = \delta_X$ avec $X \sim \mathcal{U}(S)$. Un processus ponctuel binomial est la superposition de n processus de Bernoulli indépendants. Un processus ponctuel binomial à n points est donc un processus constitué de n points répartis indépendamment selon la loi uniforme sur S . Il s'écrit : $S = \sum_{i=1}^n \delta_{X_i}$ où les X_i sont i.i.d selon $\mathcal{U}(S)$.

$$\forall A \in \mathcal{A}, N(A) \sim \mathcal{B}\left(n, \frac{\nu(A)}{\nu(S)}\right)$$

1.5.2 Processus de Poisson

Les processus ponctuels de Poisson servent de référence dans de nombreuses études (Daley et Vere-Jones [22], 2003).

Définition 1.5.1 Soit $\mu(\cdot)$ une mesure sur (S, \mathcal{A}, ν) telle que $\mu(S) > 0$ et $\mu(A) < +\infty$ pour tout borélien borné A (i.e. μ localement finie).

Un processus ponctuel N sur S est appelé processus ponctuel de Poisson de mesure intensité $\mu(\cdot)$ si :

1. $\forall A \in \mathcal{A}, N(A)$ suit une loi de Poisson de paramètre $\mu(A)$, notée $\mathcal{P}(\mu(A))$.
2. Pour k éléments disjoints deux à deux A_1, \dots, A_k de \mathcal{A} , les variables aléatoires $N(A_1), \dots, N(A_k)$ sont indépendantes.

On parle de processus de Poisson homogène (PPH) lorsqu'il existe $a > 0$ tel que $\mu(\cdot) = a\nu(\cdot)$. Ce paramètre est appelé intensité du PPH relative à ν .

L'hypothèse de processus de Poisson homogène est souvent testée lors d'analyses statistiques concernant la répartition spatiale ou temporelle d'occurrences ayant un intérêt dans l'étude effectuée.

D'un point de vue intuitif, un processus de Poisson homogène sur un espace S quelconque décrit une répartition complètement aléatoire d'occurrences caractérisée par les deux points suivants :

- homogénéité : tous les points de S ont la même chance d'être la position d'occurrence ;
- indépendance : les différentes occurrences se réalisent indépendamment les unes des autres.

1.5.3 Processus de Poisson temporel

Le cas temporel admet une approche différente en raison de l'ordre chronologique.

Définition 1.5.2 Soit $(T_i)_{i \in \mathbb{N}^*}$ un processus ponctuel temporel. Posons $N_t = \#\{T_i, T_i \leq t\}$, alors le processus $(N_t)_{t \geq 0}$ est le processus de dénombrement ou de comptage associé au processus $(T_i)_{i \in \mathbb{N}^*}$.

Définition 1.5.3 Un processus de Poisson de fonction d'intensité $\lambda(t)$ est un processus de dénombrement, $(N_t)_{t \geq 0}$, à accroissements indépendants et tel que $N(t)$ suit une loi de Poisson de paramètre $\Lambda(t)$ (notée $P(\Lambda(t))$), pour tout $t \geq 0$. $\Lambda(t)$ est la fonction d'intensité cumulée définie par : $\Lambda(t) = \int_0^t \lambda(u) du$.

Dans cette définition, on suppose que $\lambda(t) \geq 0$ pour tout $t \geq 0$ et que $\int_0^t \lambda(u) du < \infty$ pour tout $t \geq 0$. Si $\lambda(t) = a > 0$ pour tout $t \geq 0$, on a un PPH.

Nous avons donc par définition : $\Lambda(t) = E[N(t)]$. Nous supposons que $\Lambda(t)$ est continue partout et dérivable sauf en un nombre fini de points sur tout intervalle fini. En tout point où Λ est dérivable, on a $\lambda(t) = \Lambda'(t)$. Si λ est une fonction constante égale à a , on a un PPH avec $\Lambda(t) = at$.

1.5.4 Processus de renouvellement

Définition 1.5.4 *Le processus ponctuel temporel $(T_i)_{i \in \mathbb{N}^*}$ est dit de renouvellement si les temps d'attente $T_{i+1} - T_i, i \in \mathbb{N}^*$ sont des variables aléatoires indépendantes et identiquement distribuées.*

Remarquons que si $T_{i+1} - T_i$ suit une loi exponentielle d'espérance $\frac{1}{a}$, notée $\mathcal{E}(a)$, alors $(T_i)_{i \in \mathbb{N}^*}$ est un PPH d'intensité a .

1.5.5 Processus de Cox

Considérons un processus ponctuel N sur l'espace mesuré (S, \mathcal{A}, ν) dont la mesure d'intensité μ est stochastique. Si N , conditionnellement à μ , est un PPNH de Poisson alors N est appelé processus de Cox ou processus de Poisson doublement stochastique.

Ainsi les processus de Cox temporels sont des processus sans mémoire car d'intensité stochastique ne faisant pas référence au passé.

1.5.6 Processus de Hawkes

En introduisant ce processus auto-excitatif, Hawkes ([40], 1971) souligne la représentation linéaire auquel le terme auto-excitatif se réfère. Il peut être également introduit par le biais de processus de groupe. Considérons un processus de Poisson stationnaire d'intensité λ pour les centres de groupe. Supposons que les groupes forment des réalisations indépendantes d'un processus de branchement fini (figure 1.2); un parent génère une progéniture selon le processus de Poisson d'intensité inférieure à 1. Le processus auto-excitatif est ainsi défini par la génération des groupes.

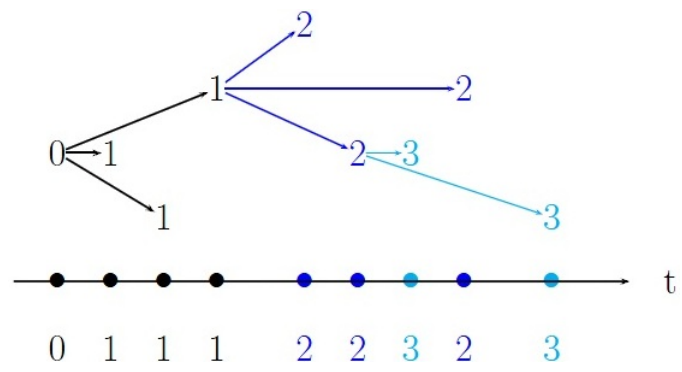


FIGURE 1.2 – Illustration du processus de branchement

ÉCHANTILLONNAGE DE LOIS A

POSTERIORI

Nous nous intéressons aux méthodes permettant d'obtenir la loi *a posteriori* des paramètres du modèle, ceci en fonction des observations disponibles sur le processus ponctuel. Selon le contexte d'étude du processus ponctuel spatio-temporel, nous avons soit les dates d'occurrences, soit les positions des occurrences, soit les deux.

Soient θ le paramètre sur lequel porte l'inférence et Θ l'espace des paramètres. L'approche bayésienne se distingue de l'approche fréquentiste par le fait qu'une mesure de probabilité est associée à l'espace Θ . Cette mesure est appelée loi *a priori* de θ . Elle représente l'information disponible sur ce paramètre. Les méthodes statistiques bayésiennes se sont développées dans les années 90 grâce à l'augmentation de la puissance des calculateurs disponibles (Droesbeke et al. [28], 2002 ; Robert [79], 2006). Le but de l'inférence bayésienne est l'obtention de la loi *a posteriori* de θ conditionnelle aux observations. Robert ([79], 2006) précise que le choix de la loi *a priori* est importante dans le cadre d'une analyse bayésienne et on peut chercher à réduire l'influence de ce choix en utilisant des lois non informatives. L'idée est de minimiser l'information apportée par la loi *a priori* par rapport à celle apportée par les données. Certains auteurs (Droesbeke et al. [28], 2002 ; Robert [78], 1996) parlent de subjectivité de l'approche bayésienne par rapport à ce choix de loi *a priori*, ceci n'est pas forcément un désavantage comme le souligne Robert ([79], 2006). Le traitement bayésien de données utilise les méthodes de Monte Carlo par chaîne de Markov (MCMC) que l'on verra au paragraphe (2.2).

Pour notre part, nous utiliserons les algorithmes de calcul d'inférence les plus efficaces fondés sur les méthodes Monte Carlo par chaînes de Markov. Ces méthodes ont suscités provoqué des développements spectaculaires récents de la statistique bayésienne

(Neal [63], 2000 ; Green et Richardson [38], 2002 ; Albert [1], 2009).

On appelle méthode de Monte-Carlo toute méthode visant à calculer une valeur numérique, en utilisant des procédés aléatoires, c'est-à-dire des techniques de simulation numérique.

2.1 Exemple d'algorithme stochastique de simulation

Méthode d'acceptation-rejet Soit f la densité de la loi à simuler et d'une densité g telle que :

$$f(x) \leq Mg(x) \quad (2.1)$$

où M est une constante et $\text{supp}(f) \subset \text{supp}(g)$,

L'algorithme consiste à imiter f par une densité g plus simple à simuler.

Algorithme 1: Algorithme d'acceptation-rejet

1. Simuler $X \sim g, U \sim \mathcal{U}_{[0,1]}$;
 2. Accepter $Y = X$ si $U \leq \frac{f(X)}{Mg(X)}$;
 3. Sinon retourner en 1.
-

Exemple 2.1.1 Génération d'une loi Normale $\mathcal{N}(0, 1)$ à partir d'une loi de Cauchy :

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \quad \text{et} \quad g(x) \approx \frac{1}{\pi(1+x^2)}$$

$$\frac{f(x)}{g(x)} = \sqrt{\frac{\pi}{2}}(1+x^2) \exp\left(-\frac{x^2}{2}\right) \leq \sqrt{\frac{2\pi}{e}} = 1.52$$

2.2 Méthodes MCMC

2.2.1 Définitions essentielles

Soit x le vecteur des observations, p la densité jointe des observations et des paramètres, p_1 la densité de θ et p_2 la densité de x sachant θ , on a :

$$p(x, \theta) = p_1(\theta)p_2(x|\theta). \quad (2.2)$$

Par conséquent, la loi *a posteriori* p_2 est proportionnelle à la loi jointe p .

Des méthodes de Monte Carlo ont été développées pour approcher au mieux la loi *a posteriori* quand elle ne peut être obtenue de façon analytique (Robert [78], 1996).

Définition 2.2.1 Soient (Ω, \mathcal{A}, P) un espace probabilisé, (E, \mathcal{E}) un espace mesurable et $\{X_t, t \in \mathbb{N}\}$ un processus stochastique à temps discret à valeurs dans (E, \mathcal{E}) , on dit qu'il s'agit d'une chaîne de Markov si

$$\forall t \in \mathbb{N}^*, \forall A \in \mathcal{E}, P(X_t \in A | X_{t-1}, \dots, X_0) = P(X_t \in A | X_{t-1})$$

avec les états de la chaîne (X_t) définis sur (E, \mathcal{E}) .

Définition 2.2.2 On dit qu'une chaîne de Markov est ϕ -irréductible, ϕ étant une mesure de probabilité σ -finie, si et seulement si

$$\forall (x, A) \in E \times \mathcal{E} \text{ avec } \phi(A) > 0, \exists n \in \mathbb{N}^* \text{ tel que } \phi(X_n \in A | X_0 = x) > 0.$$

Remarque 2.2.1 Soit $(X_t, t \in \mathbb{N})$ une chaîne de Markov π -irréductible et apériodique admettant $\pi(\cdot)$ comme loi stationnaire, alors la chaîne $(X_t; t \in \mathbb{N})$ est ergodique.

Remarque 2.2.2 La notion d'ergodicité assure que quelque soit l'état initial x_0 de la chaîne de Markov, la loi objectif π pourra toujours être atteinte.

Définition 2.2.3 L'apériodicité est une condition nécessaire pour garantir la convergence en loi d'une chaîne de Markov. La période d'un état $x \in E$ est définie par :

$$d(x) = \text{pgcd}\{m \geq 1, P^{(m)}(x, x) \geq 0\}.$$

Une chaîne de Markov X_t ϕ -irréductible est apériodique si

$$\forall x \in E, d(x) = 1.$$

Les méthodes MCMC permettent de simuler une chaîne de Markov $(X_t)_{t \geq 0}$ de noyau de transition P convergeant vers la loi d'intérêt π . Les algorithmes MCMC font appel, comme les méthodes d'acceptation-rejet et d'échantillonnage pondéré, à la simulation d'une loi instrumentale. Cette loi instrumentale peut être caractérisée par un noyau de transition $q(\cdot|\cdot)$ ou par une loi conditionnelle.

2.2.2 Efficacité et vitesse de convergence

Lors de l'utilisation des méthodes MCMC, il est primordial de s'assurer que la chaîne converge vers la loi d'intérêt, c'est-à-dire, qu'elle atteint sa phase stationnaire. Afin de statuer sur la convergence des MCMC, plusieurs solutions ont été proposées au fil des années. Les tests les plus courants sont les tests de Gelman-Rubin et de Geweke et de Raftery-Lewis (1992).

Le test de Gelman-Rubin revient à effectuer une analyse de variance entre deux ou plusieurs chaînes ayant des valeurs de départ différentes. L'idée est de chercher une multi-modalité dans l'espace des paramètres. En commençant la chaîne à différents endroits de l'espace d'état, il est possible d'obtenir des maxima locaux, et d'y rester. Toutefois, en prenant en compte plusieurs chaînes, la convergence peut être mise en évidence.

Le test de Geweke permet de statuer sur la stationnarité d'une chaîne. Il se base sur l'écart entre la moyenne des premières p observations et la moyenne des q observations subséquentes d'une seule chaîne telle que le nombre d'itérations $n = p + q$.

Le test de Raftery-Lewis indique le nombre d'itérations requis pour estimer les quantiles d'une fonction des valeurs générées ainsi que la dépendance entre ces valeurs.

2.3 Algorithme de Metropolis-Hastings

Proposé par Metropolis en 1953, puis généralisé par Hastings en 1970, cet algorithme ne requiert qu'une connaissance limitée de la densité à simuler.

Théorème 2.3.1 (*Stationnarité*) *Quelque soit le choix du noyau de transition $q(\cdot|\cdot)$, l'algorithme de Metropolis-Hasting produit une chaîne de Markov $(X_t, t \in \mathbb{N})$ qui admet comme loi stationnaire la loi objectif $\pi(\cdot)$.*

Principe L'algorithme de Metropolis-Hasting réutilise, par un mécanisme d'acceptation-rejet, des variables aléatoires Y produites selon une loi instrumentale.

On utilise souvent trois types d'algorithmes de Metropolis-Hastings :

2.3.1 Algorithme de Metropolis-Hastings indépendant

Il est basé sur l'utilisation d'une densité instrumentale $q(x^{(t)}, y)$ qui peut s'écrire $q(y)$, autrement dit le candidat y généré est indépendant de l'état courant de la chaîne $x^{(t)}$.

Algorithme 2: Algorithme de Metropolis-Hastings

1. choix de la valeur initiale $x^{(0)}$ tel que $\pi(x_0)q(y_0|x^{(0)}) > 0$
2. génération d'un candidat y_t à partir de $q(\cdot|x^{(t)})$
3. évaluation de la probabilité d'acceptation

$$\alpha(x^{(t)}, y_t) = \min\left(1, \frac{\pi(y_t) q(x^{(t)}|y_t)}{\pi(x^{(t)}) q(y_t|x^{(t)})}\right)$$

4. génération de u à partir d'une distribution uniforme sur $[0,1]$
Si $u \leq \alpha(x^{(t)}, y_t)$ alors $x^{(t+1)} = y_t$, sinon $x^{(t+1)} = x^{(t)}$
-

Le nouveau candidat y est accepté avec la probabilité,

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

Par conséquent, q doit être strictement positive sur le support de π et le support de π doit être inclu dans le support de q .

De plus, si l'on souhaite optimiser l'algorithme, il faut choisir une densité q proche de π , pour augmenter le taux d'acceptation des candidats générés.

2.3.2 Algorithme de Metropolis-Hastings à marche aléatoire

Cet algorithme est basé sur l'utilisation d'une densité instrumentale de la forme $q(x^{(t)}, y) = q(y - x^{(t)})$ où q est une densité indépendante de l'état courant $x^{(t)}$. Le candidat y est généré par le processus $y = x^{(t)} + \epsilon$ où ϵ est une perturbation ou une marche aléatoire d'incrément qui suit la loi $q(\cdot)$.

Les densités instrumentales $q(\cdot)$ les plus utilisées sont les densités centrées en 0 et facilement simulables comme les densités uniformes et les densités normales.

Notons que si q est symétrique, ce qui est généralement considéré, on a $q(\epsilon) = q(-\epsilon)$ et, par conséquent, la probabilité d'acceptation est définie par,

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

ce qui correspond à la forme originale de l'algorithme proposé par Metropolis(1953).

2.3.3 Algorithme de Metropolis-Hastings dit *une-variable-à-la-fois*

Lorsque la dimension des candidats y , à générer est très grande, il est très difficile de trouver une densité instrumentale multidimensionnelle q qui entraîne la convergence de la chaîne vers la densité cible.

Par conséquent, au lieu de remettre à jour le vecteur $x^{(t)}$ dans sa totalité, on divise $x^{(t)}$ en plusieurs composantes $(x_1^{(t)}, x_2^{(t)}, \dots, x_d^{(t)})$ de dimension plus petite et de remettre à jour ces composantes l'une après l'autre.

À chaque itération de l'algorithme, on fait successivement évoluer chacune des d composantes $x_i^{(t)}$ en utilisant, par conséquent, d étapes de type Metropolis-Hastings et donc d densités instrumentales $q_i(y_i|x_i^{(t)}, x_{-i}^{(t)})$ pour obtenir le nouveau vecteur $x^{(t+1)}$.

On peut remarquer que chacune des d étapes de l'algorithme de Metropolis-Hastings simule une chaîne de Markov dont la densité invariante est la densité conditionnelle $\pi(x_i|x_{-i}^{(t+1)})$ où

$$x_{-i}^{(t+1)} = (x_1^{(t+1)}, \dots, x_{i-1}^{(t+1)}, x_{i+1}^{(t)}, \dots, x_d^{(t)}), i = 1, \dots, d.$$

La probabilité d'acceptation du $i^{\text{ème}}$ élément est :

$$\alpha_i(x_i^{(t)}, x_{-i}^{(t+1)}, y_i) = \min \left(1, \frac{\pi(y_i|x_{-i}^{(t+1)})q_i(x_i^{(t)}|y_i, x_{-i}^{(t+1)})}{\pi(x_i^{(t)}|x_{-i}^{(t+1)})q_i(y_i|x_i^{(t)}, x_{-i}^{(t+1)})} \right).$$

Exemple : Simulation d'une loi de Poisson à partir de la loi Normale $\mathcal{N}(0, 1)$

La loi instrumentale est :

$$q(x, y) = \exp(-(y - x)^2/2)/\sqrt{2\pi},$$

on obtient alors une marche aléatoire

$$X_{n+1} = X_n + R_{n+1}$$

avec $R_{n+1} \sim \mathcal{N}(0, 1)$.

Initialisation de l'algorithme :

Premier élément de la chaîne : $u_0 = 1$

Nombre de répétitions : N

1) Avantages :

- a) Permet de simuler des distributions dont la densité est connue à un facteur multiplicatif près.
 - b) La chaîne générée converge en distribution vers la loi souhaitée.
- 2) Inconvénients :
- a) l'échantillon généré n'est pas indépendant.
 - b) il faut déterminer la période de rodage de la chaîne.
 - c) les temps de calcul peuvent être selon les cas plus ou moins long.

2.4 Échantillonneur de Gibbs

De part son universalité, la technique de Metropolis-Hastings présentée dans la section précédente est très attrayante. Toutefois, le manque de liaison entre la loi instrumentale q et la loi cible π peut nuire aux propriétés de convergence de la méthode. L'utilité de l'échantillonnage de Gibbs vient du fait qu'il soit fondé sur loi π . D'abord introduit par Geman et Geman en 1984 en restauration d'image, puis généralisé par Tanner et Wong en 1987, ensuite Gelfand en 1990 en statistique appliquée, cet algorithme est utilisé pour simuler une loi jointe $\pi(x)$, $x \in \mathbb{R}^N$ et suppose que les lois conditionnelles $\pi(x_i|x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_N)$ sont simulables aisément.

Principe D'un point de vue général, l'échantillonnage de Gibbs tire profit des structures hiérarchiques d'un modèle. Il consiste à réactualiser "composante par composante" le vecteur x où les densités génératrices de candidat sont les densités conditionnelles des composantes x_i conditionnellement à toutes les autres composantes du vecteur x

Algorithme 3: Algorithme de Gibbs

A l'étape t , disposant d'une réalisation $x^{(t)} = (x_1^{(t)}, \dots, x_N^{(t)})$, on génère $x^{(t+1)}$ de la manière suivante :

$$\begin{aligned}
 x_1^{(t+1)} &\sim \pi(x_1|x_2^{(t)}, \dots, x_N^{(t)}) \\
 x_2^{(t+1)} &\sim \pi(x_2|x_1^{(t+1)}, \dots, x_3^{(t)}, \dots, x_N^{(t)}) \\
 x_3^{(t+1)} &\sim \pi(x_3|x_1^{(t+1)}, x_2^{(t+1)}, \dots, x_4^{(t)}, \dots, x_N^{(t)}) \\
 &\vdots \\
 x_N^{(t+1)} &\sim \pi(x_N|x_1^{(t+1)}, \dots, x_{N-1}^{(t+1)})
 \end{aligned}$$

On montre que la chaîne produite converge en distribution vers la loi jointe $\pi(x)$. Par ailleurs, l'échantillonneur de Gibbs suppose une connaissance plus poussée de la

loi à simuler au sens où on doit être capable de simuler suivant les lois conditionnelles. Notons que l'échantillonneur de Gibbs est particulièrement bien adapté aux modèles hiérarchiques, que les paramètres inconnus sont munis de lois *a priori* ainsi que les hyperparamètres associés et que généralement, on introduit des lois non informatives au dernier niveau de la hiérarchie.

2.5 Méthodes hybrides

Principe La convergence de l'échantillonneur de Gibbs peut être lente car une seule composante à chaque itération est simulée. À cet effet, des méthodes hybrides, combinant les méthodes MCMC et l'échantillonnage de Gibbs, ont été développées. De plus, ces méthodes sont moins sensibles à la loi instrumentale comme l'est l'algorithme de Metropolis-Hastings.

Algorithme Un algorithme MCMC hybride est une méthode MCMC utilisant simultanément des étapes d'échantillonneur de Gibbs et des étapes de Metropolis-Hastings. Considérons la densité g telle que π soit une loi marginale de g . Remarquons que les lois conditionnelles de g sont parfois plus aisées à simuler que celles de π .

2.6 MCMC à sauts réversibles

Principe La méthode de Monte Carlo par Chaînes de Markov à sauts réversibles (RJCMC), introduite par Green ([37], 1995), puis développée par certains auteurs tels Green et Richardson ([38], 2002) et Ehlers et Brooks ([31], 2008), fournit une approche générale pour la simulation de MCMC, pour lesquelles la dimension de l'espace des paramètres peut varier au cours des itérations de la chaîne. L'échantillonneur à sauts réversibles peut être vu comme une extension de l'algorithme de Metropolis-Hastings sur des espaces d'états plus généraux.

Afin d'illustrer nos propos, supposons que pour des données observées x , nous avons un nombre fini de modèles candidats $(\mathcal{M}_k)_{k \in \mathbb{N}^*}$. L'indice k peut être considéré comme une variable indicatrice du modèle pris en compte, telle que $\mathcal{M}_{k'}$ désigne le modèle pour lequel $k = k'$. Chaque modèle \mathcal{M}_k est constitué de paramètres inconnus formant un vecteur, $\theta_k \in \Theta^{n_k}$, de dimension n_k (n_k pouvant prendre des valeurs différentes en fonction du modèle). La loi jointe *a posteriori* de (k, θ_k) sachant les observations x , est

obtenue à partir de la vraisemblance $L(x|k, \theta_k)$ et de la loi *a priori* $p(k, \theta_k) \propto p(\theta_k|k)p(k)$ et s'écrit comme suit :

$$\pi(k, \theta_k|x) = \frac{L(x|k, \theta_k)p(k, \theta_k)}{\sum_{k' \in \mathbb{N}^*} \int_{\theta_k \in \Theta^{n_{k'}}} L(x|k', \theta_{k'})p(k', \theta_{k'})d\theta_{k'}}. \quad (2.3)$$

Algorithme L'algorithme de la méthode RJMCMC utilise la loi jointe *a posteriori* définie par l'expression 2.3 comme loi cible de l'échantillonneur MCMC sur l'espace d'états $\bigcup_{k \in \mathbb{N}^*} (\{k\} \times \Theta^{n_k})$, où les états de la chaîne de Markov sont de la forme (k, θ_k) .

2.7 Applications aux processus ponctuels

Lorsque l'on observe un processus ponctuel, plusieurs situations se présentent (paragraphe 3.1) et les observations envisageables dépendent des contraintes techniques et du coût d'échantillonnage. Dans l'expression (2.2), la vraisemblance $p_2(x|\theta)$ peut prendre plusieurs formes. Par exemple, pour un processus ponctuel spatio-temporel simple dont on dispose sur un domaine X au cours d'une période $[0, T]$, de toutes les positions et dates d'occurrences, la vraisemblance sera donnée par (1.2). Un autre exemple est le cas d'observations de cartes spatiales à des dates données d'observations (paragraphe 3.2.1), la vraisemblance aura la forme indiquée par l'expression (3.7).

ASSOCIATION SPATIO-TEMPORELLE DES PROCESSUS DE COX, SHOT-NOISE ET DIRICHLET

Dans ce chapitre, nous considérons des processus de Cox dont l'intensité est associée à un processus shot noise généralisé (Brix [11], 1999 ; Brix et Chadœuf [12], 2002 ; Møller [58], 2003 ; Møller et Torrisi [61], 2005). Pour ce faire, nous nous intéressons à l'étude de phénomènes conduisant à des occurrences d'événements localisés dans l'espace et/ou dans le temps.

Les processus de Dirichlet définissent une distribution sur l'ensemble des distributions de probabilité. Ils permettent donc de définir, dans le cadre de l'estimation bayésienne, un *a priori* sur une distribution de probabilité inconnue (Ferguson [32, 33], 1973, 1974 ; Blackwell et Macquenn [8], 1973 ; Sethuraman [83], 1994 ; Teh et al. [89], 2006).

Considérons l'espace mesuré (X, \mathcal{A}_d, ν) où \mathcal{A}_d désigne la tribu des boréliens pour la distance d . Soit $A \in \mathcal{A}_d$ et Y : le processus correspondant aux occurrences observées (Y suit la loi du processus observé). On pose :

$$\lambda(x, t) = \sum_{i: a_i, x_i, t_i} a_i \phi(d((x, t), (x_i, t_i))) \quad (3.1)$$

avec

- ϕ : noyau spatio-temporel
- (x_i, a_i, t_i) : réalisation d'un processus de Poisson \mathcal{N} à valeurs dans $(X \times \mathbb{R}_+^* \times \mathbb{R}_+)$

- \mathcal{N} : processus de Poisson stationnaire d'intensité μ ;
dans notre cas, c'est un processus caché
- x_i : position spatiale du $i^{\text{ème}}$ événement caché (c'est-à-dire non observé)
- t_i : date du $i^{\text{ème}}$ événement caché
- a_i : contribution à l'intensité du $i^{\text{ème}}$ événement caché ;
 $a_i \sim \Gamma(b, 1)$ (loi conjuguée de la loi de Poisson)
- λ : intensité spatio-temporelle du processus observé
- $\lambda(\cdot, t) = \int_X \lambda(x, t) \nu(dx)$: intensité temporelle du processus observé.

Exemple 3.0.1 *On observe des ravages sur une culture, l'événement caché serait l'insecte ou le champignon responsable des dégâts.*

De façon heuristique, $\lambda(x, t) \nu(dx) dt$ est la probabilité d'avoir un événement autour de x à la date t .

Quelques propriétés distributionnelles de $Y(\mathbf{B} \times [0, \mathbf{t}])$

$Y(B \times [0, t])$ désigne le nombre d'occurrences du processus observé dans $B \times [0, t]$. On se sert du fait que Y soit un processus de Cox donc conditionnellement à λ , un processus de Poisson, $Y(B \times [0, t]) \sim \mathcal{P} \left(\int_B \int_0^t \lambda(x, u) \nu(dx) du \right)$, afin d'obtenir :

$$E(Y(B \times [0, t]) | \lambda) = \int_B \int_0^t \lambda(x, u) \nu(dx) du$$

$$V(Y(B \times [0, t]) | \lambda) = \int_B \int_0^t \lambda(x, u) \nu(dx) du.$$

On peut ainsi écrire l'espérance et la variance totale :

$$E(Y(B \times [0, t])) = E \left(\int_B \int_0^t \lambda(x, u) \nu(dx) du \right) \quad (3.2)$$

$$V(Y(B \times [0, t])) = E(V(Y(B \times [0, t]) | \lambda)) + V(E(Y(B \times [0, t]) | \lambda)) \quad (3.3)$$

$$= E \left(\int_B \int_0^t \lambda(x, u) \nu(dx) du \right) + V \left(\int_B \int_0^t \lambda(x, u) \nu(dx) du \right). \quad (3.4)$$

Lorsqu'on étudie un phénomène qui conduit à des occurrences d'événements localisés dans l'espace et dans le temps, on peut avoir les observations sous diverses formes (Vaillant [91, 92], 1991 et 1992). A cet effet, quelques exemples sont données dans la

section suivante.

3.1 Exemples de données observables

- i. Une carte complète en temps continu dans $X \times [0; T]$ c'est-à-dire pour tout événement donné, on peut avoir sa localisation et sa date d'occurrence.

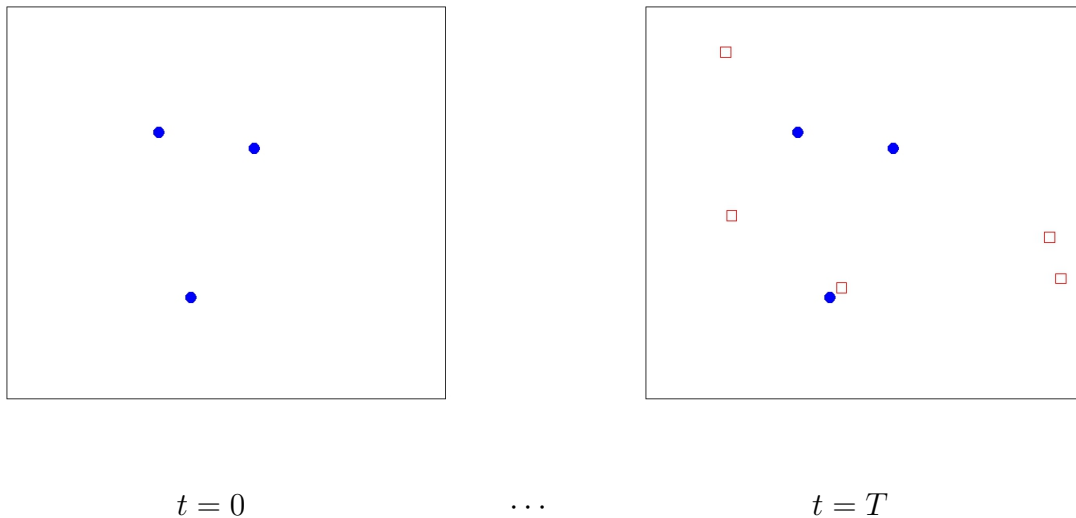


FIGURE 3.1 – Cartes exhaustives en temps continu : \bullet représente les occurrences à la date $t = 0$ et \square représente les occurrences dans $]0, T]$

- ii. Nous disposons plus fréquemment des données du type k cartes exhaustives à des dates d'observations échelonnées dans le temps $t_1 < t_2 < \dots < t_k$. En effet, dans la pratique, nous sommes rarement en temps continu, la vraisemblance du processus ponctuel fournit par son processus intensité n'est pas disponible. Par contre, pour chaque carte spatiale, nous avons cette vraisemblance à partir des intensités cumulées par la période délimitée par deux dates d'observations.

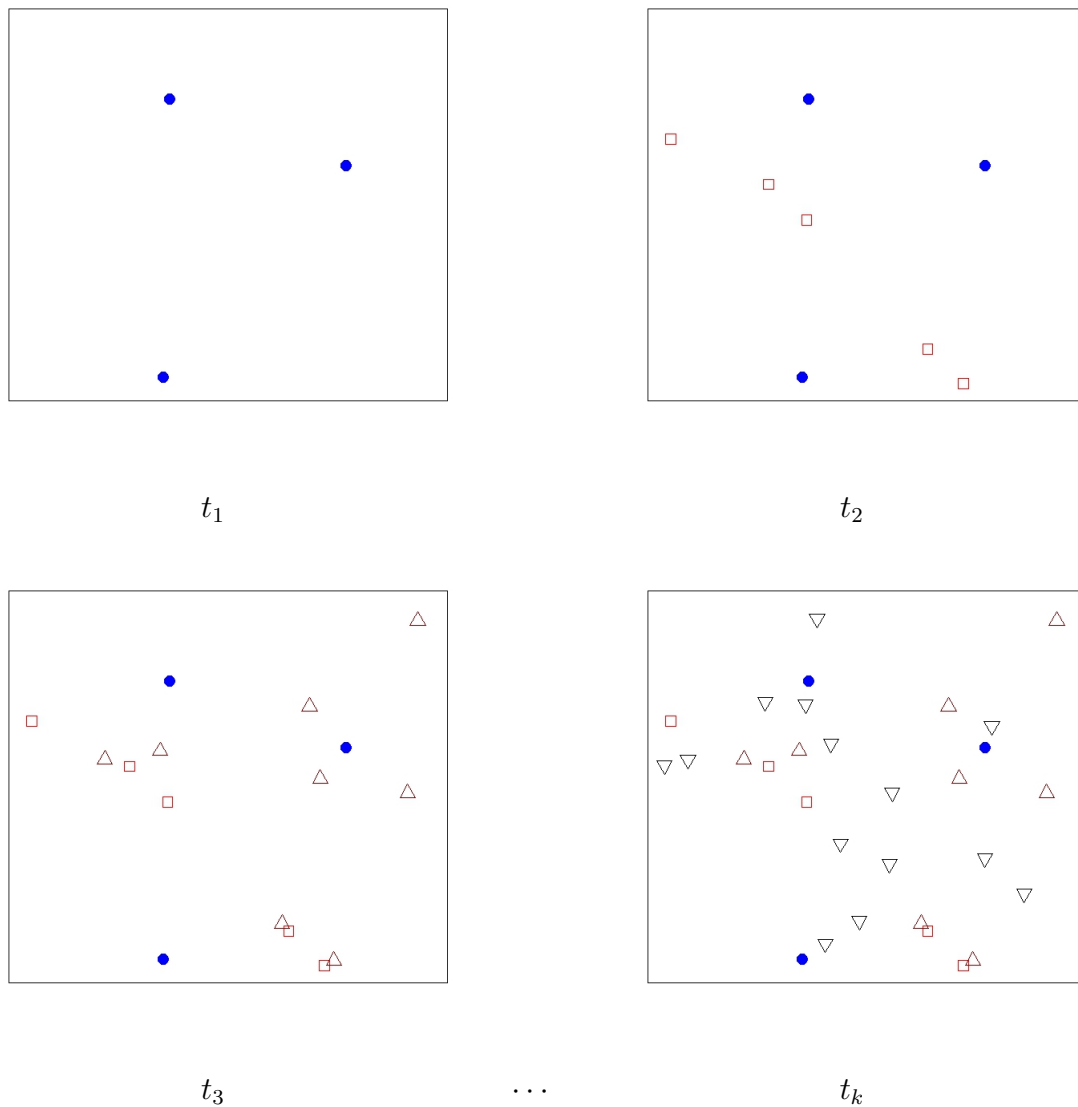


FIGURE 3.2 – Cartes exhaustives en temps continu : \bullet représente les occurrences à la date t_1 , \square : les occurrences à la date t_2 , \triangle : les occurrences à la date t_3 et ∇ : les occurrences à la date t_k

- iii. Un troisième type de données moins informatif que le précédent est lorsque l'on dispose de comptages d'événements dans des sous-espaces X_i de X formant une partition de X .

n_1	n_2	n_3	\dots
\dots	\dots	\dots	\dots
\dots	\dots	\dots	\dots
\dots	\dots	\dots	n_k

FIGURE 3.3 – Comptages dans le domaine d'étude à diverses dates

Remarque 3.1.1 *Notons qu'il est possible d'obtenir des données de comptages d'événements dans des sous-espaces X_i de X formant une partition de X avec des dates d'observations $t_1 < t_2 < \dots < t_k$ à partir des données illustrées en dans le deuxième exemple.*

3.2 Statistique d'un processus de Cox spatio-temporel associé à un processus de Dirichlet

Nous considérons la situation où des occurrences d'un phénomène se produisent dans une zone d'étude et constituent la réalisation d'un processus de Cox spatio-temporel dont l'intensité est associée à un processus shot noise généralisé. Le modèle correspond à une mesure d'intensité liée à des contributions générées par un processus caché de Poisson et qui suivent un processus de Dirichlet centré sur la loi Gamma. À partir des positions spatiales des occurrences observées entre plusieurs paires de dates d'observations consécutives, nous proposons d'inférer sur les paramètres d'intérêt à l'aide de méthodes MCMC dans le cadre d'un modèle bayésien hiérarchique.

Les processus ponctuels sont souvent utilisés comme modèles de répartitions spatiales ou spatio-temporelles d'occurrences car ce sont des mesures stochastiques à valeurs entières qui sont associées à des points d'un espace donné (Brémaud [9], 1981; Daley et Veres-Jones [22], 2003); Karr [49], 1991).

Le processus de Poisson, homogène ou non, est souvent proposé en cas d'absence présumée de corrélation spatiale ou temporelle entre occurrences. Cette classe de processus bien que riche (Kutoyants [53], 1998) est limitative. Des modèles alternatifs ont

donc été proposés en tenant compte des caractéristiques du phénomène étudié, par exemple les processus de Cox (Brix et Diggle [13], 2001 ; Kottas et Sansó [52], 2007 ; Ji et al. [45], 2009), les processus à mémoire tels que le processus d'Hawkes (Hawkes [40], 1971 ; Brémaud et Massoulié [10], 2001). Une panoplie de modèles spatiaux et de méthodes d'analyse statistique présentée par Møller et Waagepetersen ([62], 2007) permet d'évaluer la diversité des choix disponibles pour effectuer une modélisation à partir de processus ponctuels.

Lorsque l'on désire étudier les localisations spatiales et temporelles d'occurrences liées à un phénomène, les données disponibles ou observables pour une telle analyse peuvent être sous diverses formes (Cressie [20], 1993 ; Diggle [25], 2005 ; Kottas et al. [51], 2007 ; Ogata [68], 2011 ; Ripley [77], 1981 ; Vaillant et al. [96], 2011 ; Waller [109], 2005). Nous nous intéressons pour notre part à la situation où l'on dispose de cartes complètes d'une zone d'observation X à des dates successives non régulièrement espacées dans le temps. Les positions des occurrences sont donc observées mais les dates d'occurrences ne le sont qu'à un intervalle de temps près. Nous considérons d'autre part que le phénomène de répartition d'occurrences étudié obéit à un processus de Cox shot noise généralisé (Møller et Torrisi [61], 2005). Ce modèle postule que les occurrences sont conditionnellement à un environnement aléatoire, réparties indépendamment les unes des autres. Il stipule, en outre, que l'intensité spatio-temporelle est une somme lissée de contributions aléatoires associées à des événements engendrés par un phénomène sous-jacent à celui observé. Cette classe de processus ponctuels simples est donc adéquate dans bien des situations en sismologie (Vere-Jones [107], 1995 ; Ogata [66], 1998), écologie (Matern [56], 1971 ; Vaillant [93, 94], 1992 et 1994), neurophysiologie (Lansky et Vaillant [54], 2000 ; Truccolo et al. [90], 2005), hydrologie (Cowpertwaita et al. [18], 1996), analyse d'image et épidémiologie (Beneš et al. [5], 2005 ; Rosychuk et al. [81], 2006 ; Kottas et Sansó [52], 2007) ou recherche forestière (Brix et Chadoëuf [12], 2002 ; Møller et Diaz-Avalos [59], 2010 ; Stoyan et Penttinen [85], 2000) pour lesquelles on observe des occurrences sans coïncidence mais en présence d'agrégativité et de corrélations spatiales et temporelles positives nettes. Elle inclut les processus de Poisson en groupes tels que celui de Neyman-Scott et bien d'autres (Møller et Torrisi [61], 2005).

Nous proposons dans cette partie une approche bayésienne semi paramétrique pour des processus de Cox shot noise généralisés associés à un processus de Dirichlet. Ce dernier processus modélise des contributions environnementales cachées affectant l'intensité du processus ponctuel observé. Au paragraphe 2, nous introduisons les notations utilisées et le contexte de modélisation par processus ponctuel est précisé. Nous sou-

lignons que les données auxquelles nous nous intéressons sont des positions spatiales d'occurrences avec une information restreinte sur les dates de ces occurrences. Le processus de Dirichlet (DP) introduit par Ferguson ([32, 33], 1973 et 1974) est présenté au paragraphe 3 ainsi que son utilité comme loi *a priori* sur les contributions à l'intensité du processus ponctuel étudié. La nature discrète du processus de Dirichlet permet en effet de décrire les cas où ces contributions sont identiques bien qu'associées à des positions spatiales et dates différentes. Les paragraphes 4 et 5 sont consacrés respectivement à la modélisation bayésienne hiérarchique par processus de Cox et aux propriétés distributionnelles. Des méthodes MCMC sont discutées au paragraphe 6 pour l'obtention des lois *a posteriori* fournissant les outils d'inférence statistique. Les résultats des expérimentations numériques sont fournis au paragraphe 7, tandis que des données de référence concernant les ronces *Rubus fruticosus* sont traitées au paragraphe 8 pour illustrer notre approche.

3.2.1 Notations et contexte

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilité, $(X \times \mathbb{R}_+, \mathcal{B}_X \otimes \mathcal{B}_{\mathbb{R}_+}, \nu \otimes l)$ un espace mesuré où l est la mesure de Lebesgue. Nous considérons des données du type carte complète d'occurrences à des dates d'observations $t_1 < t_2 < \dots < t_I$. Pour tout élément i de $\{1, \dots, I\}$, nous notons C_i la carte observée au temps t_i avec $C_i = \{x_1, \dots, x_{N_i}\}$ ensemble des N_i positions spatiales des occurrences produites dans X jusqu'au temps t_i . Ces cartes peuvent être considérées comme les réalisations d'un processus ponctuel spatio-temporel N observé successivement aux dates t_1 à t_I . Les dates ne sont pas supposées régulièrement espacées donc une approche en temps discret n'est pas adaptée. C_I est la dernière carte observée et contient donc les observations des I cartes précédentes.

Les données sur les positions spatiales sont complètes cependant l'information sur les dates d'occurrences $\{\tau_k, k = 1, \dots, N_I\}$ sont de la forme $\tau_k \in]t_{i-1}, t_i]$ donc nous dirons qu'elles sont incomplètes.

Notons Λ la mesure intensité du processus ponctuel N considéré. Il est important de souligner que si N est un processus de Poisson, il est simple si et seulement si Λ est diffuse c'est-à-dire ne comporte aucun atome. Si Λ est une mesure stochastique et N un processus de Poisson conditionnellement à Λ , alors N est un processus de Cox dirigé par Λ et la loi de cette mesure caractérise celle de N (Karr [49], 1991). Le choix d'un modèle de mesure stochastique dirigeant un processus de Cox dépend donc des propriétés requises pour décrire de façon adéquate le phénomène étudié. Ainsi, N est

simple si et seulement si Λ est diffuse ; si Λ est stationnaire, alors N est stationnaire. D'autre part, si Λ appartient à la classe \mathcal{M}_{ca} des mesures complètement aléatoires, alors pour tout couple (A_1, A_2) d'éléments disjoints de $\mathcal{B}_X \otimes \mathcal{B}_{\mathbb{R}_+}$, les variables $N(A_1)$ et $N(A_2)$ sont indépendantes car il en est de même pour $\Lambda(A_1)$ et $\Lambda(A_2)$. Par contre, une mesure complètement aléatoire est discrète (Jordan [46], 2010) et permet de tenir compte de situations où des occurrences peuvent coïncider. Dans tous les cas, l'aléa sur Λ produit de la surdispersion (Brix [11], 1999 ; Lansky et Vaillant [54], 2000). Mais, il est possible d'obtenir un processus de Cox simple exhibant de la surdispersion et de la corrélation spatiale, par exemple en associant Λ à une famille localement finie d'éléments de \mathcal{B}_X par le biais de mesures restrictions (Vaillant [91, 92], 1991, 1992). Ainsi, pour modéliser la dépendance spatiale entre données de comptage d'un processus de Cox, Wolpert et Ickstadt ([110], 1998) ont construit une mesure intensité qui est un mélange par noyau d'une mesure stochastique infiniment divisible à incréments indépendants. Ils posent alors un modèle bayésien hiérarchique en proposant une loi *a priori* également sur le noyau.

Une façon de modéliser un processus spatio-temporel de Cox de mesure intensité Λ est, sous certaines conditions de régularité, de construire son processus intensité défini comme suit : si ν est diffuse sur X et s'il existe un processus mesurable positif λ indicé dans $X \times \mathbb{R}_+$ tel que

$$\forall (B, t) \in \mathcal{B}_X \times \mathbb{R}_+, \quad \Lambda(B \times [0, t]) = \int_B \int_0^t \lambda(x, t) \nu(dx) dt \quad (3.5)$$

alors λ est appelé intensité dirigeante de N .

En des termes simples, $\lambda(x, t) \nu(dx) dt$ peut être interprété comme la probabilité qu'une occurrence se produise dans une zone infinitésimale autour de x de surface $\nu(dx)$ et dans un laps de temps infinitésimal de durée dt . Sous certaines conditions de régularité (Daley et Vere-Jones [22], 2003 ; Karr [49], 1991), on peut donc écrire la vraisemblance du processus N conditionnelle à λ .

Dans ce qui suit, nous nous intéressons à un processus de Cox N d'intensité dirigeante λ_θ où θ est le vecteur des paramètres du modèle. En posant $t_0 = 0$, l'expression de la vraisemblance conditionnelle à λ_θ pour la carte $C_i \setminus C_{i-1}$ est :

$$L_{C_i \setminus C_{i-1}}(x_{N_{i-1}+1}, \dots, x_{N_i}; \theta) = e^{-\Lambda_\theta(X \times]t_{i-1}, t_i])} \prod_{l=N_{i-1}+1}^{N_i} \int_{t_{i-1}}^{t_i} \lambda_\theta(x_l, t) dt. \quad (3.6)$$

Notons que $C_i \setminus C_{i-1}$ correspond à la carte des occurrences qui se sont produites dans $]t_{i-1}, t_i]$.

À partir de l'expression (3.6), on obtient donc la vraisemblance globale conditionnelle au processus intensité λ_θ :

$$L(x_1, \dots, x_{N_I}; \theta) = e^{-\Lambda_\theta(X \times]0, T_I])} \prod_{i=1}^I \prod_{l=N_{i-1}+1}^{N_i} \int_{t_{i-1}}^{t_i} \lambda_\theta(x_l, t) dt. \quad (3.7)$$

Il s'agit de la vraisemblance d'un processus de Poisson non homogène car λ_θ est alors considéré non aléatoire.

Considérons maintenant la loi du processus λ_θ . Diverses approches sont proposées dans la littérature pour prendre en compte l'hétérogénéité spatiale et temporelle. On peut se référer à l'article de Green et Richardson [38] (2002) où un éventail de méthodes est discuté. Une approche bayésienne semi-paramétrique y est présentée avec des modèles de champs markoviens cachés et des applications en analyse d'image et en épidémiologie. Pour ce qui concerne les processus spatiaux de Cox, Møller et Waagepetersen ([62], 2007) ont comparé quelques modèles de mesures dirigeantes : processus log-gaussiens et processus shot noise. Ils soulignent qu'une généralisation du modèle de processus intensité donné par le processus shot noise est possible par le biais du choix de la fonction de noyau associée, et du processus ponctuel associé. Ces extensions peuvent être assimilées à des modèles de mélange pour le processus intensité et sont intéressantes quand il existe des zones de l'espace étudié pour lesquelles différents modèles d'intensité peuvent s'adapter et que ces zones ne sont pas forcément identifiables. Wolpert et Ickstadt ([110], 1998) ont effectué des travaux analogues sur les mesures intensité de processus de Cox autocorrélées positivement.

Nous étudions un processus de Cox spatio-temporel dirigé par un processus shot noise généralisé (Møller et Torrisi [61], 2005). Nous posons

$$\forall (x, t) \in X \times \mathbb{R}_+, \quad \lambda_\theta(x, t) = \int_{X \times \mathbb{R}_+} a K_{(\sigma_1, \sigma_2)}((x, t), (y, s)) d(L(y, s, a)) \quad (3.8)$$

où L est un processus ponctuel sur $(X \times \mathbb{R}_+^2, \mathcal{B}_X \otimes \mathcal{B}_{\mathbb{R}_+^2})$; et $K_{(\sigma_1, \sigma_2)}$ est un noyau tel que $K_{(\sigma_1, \sigma_2)}(\cdot, (y, s))$ soit une densité de probabilité sur $X \times \mathbb{R}_+$ pour tout (y, s) de $X \times \mathbb{R}_+$. Le modèle (3.8) peut être reformulé de la façon suivante :

$$\forall (x, t) \in X \times \mathbb{R}_+, \quad \lambda_\theta(x, t) = \sum_{j=1}^M a_j K_{(\sigma_1, \sigma_2)}((x, t), (y_j, s_j)) \quad (3.9)$$

où M est le nombre aléatoire de contributions. Les (y_j, s_j) sont la réalisation d'un processus ponctuel L_1 sur $X \times \mathbb{R}_+$, tandis que les a_j sont des marques identiquement distribuées selon une loi de probabilité G sur \mathbb{R}_+ , et indépendantes des (y_j, s_j) .

Si L_1 est associée à un processus de Poisson homogène ($PPH(\mu)$) d'intensité scalaire μ , alors λ_θ est un processus shot noise standard (Møller [58], 2003) et N est donc, conditionnellement à L_1 , la superposition de processus de Poisson indépendants d'intensité $a_j K_{(\sigma_1, \sigma_2)}(\cdot, (y_j, s_j))$. Il peut décrire une répartition d'occurrences principales dont chacune génère dans son voisinage des occurrences secondaires. Dans la pratique, il peut s'agir de répliques de tremblements de terre en sismologie, des rejetons de plantes en écologie, de contacts infectueux en épidémiologie. Un exemple de tels phénomènes agrégatifs modélisé par un processus de Poisson en groupes est discuté dans Brix et Chadœuf ([12], 2002).

Notons que dans le processus de Cox shot noise généralisé (Møller et Torrisi [61], 2005), la fonction noyau en chaque point du processus L_1 se voit attribuée une bande passante aléatoire. D'autre part, si dans la plupart des applications L_1 est un processus de Poisson, il s'avère que d'autres modèles peuvent être plus appropriés comme, par exemple, des modèles markoviens répulsifs dans le cas de comportement territorial chez certaines espèces vivantes, dû à la compétition intraspécifique.

3.2.2 Loi et processus de Dirichlet

Ferguson ([32, 33], 1973 et 1974) a introduit le processus de Dirichlet (DP) comme une mesure de probabilité sur l'espace des mesures de probabilité. Plus précisément, soit G une mesure stochastique de probabilité sur un espace mesurable (χ, \mathcal{H}) . G est un processus de Dirichlet sur (χ, \mathcal{H}) si pour toute partition mesurable de χ d'éléments de \mathcal{H} , notés B_1, \dots, B_k , le vecteur des probabilités $(G(B_1), \dots, G(B_k))$ suit la distribution de Dirichlet de paramètres $\alpha G_0(B_1), \dots, \alpha G_0(B_k)$ où α réel strictement positif et G_0 mesure de probabilité sur (χ, \mathcal{H}) sont les deux paramètres de G . On note $G \sim DP(\alpha, G_0)$. G_0 est appelée mesure de base ou espérance de G car pour tout élément B de \mathcal{H} on a $E(G(B)) = G_0(B)$. Le scalaire α est le paramètre de précision (ou degré de confiance) et caractérise la variabilité de G autour de G_0 de la façon suivante : soient A et B

appartenant à \mathcal{H} tels que $A \cap B = \emptyset$, alors $\text{cov}(G(A), G(B)) = -\frac{G_0(A)G_0(B)}{\alpha + 1}$ et $\text{Var}(G(A)) = \frac{G_0(A)(1 - G_0(A))}{\alpha + 1}$.

Une propriété importante est qu'une réalisation d'un processus de Dirichlet est discrète presque sûrement. Si $G \sim DP(\alpha, G_0)$, alors la représentation *stick-breaking* de G introduite par Sethuraman ([83], 1994) est simplement :

$$G = \sum_{k=1}^{\infty} p_k \delta_{\zeta_k} \quad (3.10)$$

où

$$\zeta_k \stackrel{i.i.d}{\sim} G_0, \quad p_k = V_k \prod_{j=1}^{k-1} (1 - V_j) \quad \text{et} \quad V_k \stackrel{i.i.d}{\sim} \text{Beta}(1, \alpha)$$

δ_{ζ_k} étant la mesure de Dirac en ζ_k .

Il est important de noter que la suite $p = (p_k)_{k \in \mathbb{N}^*}$ est telle que les p_k sont des variables aléatoires positives indépendantes des ζ_k et telles que $\sum_{k=1}^{\infty} p_k = 1$. On dit que p suit un *stick-breaking* de paramètre α et on note $p \sim SB(\alpha)$. Les réalisations d'un processus de Dirichlet peuvent donc être vues comme des mélanges infinis dénombrables de lois (Chimard [14], 2010).

3.3 Apport bayésien

3.3.1 Modélisation bayésienne hiérarchique semi paramétrique

Notons $C_i^* = C_i \setminus C_{i-1}$ le vecteur des positions d'occurrence entre les dates t_{i-1} et t_i . Conditionnellement à λ_θ , C_i^* suit un processus de Poisson non-homogène (*PPNH*). Pour ce qui concerne les contributions aléatoires a_j , dans le modèle standard de processus de Cox shot noise, elles sont indépendantes identiquement distribuées (Møller [58], 2003) ainsi que dans le modèle généralisé. La loi commune souvent utilisée est la loi Gamma. Nous proposons de conserver cette loi marginale pour les contributions a_j mais d'élargir la classe de loi *a priori* sur ces effets environnementaux. Un processus de Dirichlet de loi de base Gamma est donc retenu. On pourra consulter Teh et al. ([89], 2006) pour les éléments concernant la modélisation hiérarchique avec cette classe de processus. Enfin, le noyau spatio-temporel $K_{(\sigma_1, \sigma_2)}$ est choisi gaussien isotropique pour sa partie spatiale de bande passante σ_1 , tandis que la partie temporelle est exponentielle

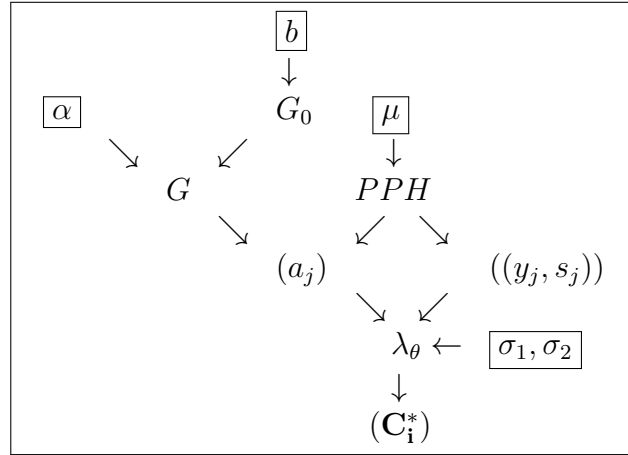


FIGURE 3.4 – Graphe orienté acyclique du modèle hiérarchique

de paramètre σ_2 (voir l'expression (3.15)). La représentation stick-breaking introduite par Sethuraman ([83], 1994) nous donne le modèle hiérarchique suivant :

$$\begin{aligned}
 C_i^* | \lambda_\theta &\stackrel{i.i.d.}{\sim} PPNH((\lambda_\theta(\cdot, t)_{t_{i-1} < t \leq t_i})), \quad i = 1, \dots, I \\
 \lambda_\theta(x, t) &= \sum_{(y_j, s_j, a_j)} a_j K_{(\sigma_1, \sigma_2)}((x, t), (y_j, s_j)), \quad \forall (x, t) \in X \times \mathbb{R}_+ \\
 a_j | G &\stackrel{i.i.d.}{\sim} G \\
 (y_j, s_j) &\sim PPH(\mu) \\
 G &\sim DP(\alpha, G_0) \\
 G_0 &= \text{Gamma}(1, b)
 \end{aligned} \tag{3.11}$$

Les a_j sont donc corrélés par le biais de G (théorème de de Finetti) et ont une loi jointe échangeable. Notons qu'il est possible de considérer, non plus un processus de Dirichlet pour la loi des a_j , mais un processus stick breaking à noyaux (KSBP) introduit par Dunson et Park ([30], 2008) et An et al. ([2], 2008) quand il est important de tenir compte de certaines formes de corrélation spatiale locale. Dans le même contexte, Rao et Teh ([75], 2009) ont proposé un modèle de dépendance entre DP par le biais d'un processus gamma spatial normalisé. Cependant, ceci rend beaucoup plus complexe les calculs permettant d'aboutir à la loi *a posteriori* des paramètres du modèle. Dans cette section, c'est la nature discrète du processus de Dirichlet qui nous intéresse. Il nous permet de décrire les situations où les valeurs des contributions a_j à l'intensité peuvent être identiques bien que générées par des occurrences positionnées différemment dans l'espace et dans le temps.

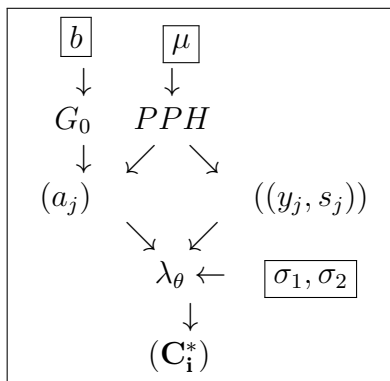


FIGURE 3.5 – Graphe orienté acyclique du modèle hiérarchique (cas iid)

Le modèle (3.11) est complété en considérant des lois *a priori* pour les hyperparamètres α , b , μ , σ_1 et σ_2 indiqués dans le graphe acyclique orienté de la figure 3.4. En les combinant à la vraisemblance globale (3.7), nous obtenons le modèle bayésien complet suivant :

$$p(\alpha)p(b)p(\mu)p(M|\mu)p(\sigma_1)p(\sigma_2)p(a_1, \dots, a_M | \alpha, b, M)L(x_1, \dots, x_{N_I}; \theta) \quad (3.12)$$

où $\theta = (\alpha, b, \mu, \sigma_1, \sigma_2, (a_j)_{j=1, \dots, M}, M)$ est le vecteur des paramètres du modèle. Soit $\tilde{\theta} = (\alpha, b, \mu, \sigma_1, \sigma_2)$ le quintuplet sur lequel porte l'inférence. La loi *a posteriori* du modèle $p(\tilde{\theta} | x_1, \dots, x_{N_I})$ émane de l'équation (3.12). Dans cette équation, la loi jointe des a_j est obtenue à partir de leur loi prédictive dont la représentation (Blackwell et Macqueen [8], 1973) est :

$$a_j | a_1, \dots, a_{j-1} \sim \frac{\alpha}{\alpha + j - 1} G_0 + \frac{1}{\alpha + j - 1} \sum_{k=1}^{j-1} \delta_{a_k}. \quad (3.13)$$

Ainsi, l'équation (3.13) nous dit que, conditionnellement aux $j - 1$ variables déjà obtenues, la probabilité que la prochaine valeur soit l'une de ces $j - 1$ valeurs est $\frac{1}{\alpha + j - 1}$, alors que la probabilité d'une toute nouvelle valeur est $\frac{\alpha}{\alpha + j - 1}$. Une réalisation d'un DP est donc discrète et cette propriété joue un rôle crucial via la classification dans les modèles de mélanges par processus de Dirichlet (Teh et al. [89], 2006).

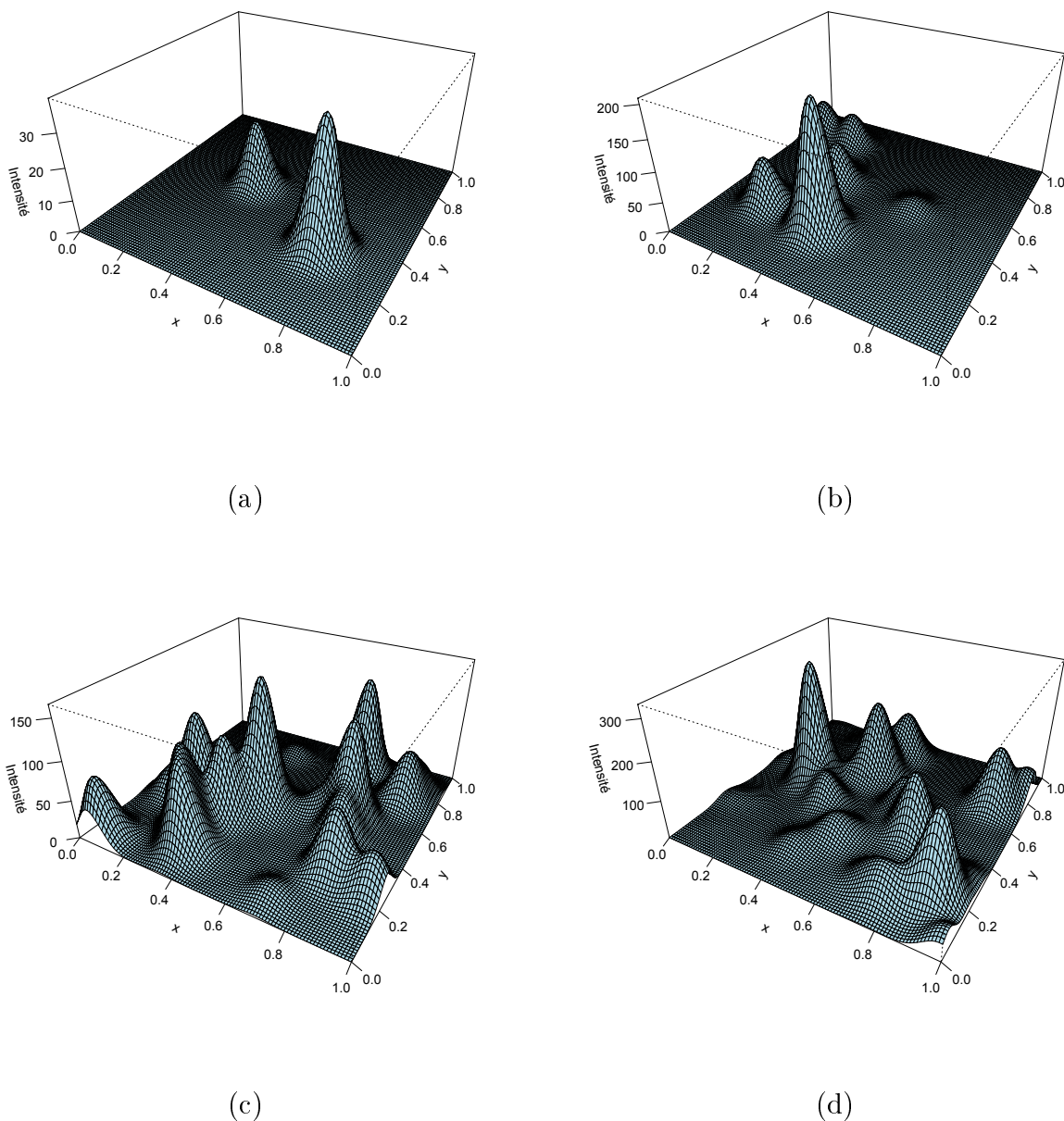


FIGURE 3.6 – Exemples d'intensité cumulée à la date $t = 1$ pour $(\alpha, b, \sigma_1, \sigma_2) = (1; 1; 0,05; 0,1)$: (a) deux contributions aléatoires ; (b) six contributions aléatoires ; (c) 18 contributions aléatoires ; (d) 32 contributions aléatoires.

Soulignons que lorsque α tend vers $+\infty$, la loi jointe asymptotique des $(a_j)_{j=1,\dots,M}$ conditionnelle à M est la mesure produit $G_0^{\otimes M}$, autrement dit, la loi jointe asymptotique des a_j est telle qu'ils sont indépendants et de même loi G_0 .

Dans le cadre des expérimentations numériques du paragraphe 7 et de l'étude du cas *Rubus fruticosus* du paragraphe 8, nous considérons les lois *a priori* suivantes : loi Gamma pour α et σ_2 , loi log-normale pour b et μ , loi uniforme pour σ_1 . Des exemples de réalisations d'intensité en dimension 2 sont présentés en figure 3.6. Les propriétés qualitatives peuvent être résumées de la manière suivante : μ est le nombre espéré de contributions du processus caché par unité de temps et d'espace ; b est l'amplitude espérée des contributions ; α est le degré d'indépendance entre contributions, en particulier plus α est proche de zéro, plus le nombre attendu de contributions ex-aequo est important ; σ_1 est le degré d'influence de la position spatiale d'une contribution sur l'intensité du processus observé tandis que σ_2 est le degré d'influence de la date d'occurrence d'une contribution. La sous-section suivante explicite des propriétés importantes du processus intensité stochastique du modèle.

3.3.2 Quelques propriétés distributionnelles

Pour mettre en place les algorithmes MCMC, il est nécessaire de préciser certaines caractéristiques du modèle telles que les intensités cumulées conditionnelles ou marginales. Les résultats distributionnels suivants permettent de proposer des valeurs initiales et des critères de mise à jour pour ces algorithmes.

Théorème 3.3.1

1) $\forall (x, t) \in X \times \mathbb{R}_+$, l'intensité cumulée en x à la date t est

$$\int_0^t \lambda_\theta(x, s) ds = \sum_{j=1}^{M_t} a_j \frac{1}{(2\pi\sigma_1^2)^{d/2}} \exp\left(-\frac{\|x - y_j\|^2}{2\sigma_1^2}\right) \left(1 - \exp\left(-\frac{t - s_j}{\sigma_2}\right)\right) \quad (3.14)$$

où $M_t = \sum_j \delta_{s_j}([0, t])$ suit une loi de Poisson de paramètre $\mu t \nu(X)$.

2) $\forall (B, t) \in \mathcal{B}_X \times \mathbb{R}_+$, $\Lambda_\theta(B \times [0, t])$ est une variable aléatoire d'espérance

$$\begin{aligned} E(\Lambda_\theta(B \times [0, t])) &= \mu b t \left(1 - \frac{\sigma_2}{t} \left(1 - \exp\left(-\frac{t}{\sigma_2}\right)\right)\right) \\ &\quad \times \int_X \int_B \frac{\exp\left(-\frac{\|x-y\|^2}{2\sigma_1^2}\right)}{(2\pi\sigma_1^2)^{d/2}} \nu(dx) \nu(dy) \end{aligned}$$

et de variance

$$\begin{aligned} V(\Lambda_\theta(B \times [0, t])) &= \mu b(1+b) \int_0^t \left(1 - \exp\left(-\frac{t-s}{\sigma_2}\right)\right)^2 ds \\ &\quad \times \int_X \left(\int_B \frac{\exp\left(-\frac{\|x-y\|^2}{2\sigma_1^2}\right)}{(2\pi\sigma_1^2)^{d/2}} \nu(dx) \right)^2 \nu(dy) \\ &\quad + \frac{1}{b(\alpha+1)} (E(\Lambda_\theta(B \times [0, t])))^2 \end{aligned}$$

Preuve 3.3.1

1) Pour tout (x, t) de $X \times \mathbb{R}_+$,

$$K_{(\sigma_1, \sigma_2)}((x, t), (y_j, s_j)) = \frac{\exp\left(-\frac{\|x-y_j\|^2}{2\sigma_1^2}\right)}{(2\pi\sigma_1^2)^{d/2}} \times \frac{1}{\sigma_2} \exp\left(-\frac{t-s_j}{\sigma_2}\right) \mathbb{1}_{\mathbb{R}_+}(t-s_j) \quad (3.15)$$

où $\|\cdot\|$ est la norme usuelle dans $X \subset \mathbb{R}^d$.

À partir de (3.9), (3.15) et en tenant compte du fait que les (y_j, s_j) sont générés par un processus de Poisson d'intensité μ , on obtient l'expression désirée pour $\int_0^t \lambda_\theta(x, s) ds$.

2) Si l'on considère maintenant la définition (3.5) et l'expression (3.14), on a

$$\forall (B, t) \in \mathcal{B}_X \times \mathbb{R}_+,$$

$$\Lambda_\theta(B \times [0, t]) = \sum_{j=1}^{M_t} a_j \left(1 - \exp\left(-\frac{t-s_j}{\sigma_2}\right)\right) \int_B \frac{\exp\left(-\frac{\|x-y_j\|^2}{2\sigma_1^2}\right)}{(2\pi\sigma_1^2)^{d/2}} \nu(dx) \quad (3.16)$$

donc l'espérance de $\Lambda_\theta(B \times [0, t])$ conditionnelle à M_t est

$$M_t b \int_0^t \frac{1}{t} \left(1 - \exp\left(-\frac{t-s}{\sigma_2}\right)\right) ds \int_X \frac{1}{\nu(X)} \int_B \frac{\exp\left(-\frac{\|x-y\|^2}{2\sigma_1^2}\right)}{(2\pi\sigma_1^2)^{d/2}} \nu(dx) \nu(dy)$$

car conditionnellement à M_t , les (y_j, s_j) sont distribués uniformément et les a_j sont d'espérance b . L'espérance non conditionnelle est alors obtenue en utilisant

$$E(M_t) = \mu t \nu(X).$$

De façon analogue, le moment d'ordre 2 de $\Lambda_\theta(B \times [0, t])$ conditionnel à M_t est

$$\begin{aligned} & M_t E(a_1^2) \int_0^t \frac{1}{t} \left(1 - \exp\left(-\frac{t-s}{\sigma_2}\right)\right)^2 ds \int_X \frac{1}{\nu(X)} \left(\int_B \frac{e\left(-\frac{\|x-y\|^2}{2\sigma_1^2}\right)}{(2\pi\sigma_1^2)^{d/2}} \nu(dx) \right)^2 \nu(dy) \\ & + M_t(M_t - 1) E(a_1 a_2) \left(\int_0^t \frac{1}{t} \left(1 - e^{-\frac{t-s}{\sigma_2}}\right) ds \int_X \frac{1}{\nu(X)} \int_B \frac{e^{-\frac{\|x-y\|^2}{2\sigma_1^2}}}{(2\pi\sigma_1^2)^{d/2}} \nu(dx) \nu(dy) \right)^2. \end{aligned}$$

$E(a_1^2) = b(1+b)$, $E(a_1 a_2) = b^2 + \frac{b}{\alpha+1}$ et $E(M_t(M_t - 1)) = (\mu t \nu(X))^2$ nous mènent au résultat final.

L'expression (3.14) intervient régulièrement dans le calcul du quotient de mise à jour de Métropolis-Hastings.

Les trois remarques suivantes nous prouvent que plus la bande passante spatiale σ_1 est petite, plus les positions des occurrences du processus observé coïncident avec celles du processus caché. De façon similaire, plus la bande passante temporelle σ_2 est petite, plus les dates d'occurrence du processus observé coïncident avec celles du processus caché.

Remarque 3.3.1 *Nous avons les convergences en loi suivantes :*

$$\forall (x, t) \in X \times \mathbb{R}_+^*,$$

$$\lambda_\theta(x, t) \xrightarrow[\sigma_1 \rightarrow 0]{\mathcal{L}} \sum_{j=1}^{M_t} a_j \delta_{y_j}(\{x\}) \times \frac{1}{\sigma_2} \exp\left(-\frac{t-s_j}{\sigma_2}\right) \quad (3.17)$$

$$\lambda_\theta(x, t) \xrightarrow[\sigma_2 \rightarrow 0]{\mathcal{L}} \sum_{j=1}^{M_t} a_j \frac{\exp\left(-\frac{\|x-y_j\|^2}{2\sigma_1^2}\right)}{(2\pi\sigma_1^2)^{d/2}} \times \delta_{s_j}(\{t\}) \quad (3.18)$$

$$\lambda_\theta(x, t) \xrightarrow[(\sigma_1, \sigma_2) \rightarrow (0,0)]{\mathcal{L}} \sum_{j=1}^{M_t} a_j \times \delta_{(y_j, s_j)}(\{x, t\}) \quad (3.19)$$

En effet, l'expression (3.15) fournit directement les résultats espérés en utilisant la convergence de la loi normale d'espérance y_j et de variance σ_1^2 vers la loi de Dirac en y_j quand σ_1 tend vers zéro. On utilise également la convergence de la loi exponentielle

d'espérance σ_2 vers la loi de Dirac quand σ_2 tend vers zéro.

Remarque 3.3.2 *Les expressions (3.17) et (3.18) conduisent directement aux limites de fonction intensité spatio-temporelle suivantes :*

1) $\forall t \in \mathbb{R}_+^*$,

- si $x \notin \{y_j, j \in \llbracket 1, M_t \rrbracket\}$, alors $\lambda_\theta(x, t) \xrightarrow[\sigma_1 \rightarrow 0]{\mathcal{L}} 0$.

- si $j \in \llbracket 1, M_t \rrbracket$ tel que $x = x_j$ alors $\lambda_\theta(x, t) \xrightarrow[\sigma_1 \rightarrow 0]{\mathcal{L}} a_j \frac{1}{\sigma_2} \exp\left(-\frac{t - s_j}{\sigma_2}\right)$.

2) $\forall x \in X$,

- si $t \notin \{s_j, j \in \llbracket 1, M_t \rrbracket\}$, alors $\lambda_\theta(x, t) \xrightarrow[\sigma_2 \rightarrow 0]{\mathcal{L}} 0$.

- si $j \in \llbracket 1, M_t \rrbracket$ tel que $t = t_j$, alors $\lambda_\theta(x, t) \xrightarrow[\sigma_2 \rightarrow 0]{\mathcal{L}} a_j \frac{\exp\left(-\frac{\|x - y_j\|^2}{2\sigma_1^2}\right)}{(2\pi\sigma_1^2)^{d/2}}$.

3.4 Génération de la loi *a posteriori*

Une approche pour obtenir la loi *a posteriori* des paramètres d'un modèle évoquant le DP est l'échantillonnage de Gibbs (Neal [63], 2000 ; Ishwaran et James [43], 2001). Pour ce qui concerne le modèle (3.11), nous adoptons une approche hybride incluant l'échantillonneur de Gibbs et les mises à jour du type Metropolis-Hastings. Pour la mise à jour du nombre d'événements cachés, une chaîne de Markov réversible est utilisée.

Notons $q((C_i^*), \theta)$ la densité jointe des paramètres et des observations. Soit $\theta^{(t)} = \left(\alpha^{(t)}, b^{(t)}, \mu^{(t)}, \sigma_1^{(t)}, \sigma_2^{(t)}, M^{(t)}, a_1^{(t)}, \dots, a_{M^{(t)}}^{(t)}\right)$ la valeur de θ à la t -ième itération de la méthode MCMC. L'inférence est effectuée en utilisant l'algorithme 4.

On remarquera que les (y_j, s_j) constituent la réalisation d'un processus de Poisson homogène de paramètre μ et que conditionnellement à M , les (y_j, s_j) sont répartis uniformément dans $X \times [0, T]$ (T étant la date de fin d'observation).

Algorithme 4: RJMCMC pour l'inférence sur $(\alpha, b, \mu, \sigma_1, \sigma_2)$ du modèle (3.11)

Pour $t = 1, \dots, N$

- Générer $\log(\alpha^{(t+1)})$ selon la loi $\mathcal{N}(\log(\alpha^{(t)}), \xi_1)$
- Retenir $\alpha^{(t+1)}$ avec la probabilité

$$\frac{q\left((C_i^*), \left(\alpha^{(t+1)}, b^{(t)}, \mu^{(t)}, \sigma_1^{(t)}, \sigma_2^{(t)}, M^{(t)}, a_1^{(t)}, \dots, a_{M^{(t)}}^{(t)}\right)\right)}{q\left((C_i^*), \left(\alpha^{(t)}, b^{(t)}, \mu^{(t)}, \sigma_1^{(t)}, \sigma_2^{(t)}, M^{(t)}, a_1^{(t)}, \dots, a_{M^{(t)}}^{(t)}\right)\right)}$$

- Générer $\log(b^{(t+1)})$ selon la loi $\mathcal{N}(\log(b^{(t)}), \xi_2)$
- Retenir $b^{(t+1)}$ avec la probabilité

$$\frac{q\left((C_i^*), \left(\alpha^{(t+1)}, b^{(t+1)}, \mu^{(t)}, \sigma_1^{(t)}, \sigma_2^{(t)}, M^{(t)}, a_1^{(t)}, \dots, a_{M^{(t)}}^{(t)}\right)\right)}{q\left((C_i^*), \left(\alpha^{(t+1)}, b^{(t)}, \mu^{(t)}, \sigma_1^{(t)}, \sigma_2^{(t)}, M^{(t)}, a_1^{(t)}, \dots, a_{M^{(t)}}^{(t)}\right)\right)}$$

- Générer $\log(\mu^{(t+1)})$ suivant la loi $\mathcal{N}(\log(\mu^{(t)}), \xi_3)$
- Retenir $\mu^{(t+1)}$ avec la probabilité

$$\frac{q\left((C_i^*), \left(\alpha^{(t+1)}, b^{(t+1)}, \mu^{(t+1)}, \sigma_1^{(t)}, \sigma_2^{(t)}, M^{(t)}, a_1^{(t)}, \dots, a_{M^{(t)}}^{(t)}\right)\right)}{q\left((C_i^*), \left(\alpha^{(t+1)}, b^{(t+1)}, \mu^{(t)}, \sigma_1^{(t)}, \sigma_2^{(t)}, M^{(t)}, a_1^{(t)}, \dots, a_{M^{(t)}}^{(t)}\right)\right)}$$

- Générer $\sigma_1^{(t+1)}$ selon la loi $\mathcal{N}(\log(\sigma_1^{(t)}), \xi_4)$
- Retenir $\sigma_1^{(t+1)}$ avec la probabilité

$$\frac{q\left((C_i^*), \left(\alpha^{(t+1)}, b^{(t+1)}, \mu^{(t+1)}, \sigma_1^{(t+1)}, \sigma_2^{(t)}, M^{(t)}, a_1^{(t)}, \dots, a_{M^{(t)}}^{(t)}\right)\right)}{q\left((C_i^*), \left(\alpha^{(t+1)}, b^{(t)}, \mu^{(t+1)}, \sigma_1^{(t)}, \sigma_2^{(t)}, M^{(t)}, a_1^{(t)}, \dots, a_{M^{(t)}}^{(t)}\right)\right)}$$

- Générer $\sigma_2^{(t+1)}$ selon la loi $\mathcal{N}(\log(\sigma_2^{(t)}), \xi_5)$
- Retenir $\sigma_2^{(t+1)}$ avec la probabilité

$$\frac{q\left((C_i^*), \left(\alpha^{(t+1)}, b^{(t+1)}, \mu^{(t+1)}, \sigma_1^{(t+1)}, \sigma_2^{(t+1)}, M^{(t)}, a_1^{(t)}, \dots, a_{M^{(t)}}^{(t)}\right)\right)}{q\left((C_i^*), \left(\alpha^{(t+1)}, b^{(t)}, \mu^{(t+1)}, \sigma_1^{(t+1)}, \sigma_2^{(t)}, M^{(t)}, a_1^{(t)}, \dots, a_{M^{(t)}}^{(t)}\right)\right)}$$

- Générer $M^{(t+1)}$ selon $\frac{1}{2} (\delta_{M^{(t)}+1} + \delta_{M^{(t)}-1})$

- Générer $(a_1^{(t+1)}, \dots, a_{M^{(t+1)}}^{(t+1)})$ de la façon suivante

- si $M^{(t+1)} = M^{(t)} - 1$ par une suppression aléatoire uniforme dans $\{a_1^{(t)}, \dots, a_{M^{(t)}}^{(t)}\}$
- si $M^{(t+1)} = M^{(t)} + 1$ par un ajout aléatoire selon la loi

$$\frac{\alpha^{(t+1)}}{\alpha^{(t+1)} + M^{(t)}} \text{Gamma}(1, b^{(t+1)}) + \frac{1}{\alpha^{(t+1)} + M^{(t)}} \sum_{k=1}^{M^{(t)}} \delta_{a_k^{(t)}}$$

- Générer $((y_1, s_1)^{(t+1)}, \dots, (y_{M^{(t+1)}}, s_{M^{(t+1)}})^{(t+1)})$ de la façon suivante

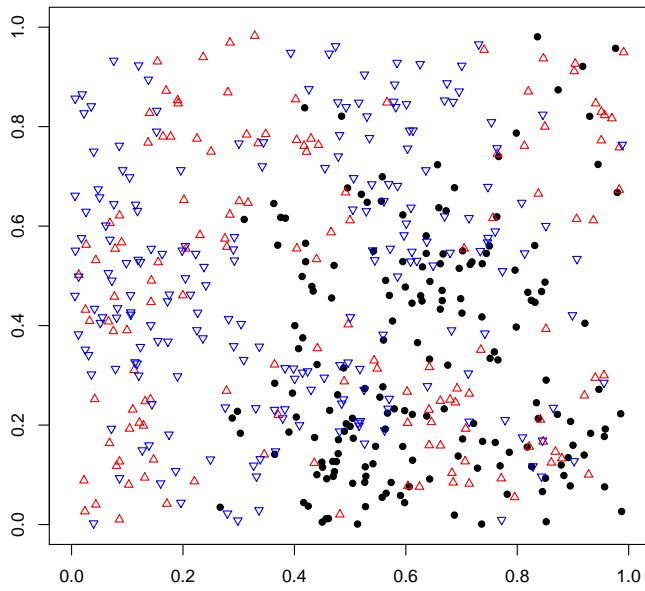
- si $M^{(t+1)} = M^{(t)} - 1$ par la suppression du (y_j, s_j) correspondant au a_j supprimé précédemment
- si $M^{(t+1)} = M^{(t)} + 1$ par un ajout $(y_{M^{(t+1)}}, s_{M^{(t+1)}})^{(t+1)}$ selon la loi $\mathcal{U}(X \times [0, T])$

3.5 Applications

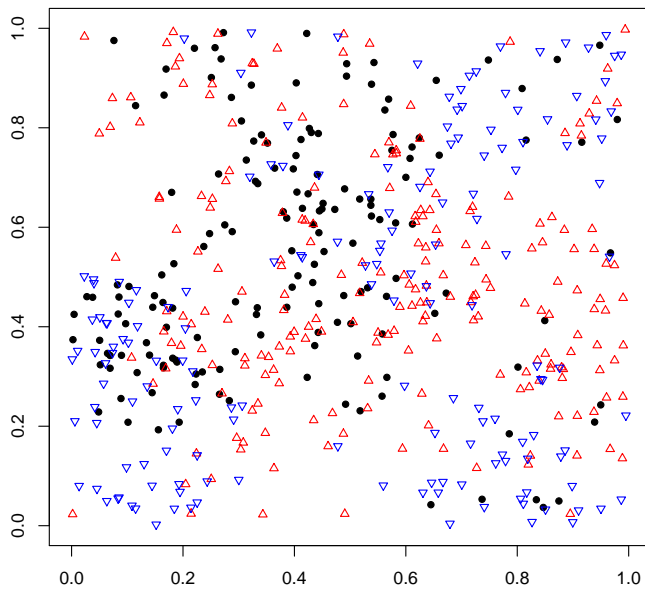
Dans les paragraphes suivants, nous présentons les résultats obtenus avec cet algorithme sur des données artificielles et des données réelles.

3.5.1 Traitement de données artificielles

Nous générons artificiellement le modèle avec des paramètres connus afin de tester l'efficacité de la méthode. Nous appliquerons la méthode à des données réelles dans le paragraphe suivant. Nous avons généré plusieurs jeux de données artificielles. Le modèle de processus ponctuel spatio-temporel défini par (3.11) peut être simulé en générant d'abord le processus intensité, puis en appliquant une méthode d'amincissement sur un processus de Poisson homogène approprié (Møller et Schoenberg, [60] 2010). Nous avons effectué des simulations numériques de ce modèle sur le carré unité de \mathbb{R}^2 et l'intervalle de temps $[0; 3]$. Plusieurs vecteurs de paramètres ont été testés en prenant en compte certaines caractéristiques des paramètres considérés et en se basant sur les résultats du paragraphe 3.3.2. Par exemple, un paramètre de concentration proche de 0 conduit à de nombreuses contributions ex-æquo, tandis que s'il est très élevé, les contributions sont plutôt indépendantes. Le vecteur de paramètres retenu pour illustrer les différentes expérimentations numériques est $\tilde{\theta} = (15; 45; 10; 0, 15; 0, 10)$. Ceci correspond à un nombre attendu d'événements cachés de 10 avec une contribution espérée de 45, les bandes passantes spatiale et temporelle étant respectivement 0, 15 et 0, 10. La procédure MCMC du paragraphe 3.4 a été appliquée avec $N = 50000$ itérations à chacun de ces jeux de données. Les figures 3.7, 3.8, 3.9, 3.10 et la table 3.1 nous montrent respectivement la carte de répartition spatio-temporelle, les résultats des chaînes obtenues, la loi *a posteriori* de $\tilde{\theta}$ et les tables de résultats pour deux de ces jeux de données. Les valeurs réelles des paramètres α, b et μ sont dans les intervalles de confiance à 95%. C'est aussi le cas des paramètres σ_1 et σ_2 pour le premier jeu de données. Cependant, pour le second jeu, il faut passer à un niveau de sécurité plus élevé (au moins 96,4%) pour que la valeur réelle de σ_1 soit dans l'intervalle de confiance. Par contre, la convergence semble beaucoup plus lente pour le paramètre σ_2 . Le fait que a soit un paramètre dans un espace de dimension variable, peut expliquer le défaut d'exploration optimal comme l'ont précisé Green et Richardson ([38], 2002). Une amélioration de l'algorithme que nous avons utilisé est en cours en s'inspirant des travaux de Ehlers et Brooks ([31], 2008), en utilisant des schémas adaptatifs appropriés.

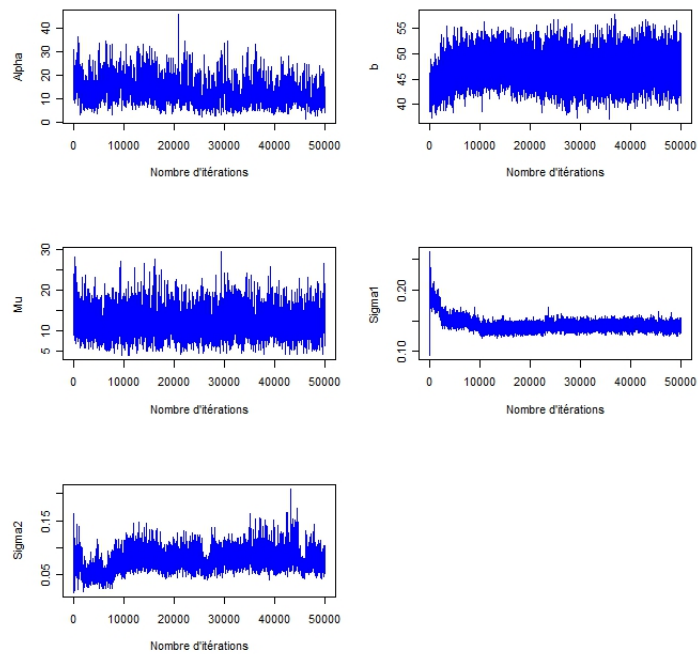


(a)

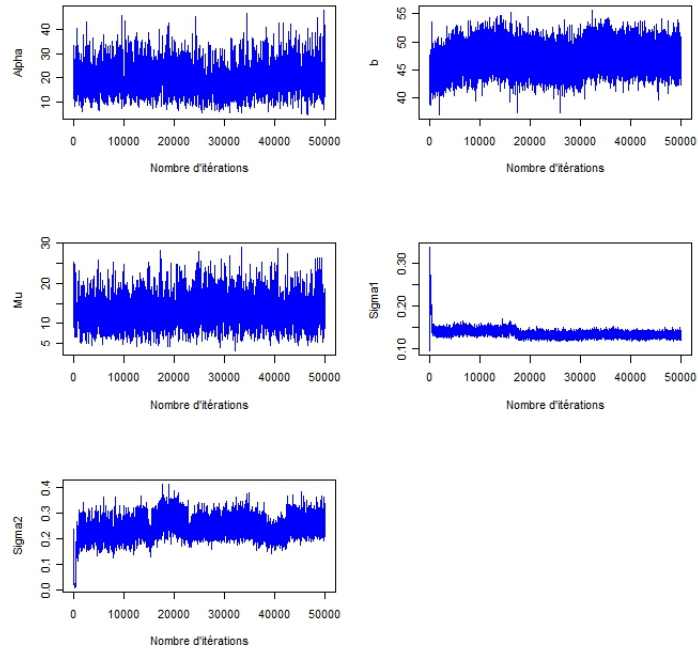


(b)

FIGURE 3.7 – Exemples de cartes de données artificielles pour $(\alpha, b, \mu, \sigma_1, \sigma_2) = (2; 20; 40; 0, 1; 0, 1)$ à 3 dates d'observation : date 1 : \bullet , date 2 : \triangle , date 3 : ∇ .

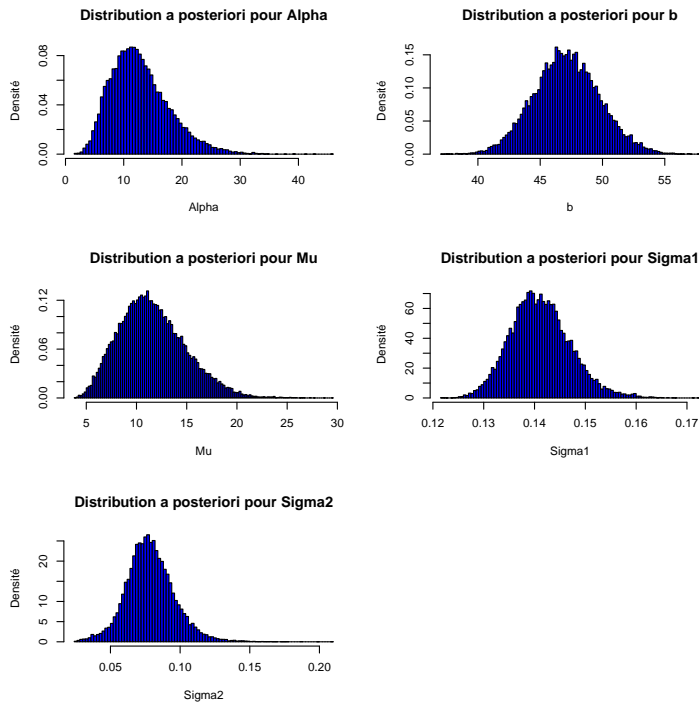


(a)

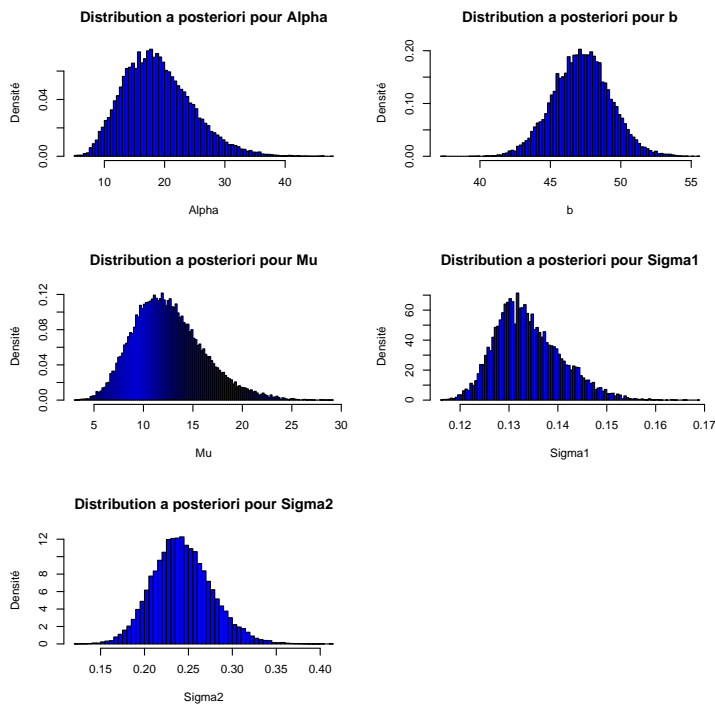


(b)

FIGURE 3.8 – Résultats MCMC obtenus par procédure hybride Gibbs-Métropolis-Hastings : (a) pour fig. 3.7a ; (b) pour fig. 3.7b

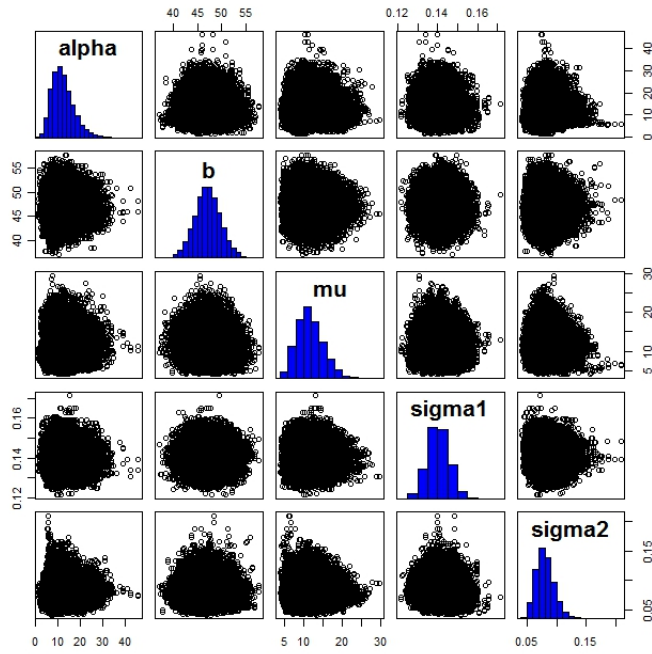


(a)

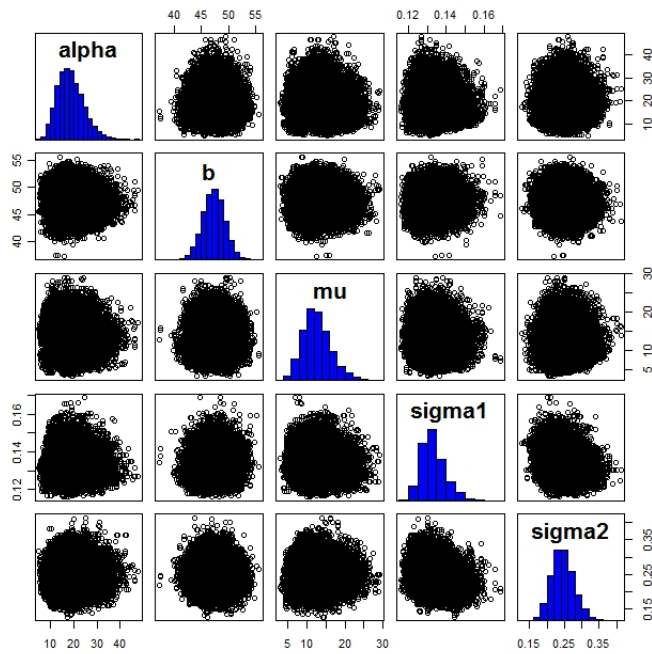


(b)

FIGURE 3.9 – Lois *a posteriori* obtenues pour les paramètres pour 50000 itérations et une période de chauffe de 5000 : (a) pour fig. 3.7a ; (b) pour fig. 3.7b



(a)



(b)

FIGURE 3.10 – Nuage de points 2 à 2 des paramètres et histogramme sur la diagonale : (a) pour fig. 3.7a ; (b) pour fig. 3.7b

	Moyenne	Médiane	Écart-type	Quantile 2,5%	Quantile 97,5%
α	12,827	12,164	4,990	5,093	24,405
b	47,127	47,092	2,640	42,060	52,460
μ	11,600	11,285	3,304	6,072	18,770
α_1	0,141	0,141	0,006	0,130	0,154
α_2	0,079	0,078	0,017	0,045	0,116

a) Traitement des données artificielles de fig. 3.7a

	Moyenne	Médiane	Écart-type	Quantile 2,5%	Quantile 97,5%
α	18,898	18,340	5,666	9,499	31,610
b	47,230	47,231	2,029	43,249	51,226
μ	12,556	12,198	3,548	6,641	20,556
σ_1	0,134	0,133	0,007	0,123	0,149
σ_2	0,243	0,241	0,034	0,181	0,315

b) Traitement des données artificielles de fig. 3.7b

TABLE 3.1 – Statistiques élémentaires pour la loi *a posteriori*

3.5.2 Traitement des données de *Rubus fruticosus*

Nous avons utilisé les données de référence de *Rubus fruticosus* (fig. 3.11) qui correspondent à la répartition spatio-temporelle de ronces dans un carré expérimental (Diggle et Milne [27],1983). Les positions spatiales aux 3 dates d'observation concernées sont indiquées en figure 3.12. Selon Padilla et al. ([71], 2012), certains mammifères et oiseaux sont responsables de la dispersion spatiale des ronces. Le mécanisme de dispersion se fait à travers l'absorption puis la déjection des semences. En outre, ces déjections peuvent se faire à une distance plus ou moins importante du lieu de cueillette. Ce phénomène peut être considéré d'un point de vue mathématique comme la réalisation d'un processus stochastique caché. μ est le nombre espéré d'événements contribuant à la répartition des semences, b traduit le niveau espéré de cette contribution et α , leur degré d'indépendance. σ_1 et σ_2 sont respectivement les degrés d'influence spatial et temporel de ces événements. Les lois *a priori* sont les suivantes : pour le paramètre de confiance α , nous avons pris une loi *Gamma* de paramètre de forme 9 et de paramètre d'échelle 2, pour l'amplitude espérée b des contributions, une loi *log-normale* de paramètres 0 et 1, pour le nombre attendu μ de contributions cachées par unité de temps et d'espace,

FIGURE 3.11 – Ronces *Rubus fruticosus*

une loi *log-normale* de paramètres 1 et 1, ce qui sous-entend un nombre espéré de 4,5 pour les occurrences cachées du shot noise généralisé par unité de temps et d'espace, avec un coefficient de variation pour μ de 1,3; le degré d'influence spatial σ_1 du processus caché a pour loi *a priori* la loi *Uniforme* sur $[0, 3]$ et celle du degré d'influence temporel σ_2 est la loi Gamma de paramètres 1 et 1. Les résultats obtenus sont présentés dans la table 3.2 et les figures 3.13, 3.14 et 3.15. L'estimateur bayésien du nombre espéré d'événements cachés est proche de 26 avec une contribution estimée de 46,7 et un paramètre de confiance proche de 4. Le paramètre d'influence spatial est estimé à 0,074, ce qui correspond à une agrégation spatiale modérée mais plus prononcée que l'agrégation temporelle qui est estimée à 1,170. Nous avons effectué un test du rapport de vraisemblance pour modèles emboîtés (Spiegelhalter et al. [84], 2002) appliqué à l'hypothèse H_0 selon laquelle les contributions cachées sont indépendantes de même loi Gamma contre l'hypothèse du modèle (3.11) où elles forment une réalisation d'un processus de Dirichlet centré sur une loi Gamma. Le rapport de vraisemblance obtenu vaut 0,277 avec une p-value de 0,109 et conduit au non rejet de H_0 . Cela signifierait que les événements cachés ont des impacts indépendants sur la densité des ronces *Rubus fruticosus*. Notons qu'à notre connaissance, les deux seules publications concernant le traitement statistique de ces données spatio-temporelles de *Rubus fruticosus* sont de Hutchings ([42], 1978) et de Diggle ([24], 1983), comme souligné par Perry et al. ([73], 2006). Hutchings a utilisé, pour chacune des trois dates d'observation, la méthode du

	Moyenne	Médiane	Écart-type	Quantile 2,5%	Quantile 97,5%
α	4,331	3,998	1,873	1,607	8,791
b	46,715	46,506	4,420	38,623	55,900
μ	26,330	25,919	5,187	17,252	37,536
σ_1	0,074	0,074	0,003	0,068	0,082
σ_2	1,170	1,169	0,108	0,962	1,382

TABLE 3.2 – Statistiques élémentaires de la loi *a posteriori* pour les données de *Rubus fruticosus*

plus proche voisin et le calcul de l'indice de Clark et Evans et a trouvé une forte agrégativité dans la répartition des ronces à chaque date. Diggle, quant à lui, n'a utilisé que les données à la date 3 et les a ajusté à un processus de Poisson en groupes ainsi qu'à un processus à interactions appariées. Ces deux modèles étant rejetés, il propose un modèle de processus de Poisson en groupes mais aminci par un processus de Poisson homogène de rayon d'action fixé. Les paramètres du modèle sont estimés par moindres carrés comme pour les deux modèles précédents et conduisent à un bon ajustement du modèle. D'autres articles concernent l'étude statistique de la répartition spatiale des ronces *Rubus fruticosus* à partir d'autres jeux de données et se limitent à des tests d'agrégativité spatiale (Sullivan et al. [86], 2006). Notre approche, elle, permet de traiter simultanément les trois dates d'observation et de tester l'effet d'occurrences cachées sur l'intensité spatio-temporelle des présences de ronces tout en intégrant le caractère agrégatif de la répartition de *Rubus fruticosus*.

3.6 Conclusion sur le modèle spatio-temporel

Nous avons proposé un modèle de processus de Cox ayant un processus intensité de type shot noise généralisé associé à un processus de Dirichlet. Il permet de modéliser un phénomène consistant en des occurrences à des positions spatiales et des dates provoquées par des événements cachés se produisant selon un processus de Poisson. La contribution de chacun de ces événements suit un processus de Dirichlet de loi espérance Γ . Un algorithme MCMC hybride Gibbs-Métropolis-Hastings a été proposé et permet d'obtenir la loi *a posteriori* des paramètres suivants : nombre attendu d'événements cachés par unité d'espace et de temps, contribution espérée par événe-

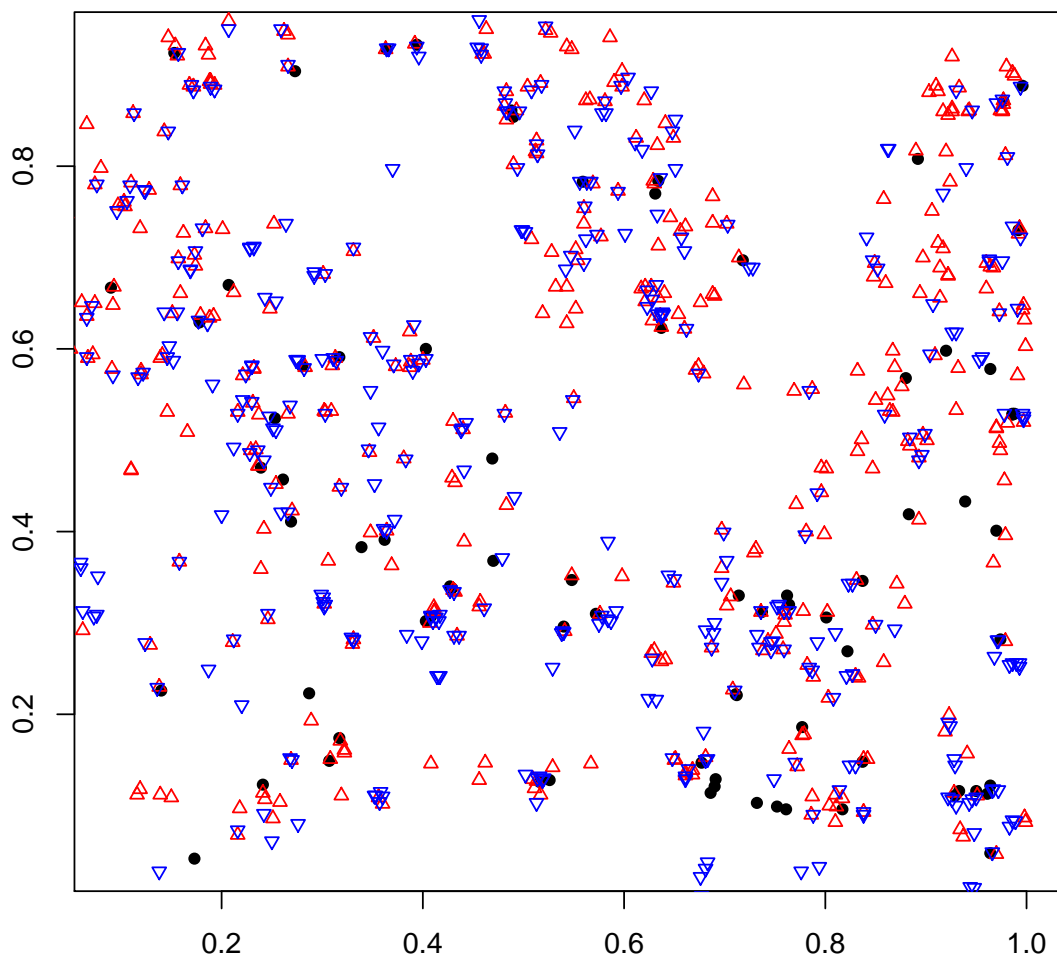


FIGURE 3.12 – Répartition spatiale des ronces *Rubus fruticosus* à 3 dates d'observation :
date 1 : ●, date 2 : △, date 3 : ▽

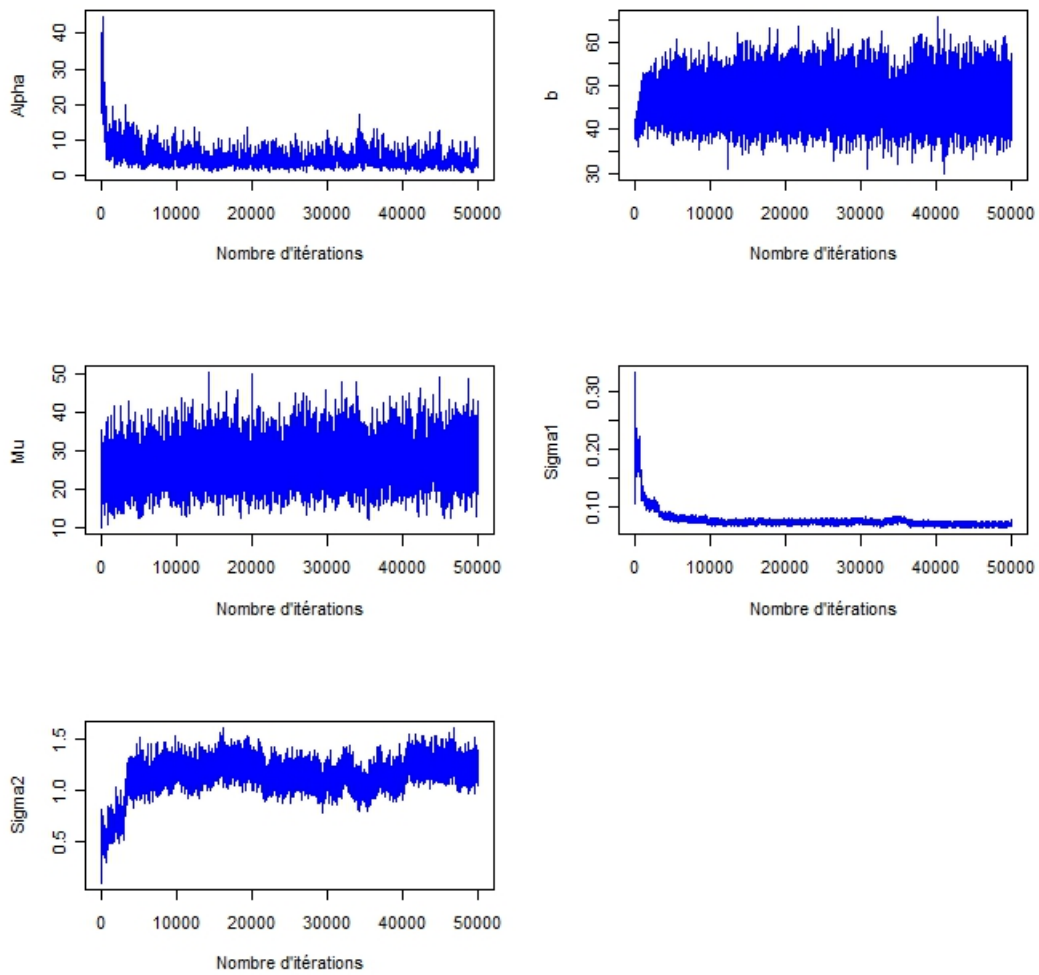


FIGURE 3.13 – Résultats MCMC obtenus par procédure hybride Gibbs-Métropolis-Hastings (Ronces *Rubus Fruticosus*)

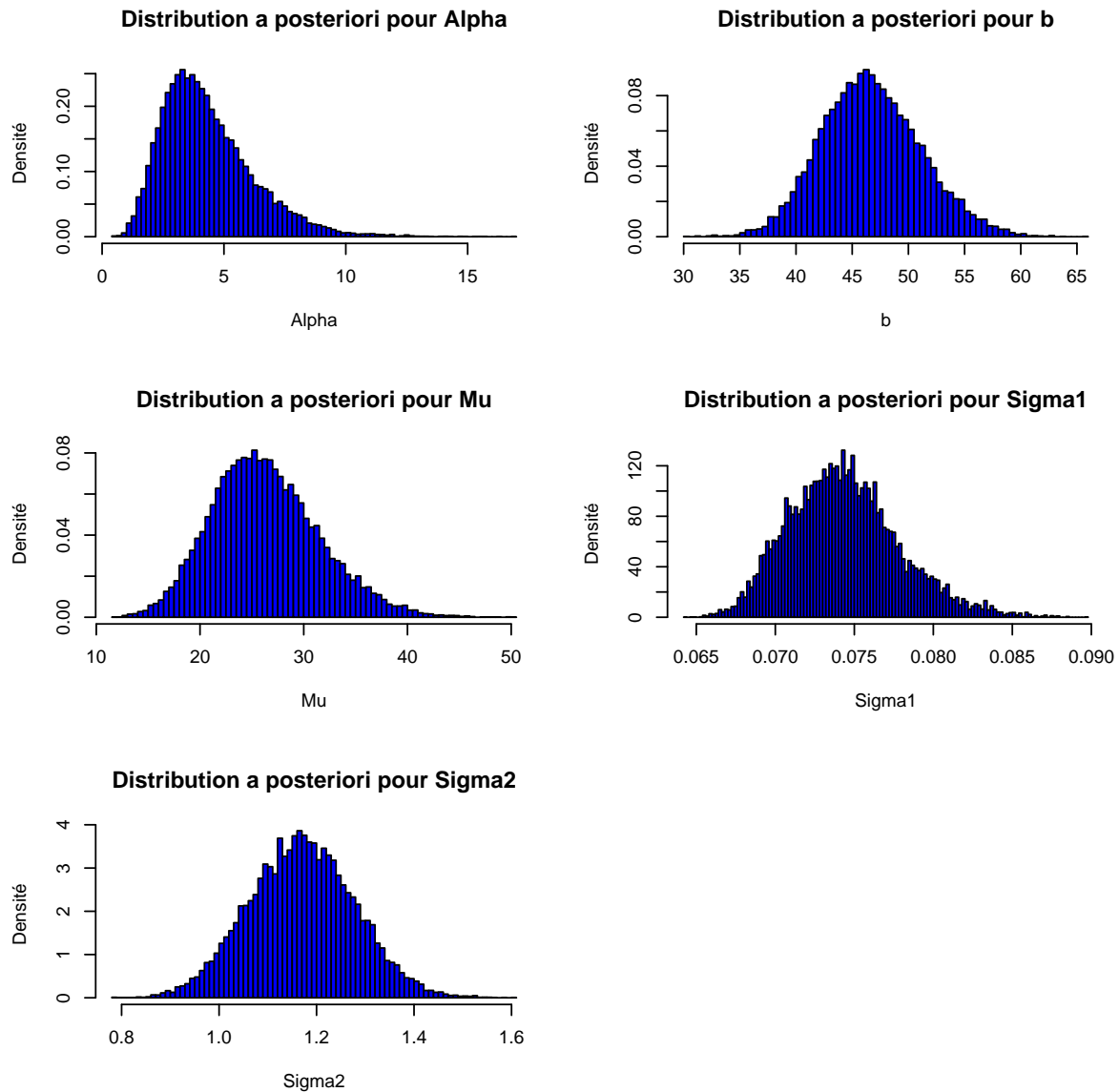


FIGURE 3.14 – Lois *a posteriori* obtenues pour les paramètres pour 50000 itérations et une période de chauffe de 5000 (Ronces *Rubus Fruticosus*)

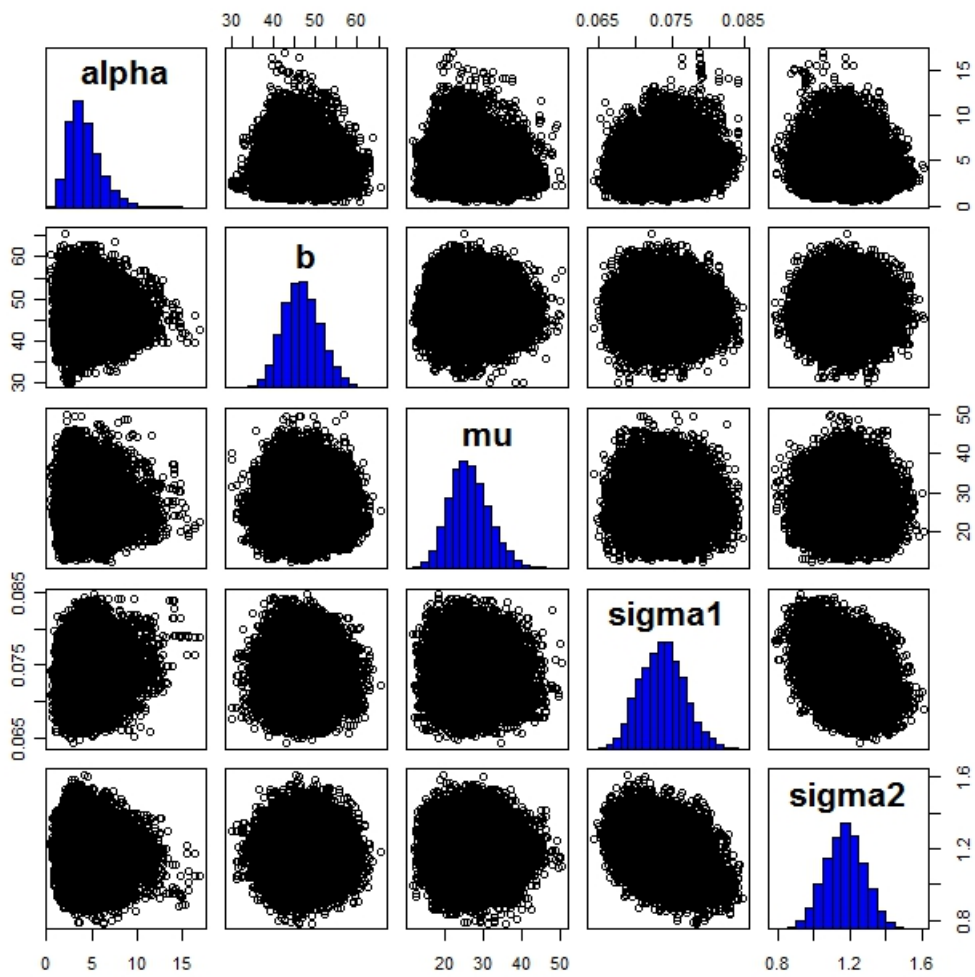


FIGURE 3.15 – Nuage de points 2 à 2 des paramètres et histogramme sur la diagonale pour les données de Ronces *Rubus Fruticosus*

ment caché, degré d'influence spatiale d'événements cachés, degré d'influence temporelle d'événements cachés et degré de corrélation des contributions des événements cachés. L'application à des données artificielles et à des données réelles montre le potentiel de la méthode.

Ce chapitre a fait l'objet d'une communication en congrès international (Valmy et Vaillant [102], 2012) et un article est en révision (Chimard et al. [15], 2012).

ASSOCIATION SPATIALE DES PROCESSUS DE COX, SHOT-NOISE ET DIRICHLET

4.1 Statistique d'un processus de Cox spatial associé à un processus de Dirichlet

En écologie et épidémiologie, la distribution d'événements tels que l'occurrence d'une maladie, l'arrivée de prédateurs ou la localisation de plantes peut être considérée comme la réalisation d'un processus ponctuel spatial, pour lequel chaque point représente un unique événement. La théorie des processus ponctuels a été développée et discutée par Brémaud ([9], 1981), Karr ([49], 1991) et Daley et Veres-Jone ([22], 2003). Certains ouvrages (Cliff et Ord [16], 1981 ; Ripley [77], 1981 ; Diggle et Milne [27], 1983) présentent des procédures statistiques pour l'analyse des réalisations d'un processus ponctuel et plusieurs publications (Diggle et al. [26], 2005 ; Brix et Chadoëuf [12], 2002 ; Vaillant et al. [96], 2011) traitent des applications dans des cas particuliers.

Certaines études sont basées sur les comptages d'événements dans des unités d'échantillonnage (Comas et al. [17], 2011), d'autres sur les positions spatiales d'événements ou les dates d'occurrence (Vaillant et al. [96], 2011), et aussi sur la distance d'échantillonnage (Melville et Welsh [57], 2001). Perry et al. ([73], 2006) ont discuté d'une sélection appropriée et de l'utilisation d'une méthode pour analyser des modèles ponctuels spatiaux en écologie. Nous pouvons nous référer à Møller et Waagepetersen ([62], 2007) pour les outils statistiques des processus ponctuels spatiaux. Comas et al. ([17], 2011) ont récemment discuté de processus ponctuels spatiaux pour un inventaire forestier mettant en avant la sur-dispersion.

Dans les cas de sur-dispersion, la modélisation du processus de Cox est un choix pra-

tique dans la mesure où cette classe de processus ponctuels est assez vaste pour prendre en considération de nombreuses caractéristiques. Ainsi, Zhang et Kou ([111], 2010) ont présenté plusieurs domaines scientifiques dans lesquels le processus de Cox, communément appelé processus doublement poissonien, se produisent. Le processus intensité $\lambda(\cdot)$ d'un tel processus sur un espace mesuré (X, \mathcal{B}, ν) est défini dans l'expression de la mesure intensité $\Lambda(\cdot)$ du processus de Cox :

$$\Lambda(B) = \int_B \lambda(x) \nu(dx), \quad (4.1)$$

pour tout élément B de \mathcal{B} . Dans l'expression (3.14), $\Lambda(B)$ est une variable aléatoire correspondant au nombre espéré de points dans B .

Ainsi, la modélisation d'un processus de Cox est équivalente à la modélisation soit de son intensité, soit de sa mesure dirigeante (Vaillant [91], 1991). Par exemple, Brix et Chadœuf ([12], 2002) ont décrit une classe spécifique de processus de Cox, appelé processus de Cox G shot-noise, en modélisant $\Lambda(\cdot)$ avec une mesure G shot-noise.

Pour notre part, nous considérons un processus shot-noise généralisé N sur un espace mesuré (Møller et Torrisi [61], 2005). Le modèle est étendu au cas où il existe des contributions à l'intensité du processus associé à un processus de Dirichlet (Ferguson [32, 33], 1973, 1974). Nous nous focalisons sur des procédures fournissant des outils d'inférence bayésienne pour des observations de données de comptages spatiaux. Ensuite, nous présentons les propriétés distributionnelles de comptages observés dans un échantillon spatial par rapport à la bande passante d'influence sur les événements spatiaux cachés. Puis, nous présentons deux études de cas : la première concerne les données de comptages spatiaux de *Ibicella Lutea*, la seconde est basée sur le nombre de décès par canton en Georgia, USA, causé soit par la grippe et la pneumonie, soit par des maladies respiratoires chroniques. Après avoir effectué des tests de dispersion et d'autocorrélation, nous nous intéressons à l'estimation du nombre espéré d'événements cachés, leur contribution espérée à l'intensité et d'autres paramètres du modèle.

4.2 Quelques propriétés distributionnelles

4.2.1 Cas général

Soit N un processus de Cox et λ son intensité sur l'espace X défini par :

$$\forall x \in X, \quad \lambda(x) = \sum_{j=1}^M a_j K_\sigma(x, y_j) \quad (4.2)$$

où K_σ est un noyau tel que $K_\sigma(\cdot, y)$ soit une fonction de densité de probabilité sur X pour tout y dans X et M est un nombre aléatoire de contributions. Les (y_j, a_j) sont une réalisation d'un processus ponctuel marqué L sur X avec des marques réelles positives a_j identiquement distribuées selon une loi de probabilité G sur \mathbb{R}_+ , et indépendantes des y_j .

Si L est un processus de Poisson homogène d'intensité μ et les a_j sont égaux, alors λ est un processus shot-noise standard (Møller [58], 2003), et N est un processus de Cox shot-noise. De plus, si la bande passante de K_σ est une variable aléatoire, alors N est un processus de Cox shot-noise généralisé (Møller et Torrisi [61], 2005). Dans l'expression (4.2), a_j est la contribution de l'événement y_j à l'intensité λ .

Nous supposons que les contributions a_j suivent un processus de Dirichlet G (Ferguson [32, 33], 1973 et 1974) de mesure G_0 et de paramètre de concentration α . Cependant, nous n'avons pas de condition d'indépendance sur les a_j comme pour beaucoup de modèles mais une faible condition d'échangeabilité fournie par le théorème de De Finetti (Diaconis et Freedman [23], 1980). En fait, les contributions a_j sont indépendantes conditionnellement à G . Ceci nous permet de considérer une corrélation entre les marques et des effets environnementaux mixtes. Les y_j sont les localisations d'occurrence d'un processus de Poisson homogène de paramètre μ , noté $HPP(\mu)$. En d'autres termes,

$$\begin{aligned} (a_j) \mid G &\stackrel{i.i.d}{\sim} G \\ (y_j) &\sim HPP(\mu) \\ G &\sim DP(\alpha, G_0). \end{aligned} \quad (4.3)$$

Les propositions suivantes donnent quelques propriétés distributionnelles de la mesure de comptage associée à N par rapport au paramètre de concentration α .

Proposition 4.2.1 *Soit N un processus de Cox sur un espace mesuré (X, \mathcal{B}, ν) d'intensité définie par les équations (4.2) et (4.3), alors*

$$\forall A \in \mathcal{B}, E(N(B)) = \mu E(a_1) \int_X \int_B K_\sigma(x, y) \nu(dx) \nu(dy)$$

et

$$V(N(B)) = E(N(B)) \left(1 - \frac{1}{\alpha + 1} E(N(B)) \right) + \mu E(a_1^2) \int_X \left(\int_B K_\sigma(x, y) \nu(dx) \right)^2 \nu(dy).$$

Preuve 4.2.1 Considérant (3.14) et (4.2), nous obtenons pour tout élément B de \mathcal{B} ,

$$\Lambda(B) = \sum_{j=1}^M a_j \int_B K_\sigma(x, y_j) \nu(dx). \quad (4.4)$$

Puisque les y_j sont uniformément distribués sur X conditionnellement à M , l'espérance de $\Lambda(B)$ conditionnellement à M est

$$ME(a_1) \int_X \frac{1}{\nu(X)} \int_B K_\sigma(x, y) \nu(dx) \nu(dy)$$

et le résultat final pour $E(N(B))$ est obtenu à partir de $E(M) = \mu \nu(X)$.

De même, le moment d'ordre deux de $\Lambda(B)$ conditionnellement à M est

$$ME(a_1^2) \int_X \frac{1}{\nu(X)} \left(\int_B K_\sigma(x, y) \nu(dx) \right)^2 \nu(dy) + M(M-1) E(a_1 a_2) \left(\int_X \frac{1}{\nu(X)} \int_B K_\sigma(x, y) \nu(dx) \nu(dy) \right)^2.$$

$E(a_1 a_2) = (E(a_1))^2 \alpha / (\alpha + 1)$; $E(M(M-1)) = (\mu \nu(X))^2$ et $V(N(B)) = V(\Lambda(B)) + E(\Lambda(B))$ nous mène au résultat final.

Proposition 4.2.2 *Sous les mêmes conditions que celles de la proposition (4.2.1), soit (B_1, B_2) un élément de \mathcal{B}^2 avec $B_1 \cap B_2 = \emptyset$. La covariance entre $N(B_1)$ et $N(B_2)$ est*

$$\begin{aligned} Cov(N(B_1), N(B_2)) &= \mu E(a_1^2) \int_X \left(\int_{B_1} K_\sigma(z, y) \nu(dz) \int_{B_2} K_\sigma(x, y) \nu(dx) \right) \nu(dy) \\ &\quad - \frac{1}{\alpha + 1} E(N(B_1)) E(N(B_2)). \end{aligned}$$

Preuve 4.2.2 Puisque N est un processus de Cox, nous avons

$$Cov(N(B_1), N(B_2)) = Cov(\Lambda(B_1), \Lambda(B_2)).$$

Par ailleurs,

$$\begin{aligned} E(\Lambda(B_1)\Lambda(B_2)) &= E\left(\sum_{j=1}^M a_j \int_{B_1} K_\sigma(x, y_j) \nu(dx) \sum_{k=1}^M a_k \int_{B_2} K_\sigma(x, y_k) \nu(dx)\right) \\ &= E\left(ME(a_1^2)E\left(\int_{B_1} K_\sigma(x, y_1) \nu(dx) \int_{B_2} K_\sigma(x, y_1) \nu(dx)\right)\right) \\ &\quad + E\left(M(M-1)E(a_1 a_2)E\left(\int_{B_1} K_\sigma(x, y_1) \nu(dx) \int_{B_2} K_\sigma(x, y_2) \nu(dx)\right)\right) \end{aligned}$$

Le résultat final est obtenu à partir de

$$E(M) = \mu\nu(X), E(a_1 a_2) = (E(a_1))^2 \alpha / (\alpha + 1) \text{ et } E(M(M-1)) = (\mu\nu(X))^2.$$

Considérons des données de comptages dans r sous-ensembles disjoints de X . Quand G_0 appartient à une famille de densités de probabilité paramétrée par b et K_σ une famille de noyaux paramétrée par σ , nous avons le résultat suivant :

Proposition 4.2.3 *Soit N un processus de Cox sur un espace mesuré (X, \mathcal{B}, ν) d'intensité définie par les équations (4.2) et (4.3) et soient B_1, \dots, B_r , r éléments disjoints de \mathcal{B} . Considérons G_0 paramétré par b et K_σ paramétré par σ . Sous des a priori indépendantes pour α, b, μ et σ , la distribution a posteriori de $(\alpha, b, \mu, \sigma, (a_j)_{j=1, \dots, M}, M)$, conditionnellement aux comptages dans B_1, \dots, B_r , est proportionnelle à :*

$$p(\alpha)p(b)p(\sigma)p(\mu)p(a_1, \dots, a_M \mid \alpha, b, M) \frac{(\mu\nu(X))^M}{M!} e^{-\mu\nu(X)} \prod_{i=1}^r \Lambda(B_i)^{N(B_i)} e^{-\Lambda(B_i)} \quad (4.5)$$

avec la distribution conjointe conditionnelle des a_j obtenue à partir de l'expression suivante :

$$a_j \mid a_1, \dots, a_{j-1} \sim \frac{\alpha}{\alpha + j - 1} G_0 + \frac{1}{\alpha + j - 1} \sum_{k=1}^{j-1} \delta_{a_k}. \quad (4.6)$$

Les $\Lambda(B_i)$ dépendent de $(\sigma, (a_j)_{j=1, \dots, M}, M)$ comme décrit dans l'équation (4.4).

Preuve 4.2.3 Conditionnellement au processus λ , les $N(B_i)$ sont des comptages indépendants suivant respectivement une distribution de Poisson de paramètre $\Lambda(B_i)$. La vraisemblance conditionnelle associée est alors multipliée par la distribution jointe a priori des paramètres $\alpha, b, \mu, \sigma, a_1, \dots, a_M$ et M . Le théorème de Bayes est alors appliqué.

L'équation (4.5) est obtenue à partir de la représentation de la loi prédictive présentée par Blackwell et Macqueen ([8], 1973).

Dans ce qui suit, nous supposons que K_σ est un noyau gaussien isotropique de bande passante σ . Un autre aspect du modèle abordé concerne les limites quand le paramètre d'influence spatiale σ tend vers zéro ou quand le paramètre de concentration α tend vers zéro ou l'infini.

Proposition 4.2.4 *Sous les mêmes conditions que celles de la proposition (4.2.1), considérons B et B' deux éléments de \mathcal{B} et soit K_σ un noyau gaussien isotropique de bande passante σ . Alors,*

$$1) \lim_{\sigma \rightarrow 0} E(N(B)) = \mu E(a_1) \nu(B),$$

$$2) \lim_{\sigma \rightarrow 0} \frac{V(N(B))}{E(N(B))} = 1 - \frac{\mu E(a_1)}{\alpha + 1} \nu(B) + \frac{E(a_1^2)}{E(a_1)},$$

$$3) \lim_{\sigma \rightarrow 0} Cov(N(B), N(B')) = \mu^2 (E(a_1))^2 \frac{\alpha}{\alpha + 1} \nu(B) \nu(B'),$$

4) *Le rapport de la variance par rapport à l'espérance $\frac{V(N(B))}{E(N(B))}$ est une fonction croissante de α .*

Preuve 4.2.4

1) $\Lambda(B) = \sum_{j=1}^M a_j \int_B K_\sigma(x, y_j) \nu(dx)$ et pour tout $(x, j) \in X \times \{1, \dots, M\}$, nous avons $K_\sigma(x, y_j)$ qui converge vers $\delta_{y_j}(\{x\})$ quand σ converge vers 0. Ce qui nous donne :

$$\Lambda(B) \xrightarrow[\sigma \rightarrow 0]{\mathcal{L}} \sum_{j=1}^M a_j \delta_{y_j}(B) \quad (4.7)$$

$E(N(B)) = E(\Lambda(B)) = E(ME(a_1|M)E(\delta_{y_j}(B)|M))$. L'indépendance des a_j et y_j conditionnellement à M mènent à $\lim_{\sigma \rightarrow 0} E(N(B)) = \mu E(a_1) \nu(B)$.

- 2) $\int_B K_\sigma(x, y)\nu(dx)$ converge vers $\mathbf{1}_B(y)$ quand σ tend vers 0. Cependant
- $$\lim_{\sigma \rightarrow 0} \int_X \left(\int_B K_\sigma(x, y)\nu(dx) \right)^2 \nu(dy) = \nu(B),$$
- qui nous mène au résultat final.
- 3) À partir de l'équation (4.7), nous obtenons $E(\Lambda(B), \Lambda(B')) = \mu^2 \nu(B)\nu(B')E(a_1 a_2)$.

Par ailleurs,

$$\text{Cov}(N(B), N(B')) = \text{Cov}(\Lambda(B), \Lambda(B')).$$

$$\text{Donc, } \text{Cov}(N(B), N(B')) = \mu^2 \nu(B)\nu(B')\text{Cov}(a_1, a_2).$$

- 4) À partir de la proposition 4.2.1, nous obtenons

$$\frac{V(N(B))}{E(N(B))} = 1 - \frac{1}{\alpha + 1}E(N(B)) + \frac{E(a_1^2) \int_X \left(\int_B K_\sigma(x, y)\nu(dx) \right)^2 \nu(dy)}{E(a_1) \int_X \int_B K_\sigma(x, y)\nu(dx)\nu(dy)}.$$

De plus, dans l'équation ci-dessus, le second terme du membre de droite et $E(N(B))$ ne dépendent pas de α . Par conséquent, le rapport de la variance par rapport à l'espérance est une fonction croissante de α .

La proposition 4.2.4 montre que même si l'influence spatiale des événements cachés est négligeable (σ proche de zéro), la corrélation des comptages peut être élevée selon la valeur prise par le paramètre α . Soulignons que quand α vaut zéro, les a_j sont tous égaux tandis que quand α tend vers l'infini, les a_j sont indépendants identiquement distribués selon G_0 . Ainsi, la propriété 4 de la proposition 4.2.4 indique que α est un paramètre de dispersion. Plus α est faible, plus la corrélation entre les a_j est élevée et la surdispersion est faible.

Remarque 4.2.1 *Sous les mêmes conditions que celles de la proposition (4.2.1), si K_σ est un noyau gaussien isotropique de bande passante σ alors*

$$\forall x \in X, \lambda(x) \leq \frac{M}{2\pi\sigma^2} \max_{j \in [1, M]} (a_j). \quad (4.8)$$

Il est aisé de le prouver en prenant en compte l'expression (4.2) et le fait que $\frac{1}{2\pi\sigma^2}$ est une borne supérieure de K_σ .

Remarque 4.2.2 *Le rapport de vraisemblance entre le processus de comptage N et un processus homogène d'intensité 1 est donné par l'expression suivante*

$$\frac{e^{-\Lambda(X)} \prod_{i=1}^r \Lambda(B_i)^{N(B_i)} / N(B_i)!}{e^{-\nu(X)} \prod_{i=1}^r \nu(B_i)^{N(B_i)} / N(B_i)!}$$

Ainsi, le log du rapport ci-dessus est obtenu :

$$-\Lambda(X) + \nu(X) + \sum_{j=1}^r N(B_j) \log \frac{\Lambda(B_j)}{\nu(B_j)}.$$

Dans la suite, nous notons $\text{Gamma}(a, b)$ la distribution Gamma de paramètre d'échelle a et de paramètre de forme b .

4.2.2 Cas particuliers

Notons $\mathcal{M}_{\alpha, b, \mu, \sigma}$ le modèle du cas général défini par les expressions 4.2 et 4.3. Nous considérons également quatre cas particuliers, notés $\mathcal{M}_{0, b, \mu, \sigma}$, $\mathcal{M}_{\infty, b, \mu, \sigma}$, $\mathcal{M}_{\alpha, b, \mu, 0}$ et $\mathcal{M}_{0, b, \mu, 0}$, définis respectivement par :

$$\forall x \in X, \quad \lambda(x) = a_1 \sum_{j=1}^M K_{\sigma}(x, y_j) \quad (4.9)$$

avec $a_1 \sim G_0$ et $y_j \sim \text{HPP}(\mu)$

$$\forall x \in X, \quad \lambda(x) = \sum_{j=1}^M a_j K_{\sigma}(x, y_j) \quad (4.10)$$

avec $a_j \stackrel{i.i.d}{\sim} G_0$ et $y_j \sim \text{HPP}(\mu)$

$$\forall x \in X, \quad \lambda(x) = \sum_{j=1}^M a_j \delta_{y_j}(x) \quad (4.11)$$

avec $a_j \sim G_0$ et $y_j \sim \text{HPP}(\mu)$

$$\forall x \in X, \quad \lambda(x) = a_1 \sum_{j=1}^M \delta_{y_j}(x) \quad (4.12)$$

avec $a_1 \sim G_0$ et $y_j \sim \text{HPP}(\mu)$. Ce cas définit un processus de Poisson mixte de distribution mélangeante G_0 et de processus de Poisson sur un support aléatoire discret $\{y_j; j = 1, \dots, M\}$.

Proposition 4.2.5 Soit N un processus de Cox sur un espace mesuré (X, \mathcal{B}, ν) d'intensité définie par les équations (4.2) et (4.3) et soient B_1, \dots, B_r , r éléments disjoints de \mathcal{B} . Considérons G_0 paramétré par b et K_σ paramétré par σ . Sous des lois a priori indépendantes pour α, b, μ, σ , la loi a posteriori pour $(\alpha, b, \mu, \sigma, (a_j)_{j=1, \dots, M}, M)$ conditionnellement à $(N(B_1), \dots, N(B_r))$ est proportionnelle à :

(i) dans le cas du modèle $\mathcal{M}_{0,b,\mu,\sigma}$:

$$f(b, \mu, \sigma, \mu) p(a_1|b) p(y_1, \dots, y_M|M) \prod_{i=1}^r p(N(B_i)|\sigma, a_1, y_1, \dots, y_M, M).$$

(ii) dans le cas du modèle $\mathcal{M}_{\infty,b,\mu,\sigma}$:

$$f(b, \mu, \sigma, \mu) \prod_{j=1}^M p(a_j|\alpha, b, M) p(y_1, \dots, y_M|M) \prod_{i=1}^r p(N(B_i)|\sigma, a_1, \dots, a_M, y_1, \dots, y_M, M)$$

où $f(b, \mu, \sigma, \mu) = p(b)p(\mu)p(\sigma)p(M|\mu)$.

Preuve 4.2.5 En considérant l'expression (4.5) du modèle $\mathcal{M}_{\alpha,b,\mu,\sigma}$, on parvient aux résultats.

4.3 Applications

4.3.1 Analyse des données de comptages *Ibicella lutea*

Nous considérons les données fournies par Melville et Welsh ([57], 2001) constituées de 742 données de comptages d'*Ibicella lutea* (fig. 4.1) sur une parcelle de ferme australienne de surface 1200×1200 mètre carré. La dispersion de cette plante s'effectue par la chute de graines issues de fruits pourris. De plus, les fruits sont également portés par des animaux. Ce dernier événement caché peut être mathématiquement considéré comme la réalisation d'un processus caché. Les figures 4.2, 4.3 et 4.4 permettent de visualiser les données de comptages d'*Ibicella lutea* par taille de grille considérée (8×8 , 16×16 et 32×32). Dans le but de mesurer la surdispersion et l'autocorrélation spatiale à différentes échelles, les comptages ont été regroupés avec différentes tailles de grille, les indices de dispersion et d'autocorrélation de Moran ont été calculés et testés pour chaque taille de grille (table 4.1).

Une surdispersion significative a été trouvée pour chacune des échelles spatiales

FIGURE 4.1 – Plante *Ibicella Lutea*

considérées. L'autocorrélation est significativement positive pour les grilles de taille $2^k \times 2^k, k \geq 2$ (table 4.1). Ceci nous mène à appliquer le modèle du processus de Cox défini à partir de 4.2 et 4.5. Ainsi μ désigne le nombre espéré des événements de dispersion tandis que σ est un paramètre de dispersion de semence : plus σ est proche de zéro, plus la distance de dispersion est courte. Nous supposons que les contributions a_j à l'intensité suit un processus de Dirichlet centré sur une distribution Gamma de paramètre d'échelle égal à l'unité et de paramètre de forme b , noté $Gamma(1, b)$. En fait, la distribution Gamma est fréquemment utilisée pour la modélisation des effets environnementaux (Vaillant [91], 1991). Dans notre cas, le paramètre b est la contribution espérée à l'intensité pour un événement de dispersion, $Gamma(1, b)$ est la distribution marginale pour tout a_j . Quand le paramètre de concentration α tend vers zéro, nous avons des corrélations positives fortes entre les contributions a_j . Quand α tend vers l'infini, ces contributions sont indépendantes.

Les résultats d'inférence bayésienne effectuée pour les trois grilles les plus fines sont présentés dans la table 4.2. Nous avons des estimations similaires aux trois échelles d'étude. Le nombre moyen d'événements cachés varie entre 26 et 27, celui de la contribution à l'intensité, entre 1 et 2. Les estimations du paramètre d'influence sont entre 0.06 et 0.071 et du paramètre de concentration, entre 17 et 29. Notre procédure MCMC a été appliquée et les figures 4.5, 4.6, 4.5, 3.10, 4.8, 4.9, 4.10, 4.11, 4.12 et 4.13 nous montrent respectivement les résultats des chaînes obtenues et la loi *a posteriori* de $\tilde{\theta}$. Dans le cas des grilles de taille 8×8 et 16×16 , nous constatons une bimodalité pour σ il est donc nécessaire d'augmenter la période de chauffe, initialement de 5000, à 30000 dans le cas de la grille 8×8 et 10000 dans le cas de la grille 16×16 (figures 4.8, 4.9, 4.11 et 4.12). A cet effet, les résultats d'inférence bayésienne tenant compte de ces périodes de chauffe sont visibles également dans la table 4.2.

Taille de grille	Moyenne	Variance	Indice de dispersion	p-value (dispersion)	Indice de Moran	p-value (Moran)
2×2	185,00	1234,00	6,67	0,00017	-0,027	0,47541
4×4	46,25	142,33	3,08	0,00005	0,332	0,03891
8×8	11,56	18,60	1,61	0,00157	0,174	0,04159
16×16	2,89	3,98	1,38	0,00006	0,150	0,00075
32×32	0,72	0,87	1,20	0,00001	0,064	0,00369

TABLE 4.1 – Résultats des tests de dispersion et d'autocorrélation spatiale pour *Ibicella lutea* par rapport à la taille de grille

La déviance bayésienne a été choisie comme critère d'ajustement du modèle (Spiegelhalter et al. [84], 2002). La déviance la plus petite a été obtenue pour la grille de taille 16×16 . Néanmoins, pour chaque paramètre, les intervalles de confiance pour les trois grilles se chevauchent largement. La table 4.3 indique les résultats du calcul de la déviance et montre que l'hypothèse d'égalité des contributions est plus vraisemblable que l'hypothèse d'indépendance.

En plus de l'information sur la structure des plantes *Ibicella Lutea* dans les cultures, ces résultats peuvent donner des informations pertinentes sur des événements cachés dans le cadre de contrôle de gestion des plantes *Ibicella Lutea*. Notre approche donne des estimations des paramètres du processus caché influençant la dispersion des *Ibicella Lutea* et peut être intégré dans des programmes de gestion de ces plantes. Soulignons que cette plante est utilisée dans certains pays comme l'Uruguay à des fins médicinales. Ainsi, García et al. ([35],2010) ont récemment étudié l'activité antibactérienne de certains composants chimiques de cette plante et leurs intérêts pour la lutte contre les maladies infectieuses de la peau et des yeux.

4.3.2 Analyse de données de décès dus à la grippe/ pneumonie et aux maladies chroniques respiratoires

Deux jeux de données, disponibles sur le site www.georgiastats.uga.edu, sont analysés. Ces données sont composées de comptages de cas de décès dans chacun des 159 cantons en Georgia, USA, pour la grippe/ pneumonie (IP) et les maladies respiratoires chroniques (CLRD) en 2007. Pour les deux maladies, les occurrences de décès sont quelquefois la conséquence d'événements qui ne sont pas observés dus à des effets de facteurs environnementaux générant des groupes. Cependant, il est possible de formuler

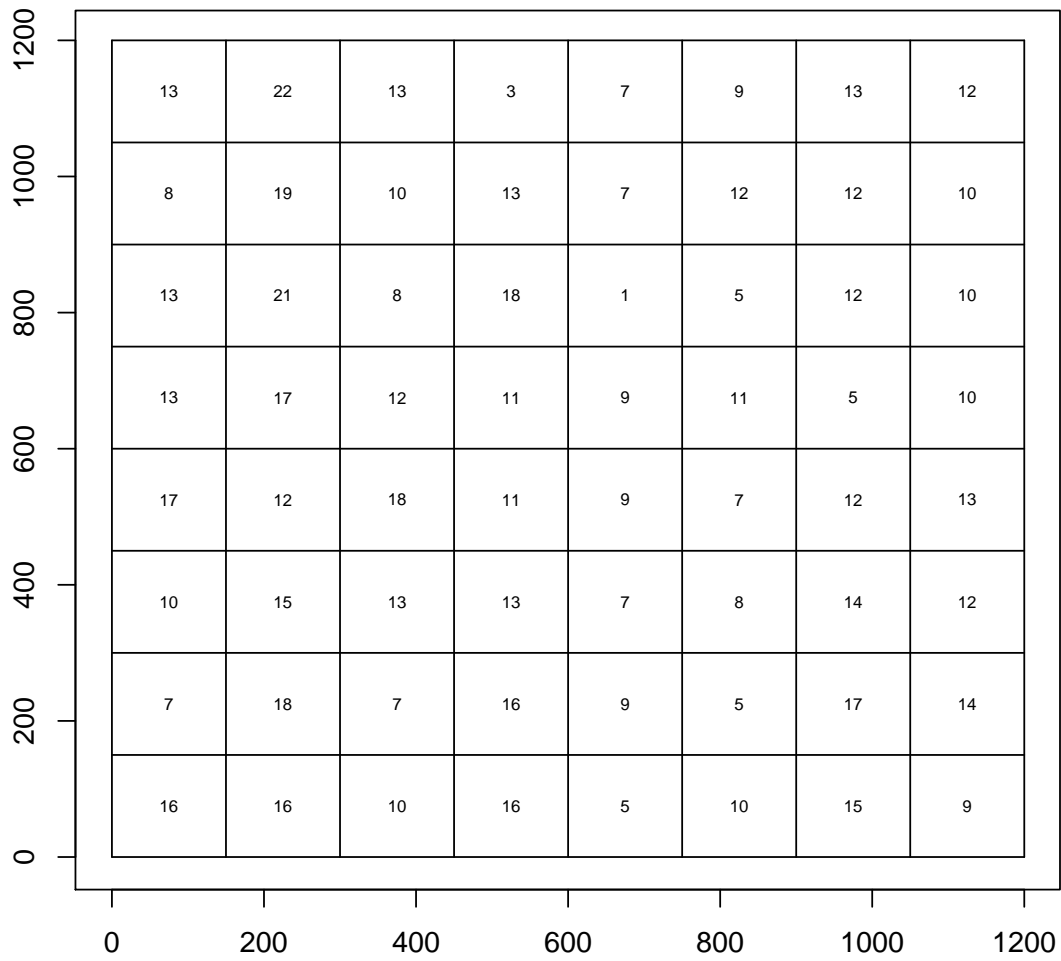


FIGURE 4.2 – Grille régulière 8×8 superposée sur la distribution spatiale de *Ibicella lutea*

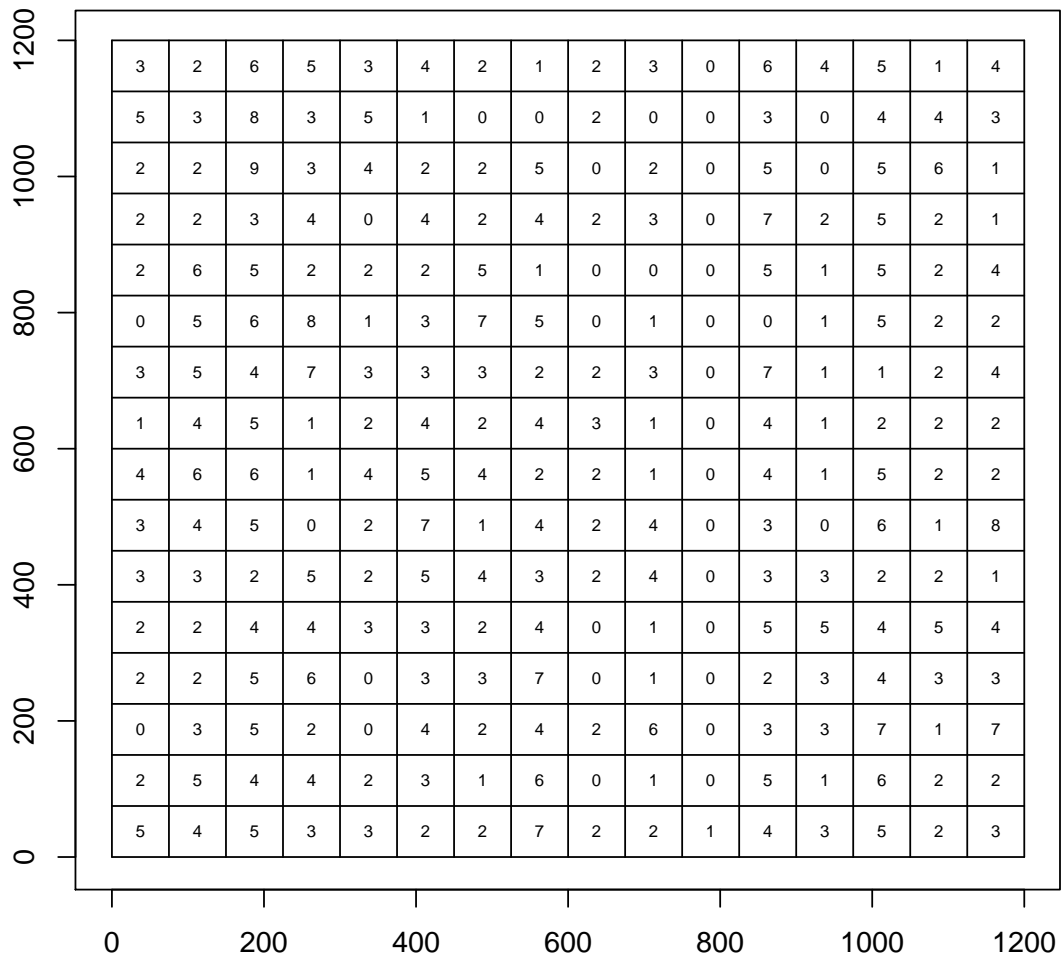


FIGURE 4.3 – Grille régulière 16×16 superposée sur la distribution spatiale de *Ibicella lutea*

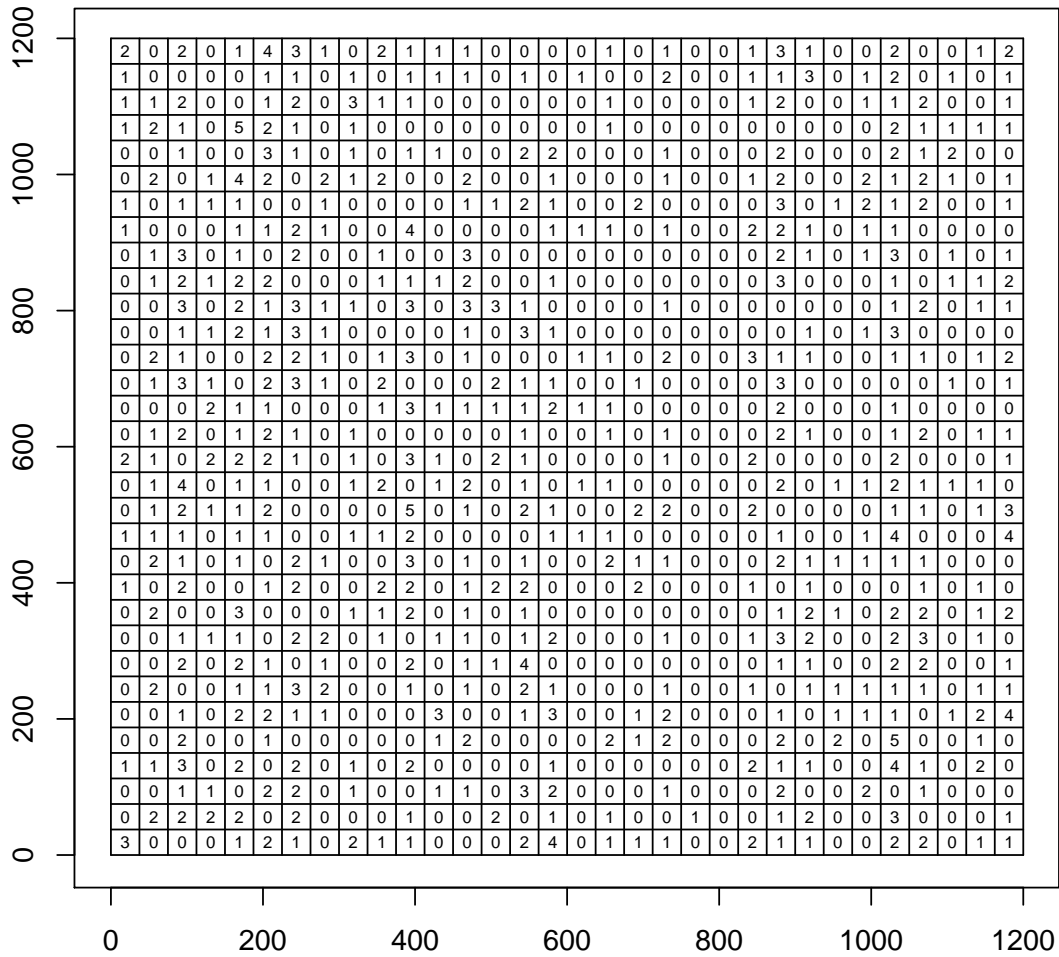


FIGURE 4.4 – Grille régulière 32 × 32 superposée sur la distribution spatiale de *Ibicella lutea*

Grille de taille 8×8 (période de chauffe de 5000)

	Moyenne	Médiane	Écart-type	Quantile 2,5%	Quantile 97,5%
α	26,103	25,472	6,374	15,497	40,046
b	1,717	1,692	0,384	1,081	2,498
μ	26,870	26,585	5,114	17,728	37,796
σ	0,057	0,054	0,009	0,044	0,071

Grille de taille 8×8 (période de chauffe de 30000)

	Moyenne	Médiane	Écart-type	Quantile 2,5%	Quantile 97,5%
α	24,528	23,890	5,794	14,873	37,150
b	2,012	2,004	0,295	1,456	2,609
μ	26,720	26,491	5,117	17,520	37,495
σ	0,067	0,067	0,003	0,062	0,072

Grille de taille 16×16 (période de chauffe de 5000)

	Moyenne	Médiane	Écart-type	Quantile 2,5%	Quantile 97,5%
α	29,644	29,046	6,837	18,156	44,746
b	2,159	2,085	0,445	1,465	3,203
μ	26,816	26,476	5,211	17,601	37,877
σ	0,050	0,048	0,006	0,043	0,066

Grille de taille 16×16 (Période de chauffe de 10000)

	Moyenne	Médiane	Écart-type	Quantile 2,5%	Quantile 97,5%
α	29,433	28,839	6,789	18,085	44,492
b	2,204	2,135	0,446	1,488	3,228
μ	26,769	26,456	5,233	17,455	37,870
σ	0,048	0,048	0,004	0,043	0,057

Grille de taille 32×32 (période de chauffe de 5000)

	Moyenne	Médiane	Écart-type	Quantile 2,5%	Quantile 97,5%
α	30,898	30,247	6,975	19,092	46,110
b	1,720	1,714	0,238	1,272	2,200
μ	26,744	26,425	5,189	17,643	37,955
σ	0,053	0,053	0,002	0,050	0,057

TABLE 4.2 – Résultats de l'inférence bayésienne pour *Ibicella lutea* en fonction de la taille de grille

Grille de taille	Déviante bayésienne ($\alpha = 0$)	Déviante bayésienne ($\alpha = +\infty$)
8×8	2131,955	2230,367
16×16	1383,272	1446,275
32×32	1391,667	1472,676

TABLE 4.3 – Résultats du calcul de la déviante bayésienne pour les données de comptages *Ibicella Lutea* en fonction de la taille de grille.

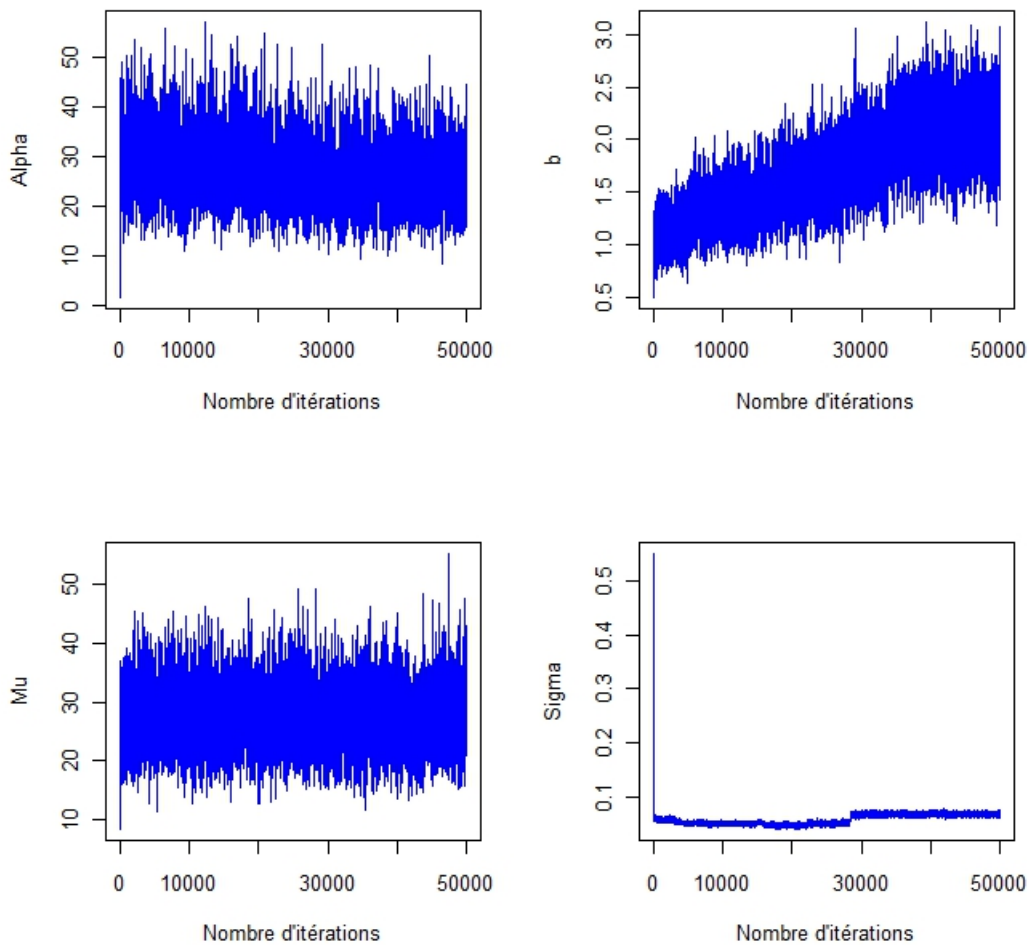


FIGURE 4.5 – Résultats MCMC obtenus par procédure hybride Gibbs-Métropolis-Hastings pour la grille de taille 8×8 de *Ibicella lutea*

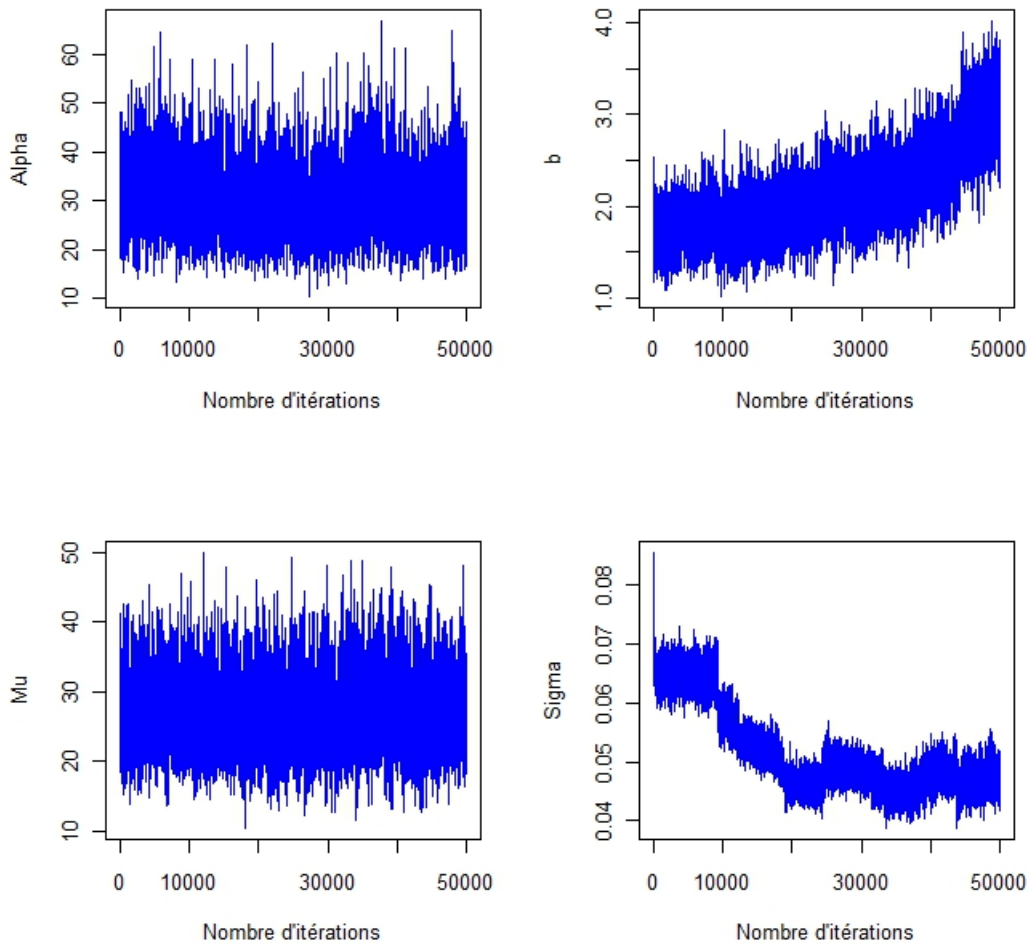


FIGURE 4.6 – Résultats MCMC obtenus par procédure hybride Gibbs-Métropolis-Hastings pour la grille de taille 16×16 de *Ibicella lutea*

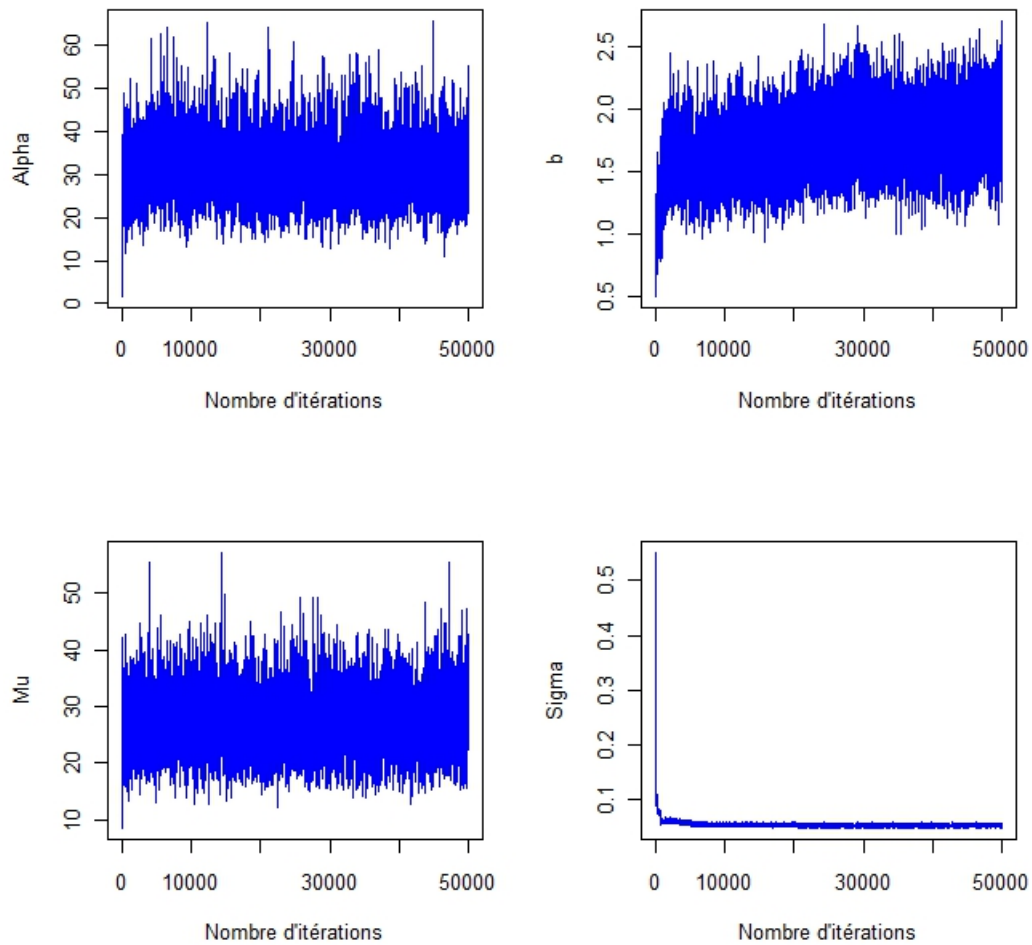


FIGURE 4.7 – Résultats MCMC obtenus par procédure hybride Gibbs-Métropolis-Hastings pour la grille de taille 32×32 de *Ibicella lutea*

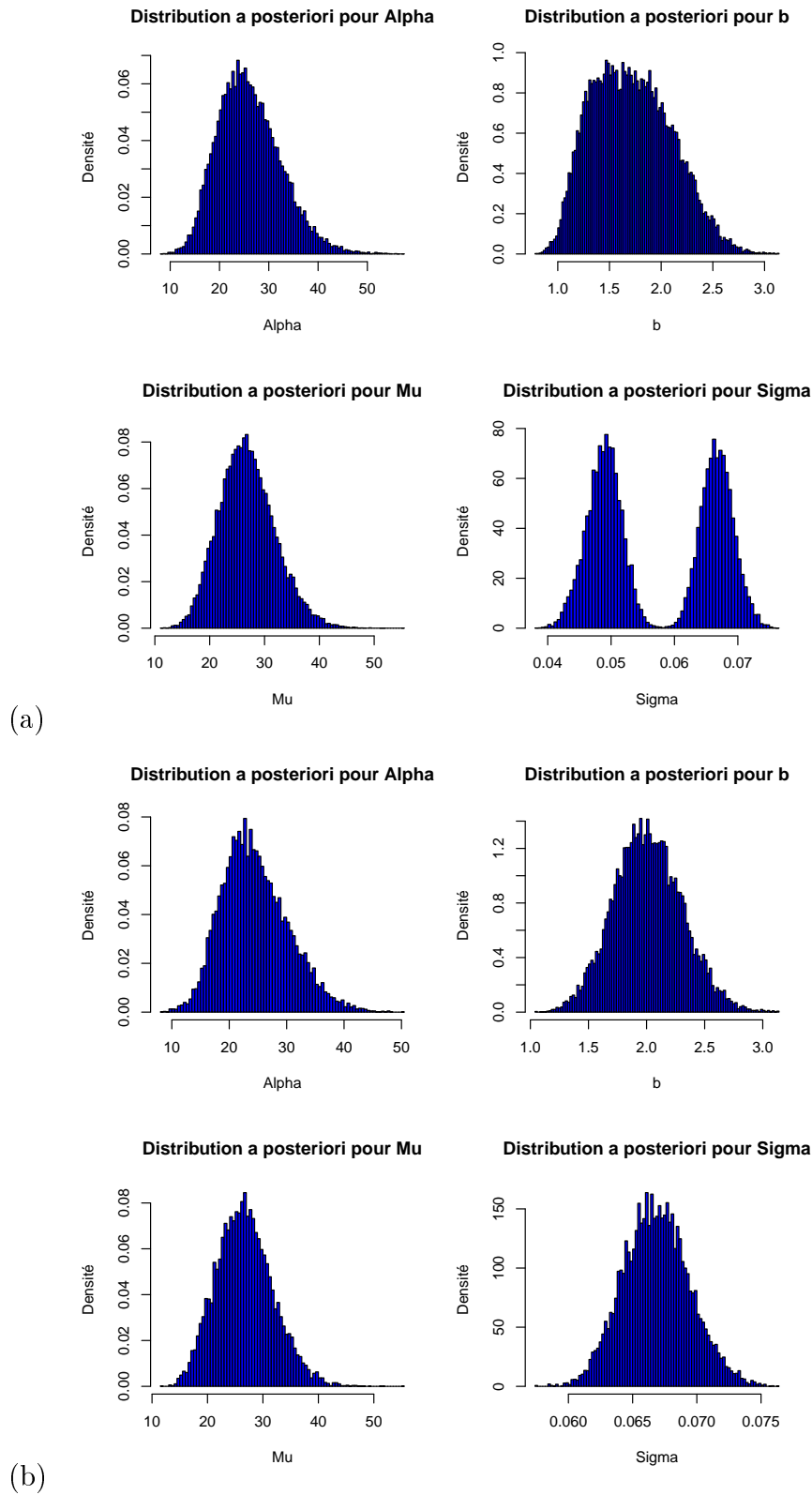


FIGURE 4.8 – Lois *a posteriori* obtenues pour les paramètres pour 50000 itérations et une période de chauffe de, respectivement, 5000 pour (a) et 30000 pour (b) (grille de taille 8×8 et les données de *Ibicella lutea*)

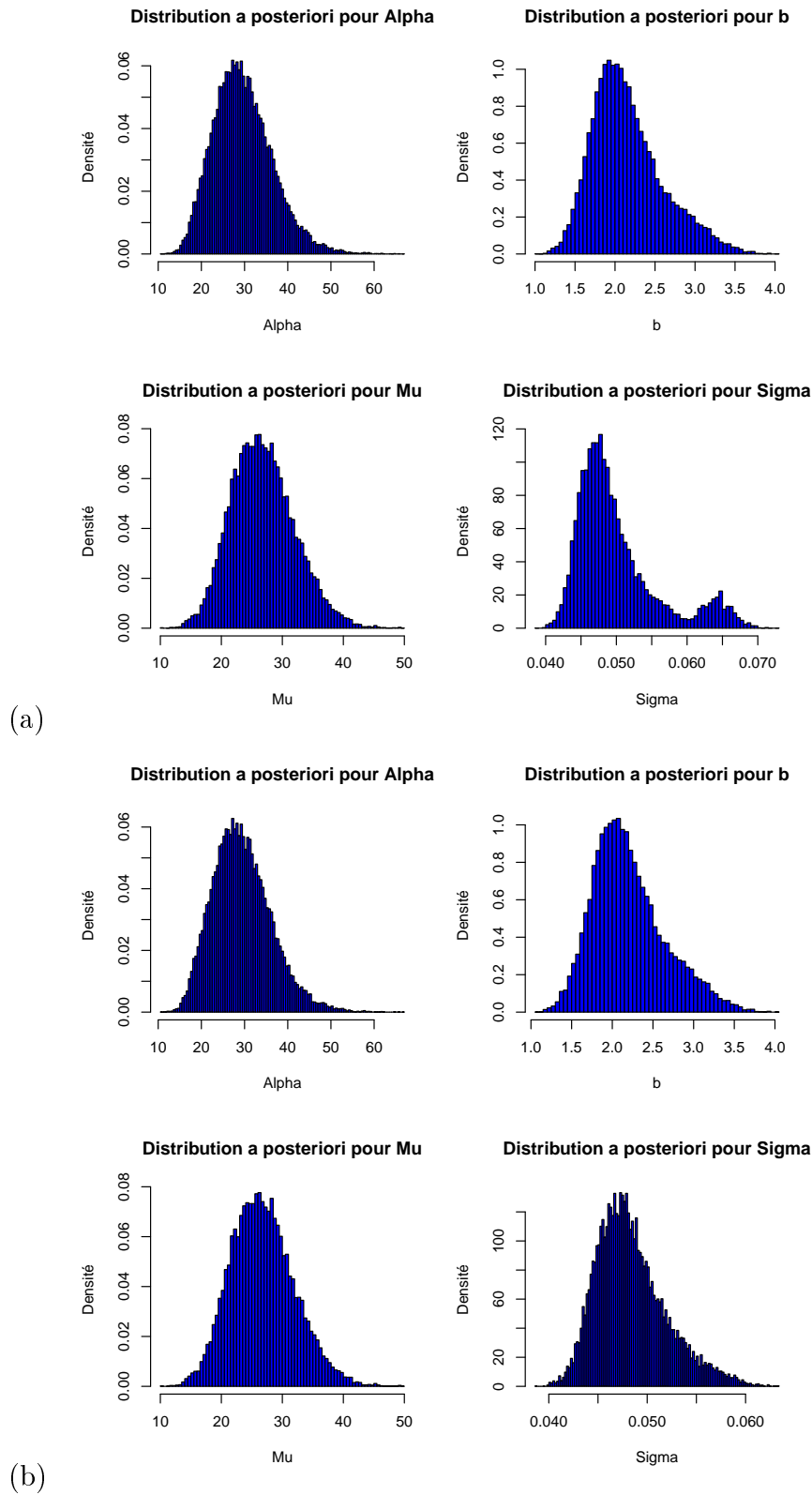


FIGURE 4.9 – Lois *a posteriori* obtenues pour les paramètres pour 50000 itérations et une période de chauffe de, respectivement, 5000 pour (a) et 10000 pour (b) (grille de taille 16×16 et les données de *Ibicella lutea*)

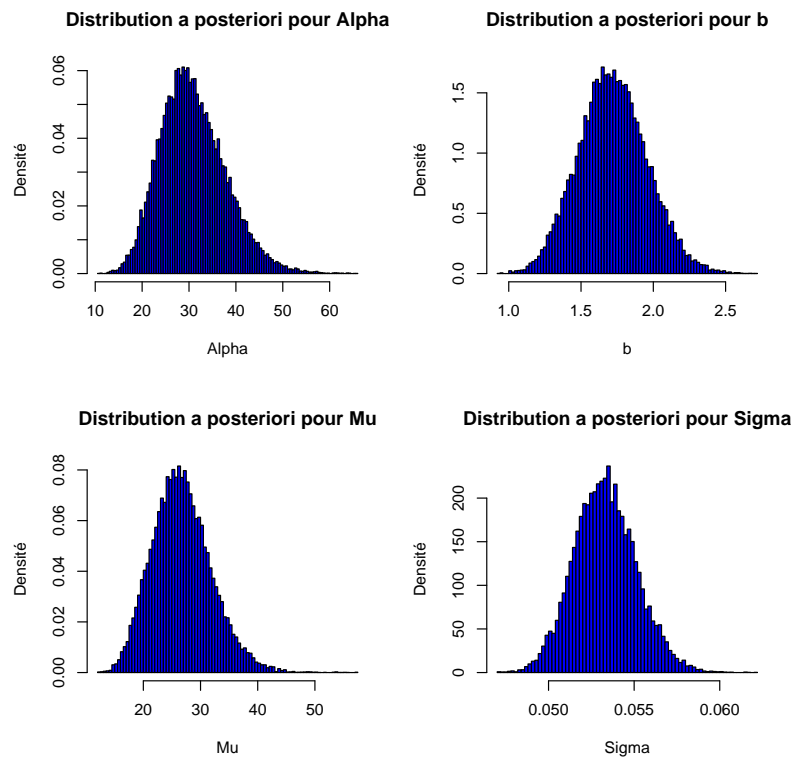
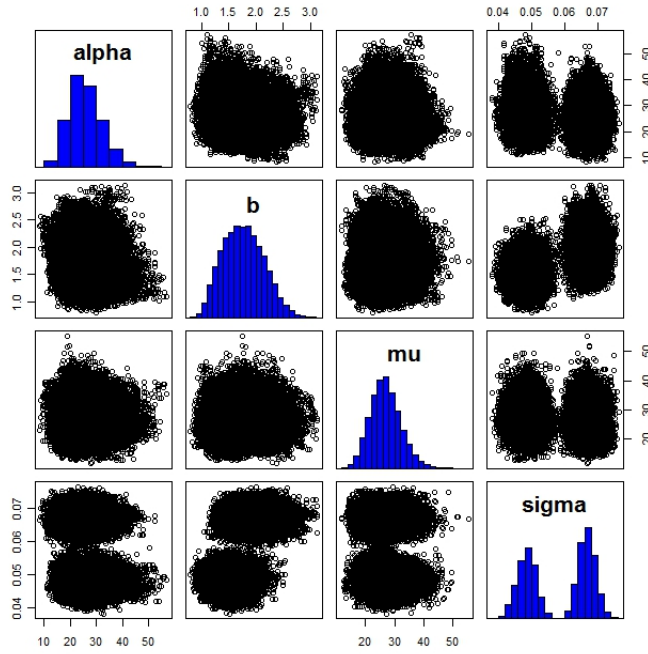
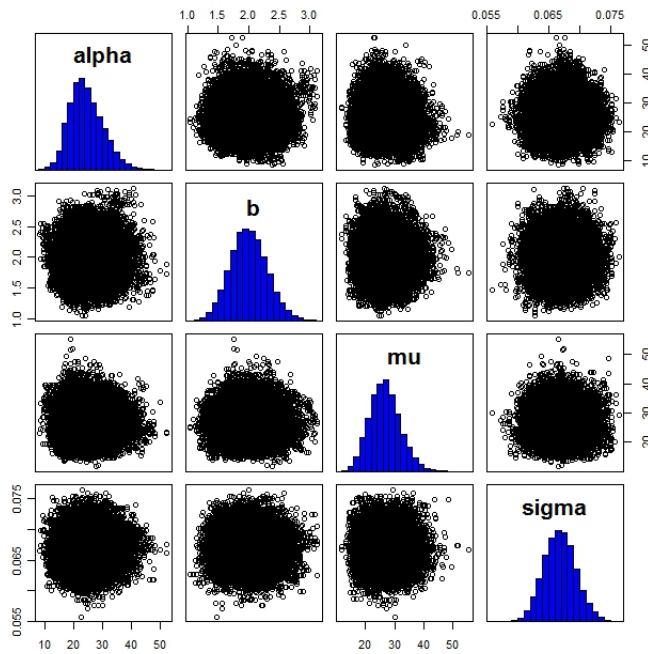


FIGURE 4.10 – Lois *a posteriori* obtenues pour les paramètres pour 50000 itérations et une période de chauffe de 5000 (grille de taille 32×32 et les données de *Ibicella lutea*)

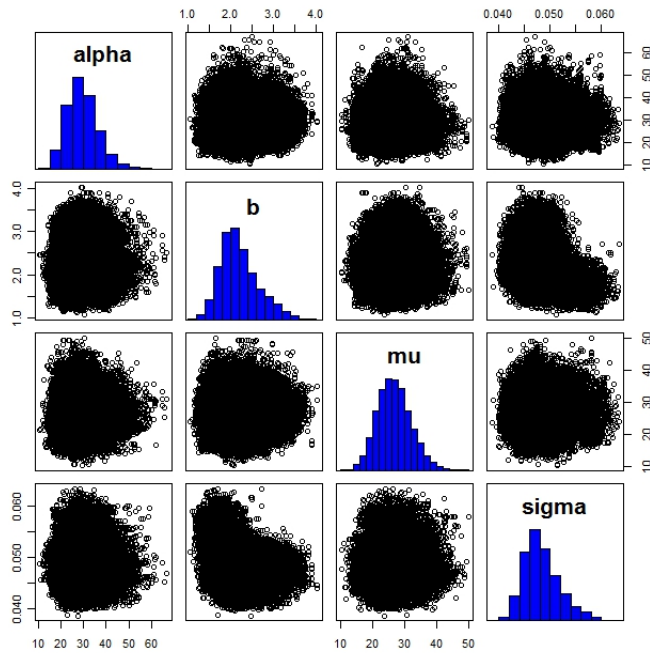


(a)

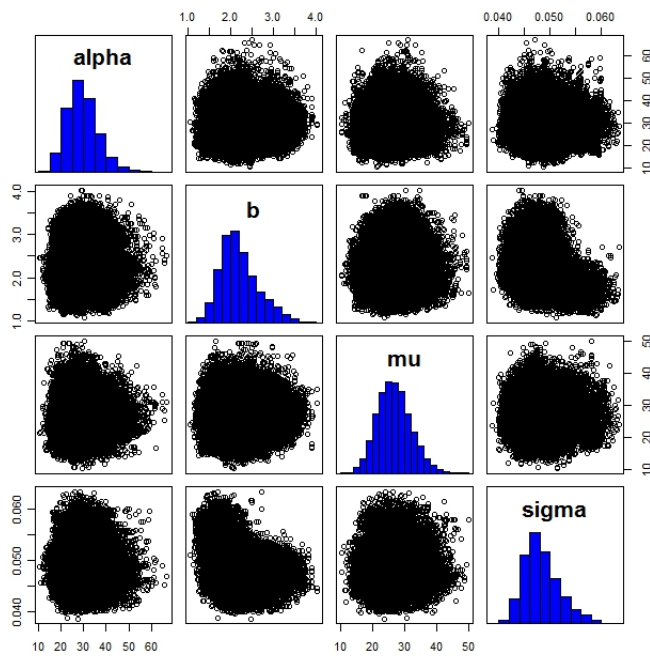


(b)

FIGURE 4.11 – Nuage de points 2 à 2 des paramètres et histogramme sur la diagonale (grille de taille 8×8 , données de *Ibicella lutea*, période de chauffe de, respectivement, 5000 pour (a) et 30000 pour (b))



(a)



(b)

FIGURE 4.12 – Nuage de points 2 à 2 des paramètres et histogramme sur la diagonale (grille de taille 16×16 , données de *Ibicella lutea*, période de chauffe de, respectivement, 5000 pour (a) et 10000 pour (b))

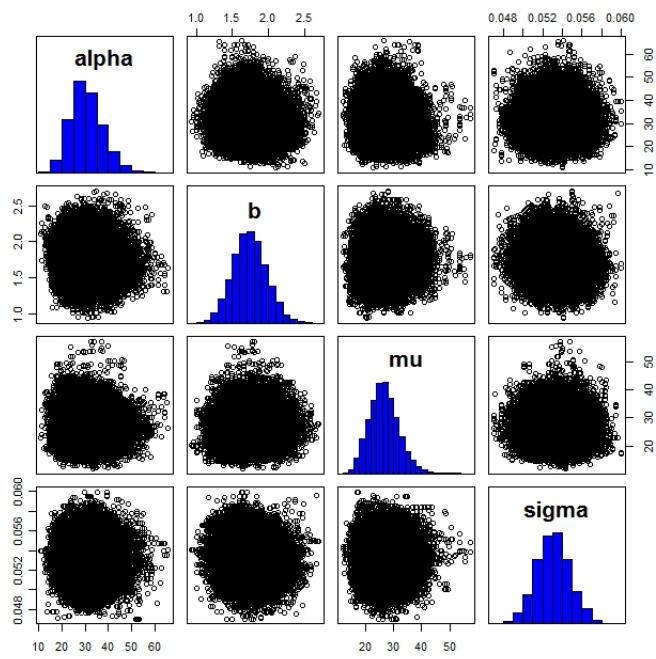


FIGURE 4.13 – Nuage de points 2 à 2 des paramètres et histogramme sur la diagonale (grille de taille 32×32 , données de *Ibicella lutea*)

Cause de décès	Moyenne	Variance	Indice de dispersion	p-value (dispersion)	Indice de Moran	p-value (Moran)
IP	8,76	191,13	21,82	0,00000	0,337	0,00000
CLRD	46,25	789,04	37,36	0,00000	0,405	0,00000

TABLE 4.4 – Résultats des tests de dispersion et d'autocorrélation spatiale pour le nombre de décès par canton en Georgia, USA en 2007 pour IP et CLRD

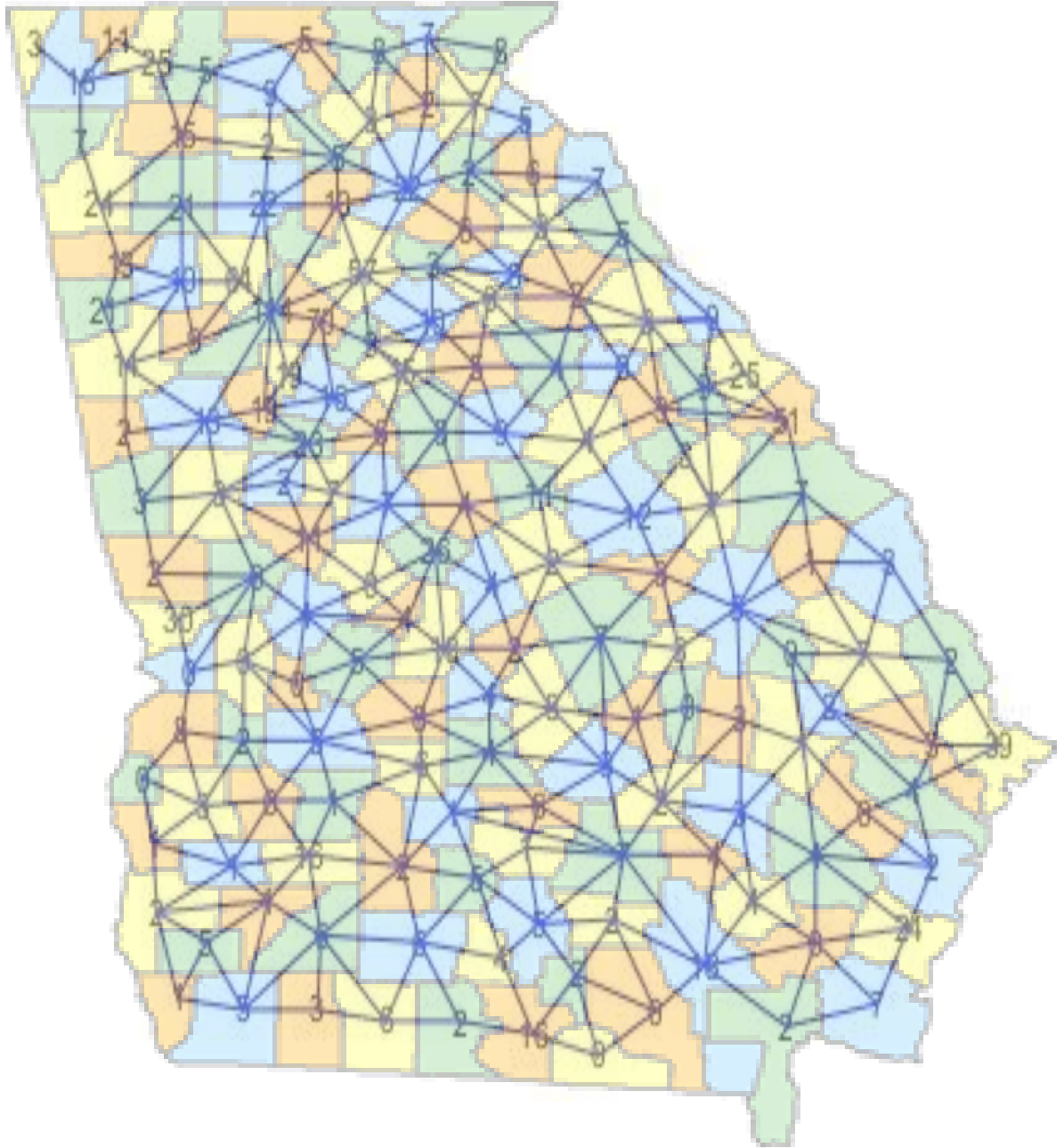
un modèle de groupe bayésien prenant en compte de tels processus non observés pour des centres de de groupes (Lawson [55], 2009).

La figure 4.14 montre les comptages pour chaque canton concernant IP et CLRD ainsi que les liens de voisinages entre cantons. Le poids des voisinages se base sur le fait que deux cantons partagent une frontière commune ou pas. Les indices de dispersion et spatial de Moran basé sur ces voisinages ont été calculés (table 4.4). Les p-values correspondantes indiquent une surdispersion très forte et une autocorrélation spatiale positive pour les deux causes de décès : IP et CLRD.

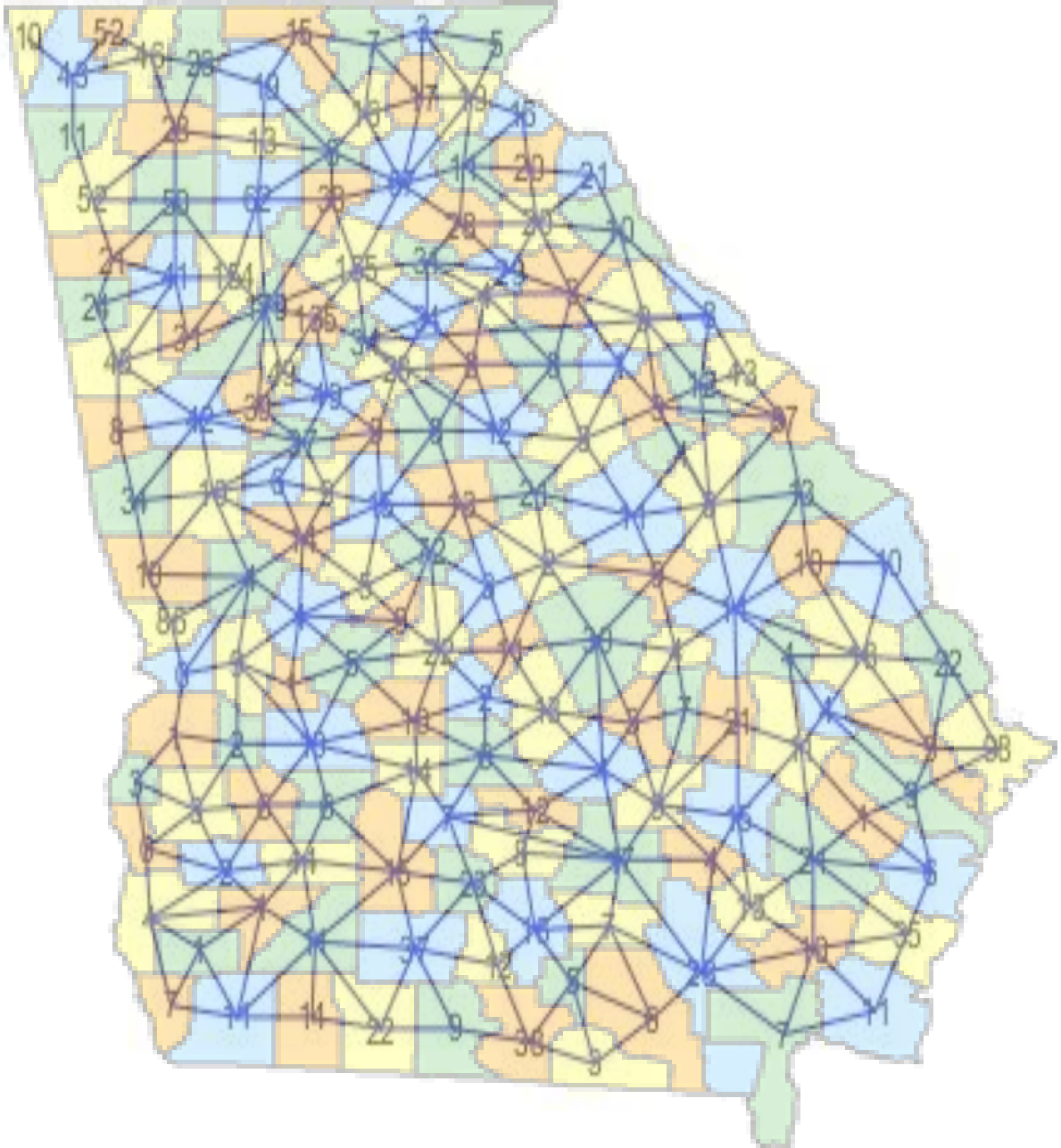
Les figures 4.15, 4.16 et 4.17 nous montrent respectivement les résultats des chaînes obtenues et la loi *a posteriori* de $\tilde{\theta}$. La table 4.5 montre les résultats de l'inférence bayésienne pour les deux maladies. Le nombre moyen d'événements cachés est 26,778 pour la grippe/pneumonie et 26,745 pour les maladies chroniques respiratoires. Ces résultats laissent supposer que les contributions cachées pour les deux causes de décès sont similaires. Les contributions moyennes sont significativement différentes pour les deux maladies, c'est également le cas pour les estimations bayésiennes du paramètre d'influence spatial. Il n'y a pas de différences significatives entre les deux maladies quant aux estimations bayésiennes du paramètre de concentration.

Selon la table 4.6, la déviance bayésienne est plus faible sous l'hypothèse d'indépendance des contributions que sous l'hypothèse d'égalité. En d'autres termes, l'égalité des contributions est moins vraisemblable que leur indépendance.

Ce chapitre a fait l'objet d'une communication en congrès international (Valmy et Vaillant [103], 2012) et un article est en cours d'achèvement (Valmy et Vaillant [101], 2012).



(a)



(b)

FIGURE 4.14 – Lien de voisinage et nombre de décès par canton pour l'état de Georgia, USA en 2007 pour : (a) IP ; (b) CLRD.

	Moyenne	Médiane	Écart-type	Quantile 2,5%	Quantile 97,5%
(a) α	23,285	22,627	6,132	12,991	36,963
b	1,650	1,637	0,264	1,174	2,206
μ	26,778	26,468	5,018	17,788	37,396
σ	0,069	0,069	0,004	0,061	0,077
	Moyenne	Médiane	Écart-type	Quantile 2,5%	Quantile 97,5%
(b) α	21,925	21,315	6,202	11,624	35,841
b	0,783	0,777	0,153	0,500	1,102
μ	26,745	26,416	5,057	17,845	37,540
σ	0,042	0,042	0,001	0,040	0,045

TABLE 4.5 – Résultats de l'inférence bayésienne pour le nombre de décès par canton en Georgia, USA en 2007 for : (a) IP et (b) CLRD.

Cause de décès	Déviance bayésienne ($\alpha = 0$)	Déviance bayésienne ($\alpha = +\infty$)
IP	10694, 56	10688, 000
CLRD	6629, 756	6617, 600

TABLE 4.6 – Résultats du calcul de la déviance bayésienne pour le nombre de décès de Georgia, USA, par canton pour IP et CLRD.

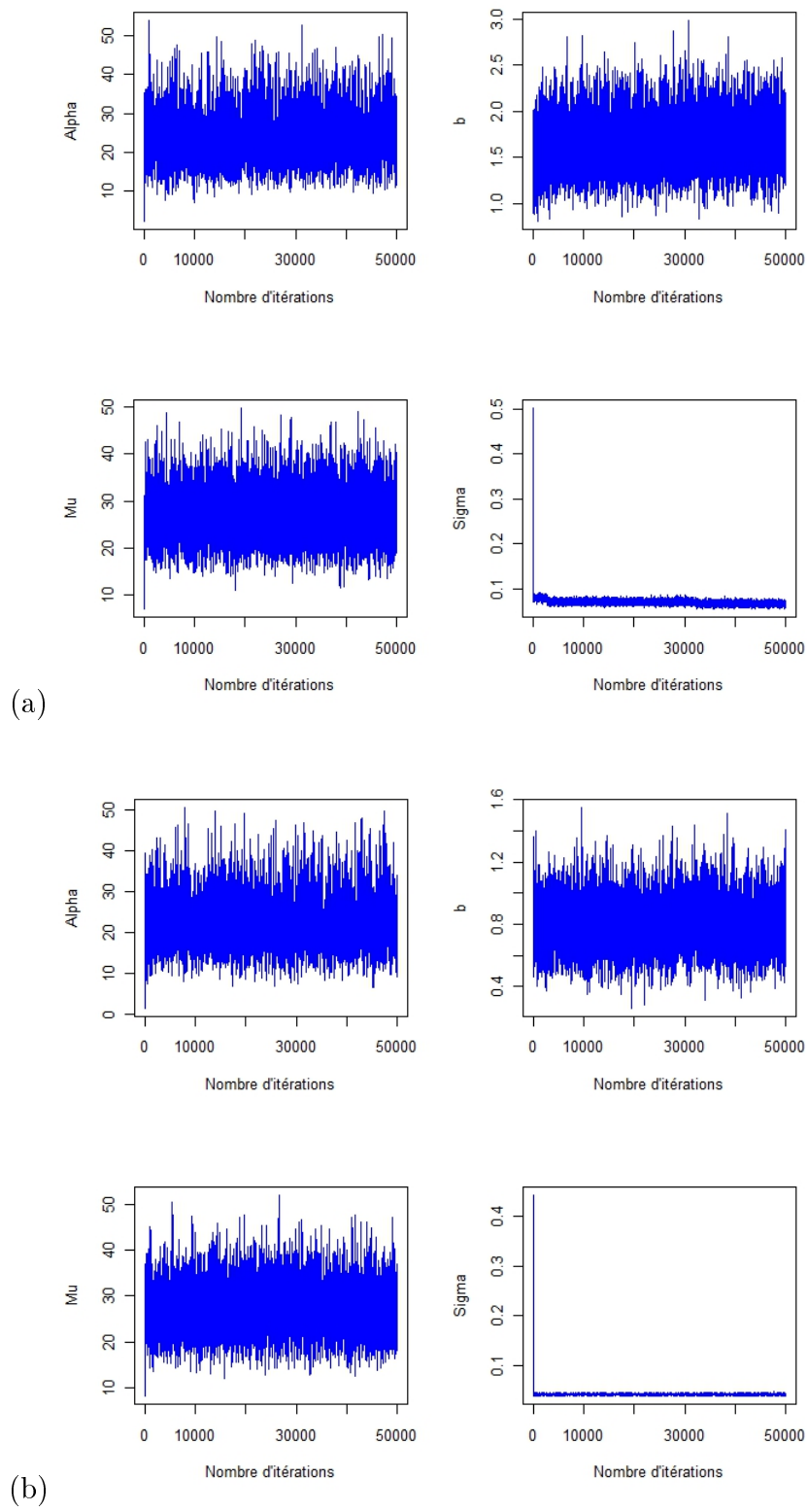


FIGURE 4.15 – Résultats MCMC obtenus par procédure hybride Gibbs-Métropolis-Hastings pour les données de : (a) IP ; (b) CLRD.

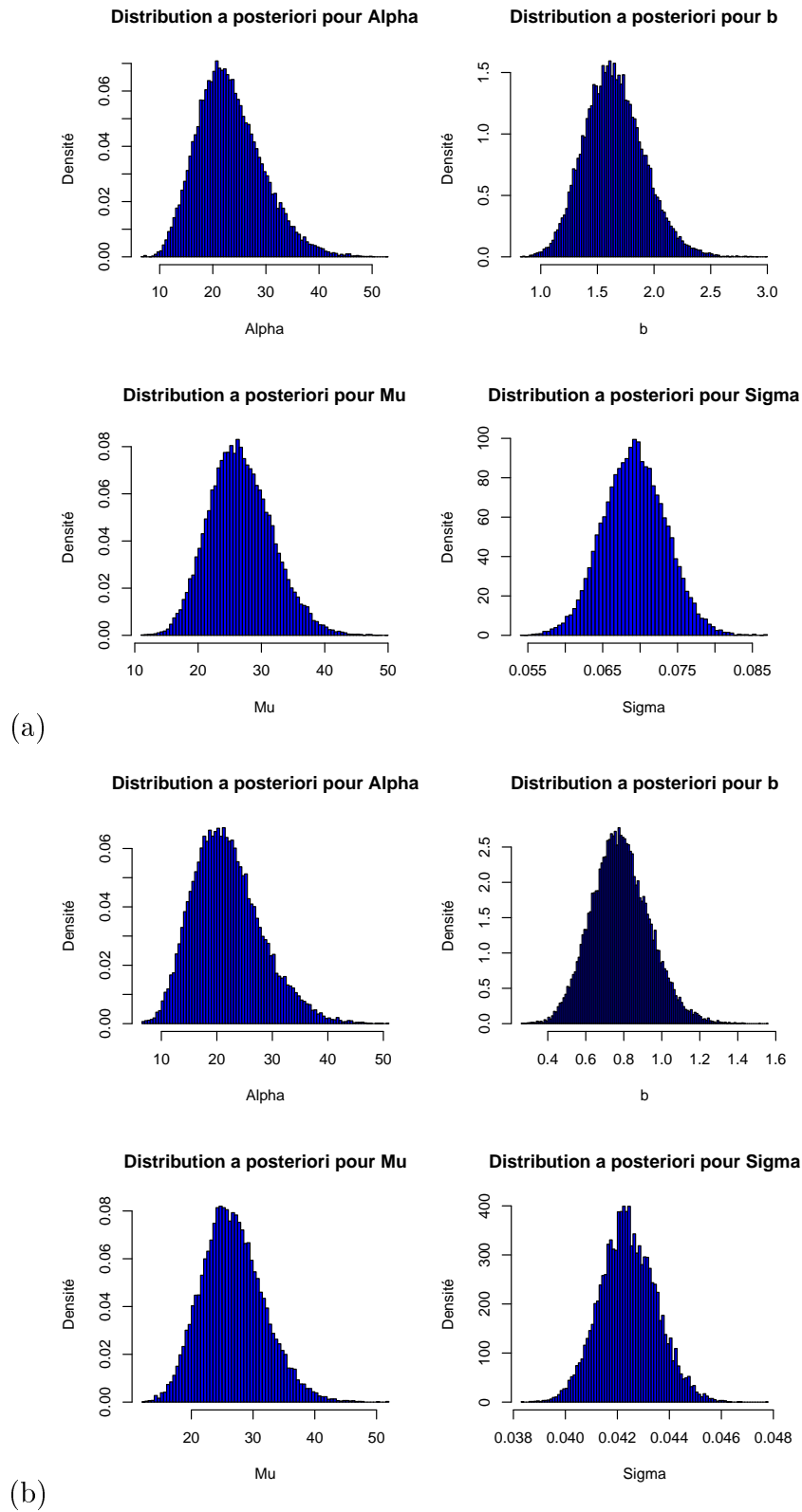
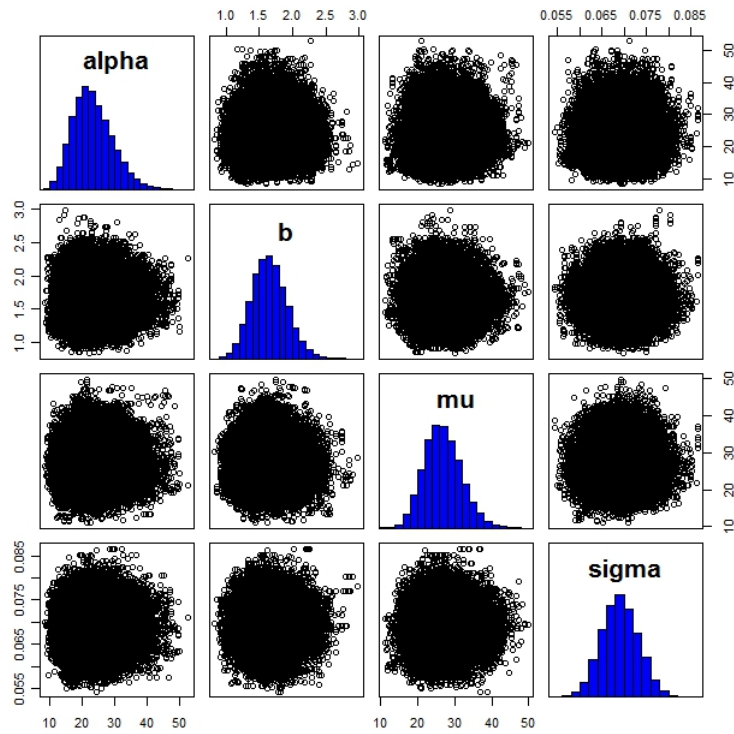
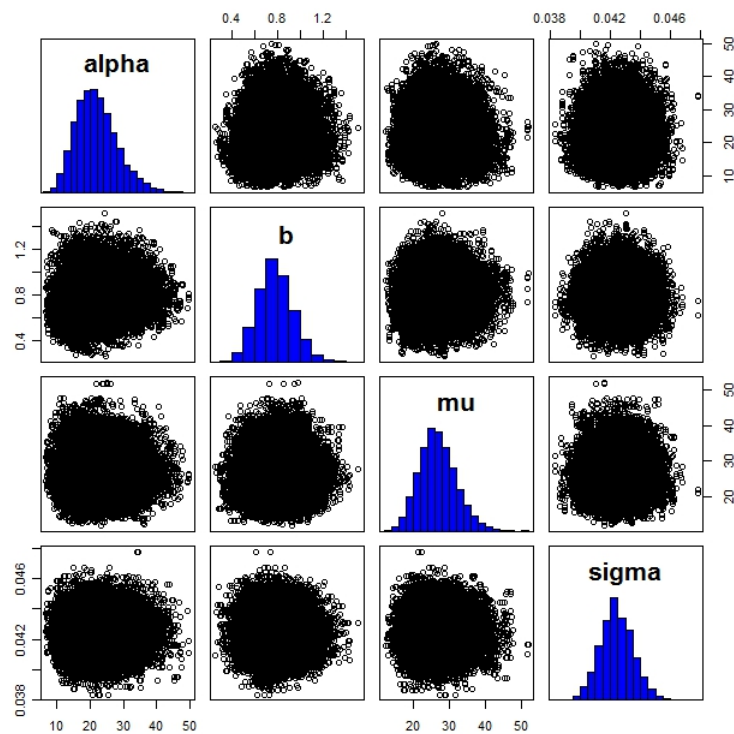


FIGURE 4.16 – Lois *a posteriori* obtenues pour les paramètres pour 50000 itérations et une période de chauffe de 5000 pour : (a) IP ; (b) CLRD.



(a)



(b)

FIGURE 4.17 – Nuage de points 2 à 2 des paramètres et histogramme sur la diagonale pour : (a) IP ; (b) CLRD.

PROCESSUS PONCTUELS EN SISMOLOGIE

Les notations que nous utiliserons dans l'ensemble de ce chapitre sont les suivantes :

- M : magnitude d'un séisme ;
- $H_t = \{t_i, x_i, y_i, M_i; t_i < t\}$: histoire des dates d'occurrence t_i jusqu'à la date t ;
- $P_{\Delta t, \Delta x, \Delta y}(t, x, y | H_t)$: probabilité connaissant le passé du processus qu'un séisme se produise dans l'intervalle de temps $[t, t + \Delta t]$, dans une région $[x, x + \Delta x] \times [y, y + \Delta y]$;
- S : aire de la région des répercussions d'un séisme ;
- $dN(s, \epsilon, M) = \begin{cases} 1 & \text{si un évènement}(t, x, y, M) \text{ est contenu} \\ & \text{dans un élément infinitésimal } d(s, \epsilon, M), \\ 0 & \text{sinon.} \end{cases}$
- μ : fonction intensité d'arrière-plan ;
- $\kappa(M)$: nombre attendu d'évènements déclenchés par un séisme de magnitude M ;
- M_c : seuil de magnitude ;
- $g(\cdot)$: fonction densité de probabilité des dates d'occurrences des répliques ;
- $f(\cdot)$: fonction de répartition des répliques ;
- $\log_{10} N(M \geq M_c) = a - bM_c$ (avec a : indicateur du taux de sismicité et b : rapport entre grands et petits séismes) : loi de Gutenberg-Richter ;
- $\beta = b \log 10$;
- $J(\cdot)$: fonction densité de probabilités des magnitudes des évènements ayant $M_c \leq M$ telle que $J(M) = \beta e^{-\beta(M-M_c)}$;
- $B_{(x,r)}$: boule de centre x et de rayon r .

5.1 Introduction

Le calcul de risque sismique dans une région s'appuie sur des modèles stochastiques décrivant au mieux l'activité sismique. Il s'agit d'étudier les probabilités d'occurrences de tremblements de terre, et d'analyser l'évolution spatio-temporelle de ces probabilités. En statistique sismologique, l'outil principal de modélisation est la théorie des processus ponctuels qui permet, pour une date donnée, d'intégrer l'information jusqu'à cette date (exclue) dans l'expression du risque conditionnel. Ainsi, des outils probabilistes et statistiques ont été développés par des équipes de recherche localisées en des lieux de haute sismicité, principalement en Californie (Turcotte), Japon (Ogata) et Nouvelle-Zélande (Vere-Jones). Les modèles de processus ponctuel permettent d'étudier les distributions spatiales et temporelles des événements sismiques (Vere-Jone [107], 1995; Ogata [65], 1988). Les techniques d'inférence statistique sont basées sur la vraisemblance des observations. Afin d'analyser des données de plus en plus nombreuses, des méthodes statistiques appropriées sont requises. Les méthodes abordées dans ce document sont celles basées sur la vraisemblance d'un processus ponctuel (paragraphe 5.2.3). Cet outil fournit des moyens d'estimation des paramètres du modèle considéré et permet de réaliser des tests d'ajustement. Cette approche possède des propriétés optimales en analyse de données, en particulier, pour les observations de processus ponctuels telles que les événements sismiques.

Dans ce qui suit, nous présenterons quelques caractéristiques générales d'un processus ponctuel dans le cadre sismique, quelques propriétés de la vraisemblance, ainsi que des tests d'ajustements (Schœnberg [82], 2003; Zhuang [112], 2006). Puis nous nous intéresserons plus spécifiquement au modèle ETAS (Epidemic Type Aftershock Sequence) (Ogata [66], 1998; Zhuang et al. [114], 2005) en considérant le modèle d'intensité conditionnelle spatio-temporelle telle que proposée par Ogata ([68], 2011) en y ajoutant une intensité d'arrière-plan, non plus déterministe, mais de loi un processus de Dirichlet centré sur un processus spatial log-gaussien. Dans un dernier temps, nous focaliserons sur la comparaison de différentes fonctions déclenchantes en analysant des données sismiques de l'Arc des Petites Antilles.

5.2 Processus ponctuels en sismologie

Notons $]0, T]$ (T désigne un réel positif) la période d'observation et X la région d'étude. Soit $N(]0, t], B_{(x,r)})$ le nombre d'événements sismiques ayant eu lieu durant la

période $]0, T]$ dans l'unité spatiale $B_{(x,r)}$ inclus dans X . L'intensité conditionnelle pour t appartenant à $]0, T]$ et x dans X est définie comme suit :

$$\lambda(x, t|H_{t-}) = \lim_{\substack{h \rightarrow 0^+ \\ r \rightarrow 0^+}} \frac{1}{h \times r} P(N(]t, t+h], B_{(x,r)}) = 1|H_{t-})$$

où H_{t-} est l'histoire des événements sismiques jusqu'à la date t .

$\lambda(x, t|H_t)$ est aussi appelée intensité stochastique (Andersen et al. [3], 1993 ; Daley et Vere-Jones [22], 2003). $\lambda(x, t|H_t)dt dx$ peut être vue comme la probabilité d'un unique événement sismique durant la période $]t, t+dt]$ et situé à une distance de x inférieure à dx sachant l'histoire jusqu'à la date t . De plus, nous considérons que la probabilité que deux ou plusieurs événements se produisent dans un petit intervalle est négligeable. Sous de faibles conditions, $\lambda(x, t|H_t)$ caractérise complètement la structure stochastique des événements sismiques.

Un catalogue sismique est un ensemble $\{(x_i, t_i, M_i), T_0 < t_i \leq T\}$ constitué de tous les séismes observés au dessus d'une certaine magnitude dans une région spécifiée, et dans un intervalle de temps $]T_0, T]$. Notons que x_i représente dans notre cas le vecteur composé de la longitude, la latitude et la profondeur. Holden et al. ([41], 2003) se sont intéressés à des modèles de processus ponctuels marqués pour lesquels chaque séisme est représenté par sa magnitude, ses coordonnées spatiales et sa date d'occurrence, afin de prendre en compte les répliques.

Un modèle très populaire en sismologie est le modèle ETAS (Epidemic Type After-shock Sequence) dont l'historique est présenté dans les paragraphes 5.2.1 et 5.2.2.

5.2.1 Modèle temporel

Hawkes ([40], 1971) introduit un processus auto-excitatif N d'intensité conditionnelle :

$$\lambda(t) = \mu + \int_{-\infty}^t g(t-s)dN(s) \quad (5.1)$$

où :

- μ est l'intensité d'arrière plan correspondant, par exemple en sismologie, aux secousses principales
- $\int_{-\infty}^t g(t-s)dN(s)$ est l'intensité due au passé du processus (les répliques en sismologie).

Il s'agit donc d'un processus purement temporel. Puri et al. ([74], 1986) étudie le fait que, dans la pratique, on n'a pas accès à la totalité de l'histoire du processus et que l'on considère donc cette histoire à partir d'une date initiale choisie égale à zéro. Ils regardent l'existence d'éventuels effets sur l'efficacité des techniques de vraisemblance dus à l'approximation de $\lambda(\cdot)$ par

$$\tilde{\lambda}(t) = \mu + \int_0^t g(t-s)dN(s).$$

telle que

$$0 < \int_0^\infty g(t)dt < 1.$$

5.2.2 Modèle spatio-temporel marqué

Ogata ([65], 1988) a développé le modèle temporel ETAS puis il s'est intéressé avec (Ogata et Zhuang [70], 2006) au spatio-temporel avec marques. Il considère un ensemble de marques \mathcal{M} , un espace $X \subset \mathbb{R}^n$, un réel strictement positif T et pose :

$$\forall(t, x, M) \in [0, T] \times X \times \mathcal{M}$$

$$\lambda^*(t, x, M) = \lambda(t, x|H_{t-}) \times J(M) \tag{5.2}$$

où $J(\cdot)$ est la fonction densité de probabilités des magnitudes des évènements ayant une magnitude M telle que : $M_c \leq M$ et où M_c est un seuil de magnitude. H_{t-} est l'histoire du processus N jusqu'à la date t exclue.

$$\begin{aligned} \lambda(t, x|H_{t-}) &= \mu(x) + \int_0^t \int_X \int_{\mathcal{M}} \kappa(M) \times g(t-s) \times f(x-\epsilon|M) \times dN(s, \epsilon, M) \\ &= \mu(x) + \sum_{t_i < t} \kappa(M_i)g(t-t_i)f(x-x_i|M_i) \end{aligned} \tag{5.3}$$

Dans le membre de droite de (5.3), interviennent

- $\mu(\cdot)$: fonction intensité d'arrière-plan,
- $\kappa(\cdot)$: nombre attendu d'évènements déclenchés par un séisme de magnitude M ,
- $g(\cdot)$: fonction densité de probabilités des dates d'occurrences des répliques,
- $f(\cdot)$: fonction de répartition spatiale conditionnelle des répliques

avec les conditions

$$\int_X \mu(x) dN(x) < +\infty$$

et

$$\int_0^{+\infty} \int_X \int_{\mathcal{M}} \kappa(M) \times g(t-s) \times f(x-\epsilon|M) \times dN(s, \epsilon, M) < 1.$$

La fonction κ est souvent choisie de la forme $Ae^{\alpha(M-M_c)}$, où M_c est une magnitude seuil et A, α désignent des constantes. Le noyau spatial f , lui, est proportionnel à la densité gaussienne (Ogata et Zhuang [70], 2006) :

$$f(x - x_i|M_i) = \exp\left(-\frac{(l - l_i)^2}{d_i^2 \exp(\beta(M_i - M_c))} - \frac{(L - L_i)^2}{e_i^2 \exp(\beta(M_i - M_c))}\right)$$

avec $x = (l, L)$, $x_i = (l_i, L_i)$, $d_i > 0$ et $e_i > 0$.

Remarque 5.2.1 L'intégrale $\int_X \lambda(x, t|H_t) dx$ désigne l'intensité $\lambda(t|H_t)$ à la date t du modèle temporel ETAS pour lequel seul le noyau temporel est pris en compte.

Pour notre part, l'aspect spatial pouvant être occulté, nous considérons l'expression suivante de l'intensité conditionnelle en indexant par θ et en n'explicitant pas le conditionnement par rapport à H_{t-} pour plus de clarté dans les résultats asymptotiques qui suivent :

$$\lambda_\theta(t) = \mu + \int_{-\infty}^t \kappa(M) g((t-s)) dN(s). \quad (5.4)$$

L'approximation de cette intensité liée à la non observation des dates antérieures à 0 est :

$$\lambda_\theta(t) = \mu + \int_0^t \kappa(M) g((t-s)) dN(s). \quad (5.5)$$

5.2.3 Techniques de vraisemblance

La vraisemblance d'une série d'événements sismiques ou d'un modèle statistique est définie par la densité de probabilité jointe des données. Pour un processus ponctuel quelconque, cette densité est obtenue à partir de son intensité conditionnelle (Daley et Vere-Jones [22], 2003). La densité de probabilité dépend d'un paramètre inconnu, noté θ , qu'il faut estimer ; il s'agit donc d'une fonction de θ . Soit $\{(x_i, t_i), i = 1, \dots, n\}$ un ensemble d'événements sismiques sur la période $(0, T]$ et l'espace X . Afin de faciliter l'estimation des paramètres du modèle, il est usuel de considérer la fonction

log-vraisemblance au lieu de la vraisemblance. Dans le cas d'un processus ponctuel, d'intensité notée λ_θ , la log-vraisemblance \mathcal{L} est définie comme suit :

$$\mathcal{L} = \log L(x, t | H_t) = \sum_{t_j \leq T} \log \lambda_\theta(x_j, t_j | H_t) - \int_0^T \int_X \lambda_\theta(x, t | H_t) dx dt. \quad (5.6)$$

En maximisant la log-vraisemblance par rapport à θ , il est possible d'estimer les paramètres du modèle. Cette méthode est appelée estimation du maximum de vraisemblance. Les estimations du maximum de vraisemblance $\hat{\theta}$ de θ sont les valeurs maximisant la log-vraisemblance \mathcal{L} .

Au début des années 70, Vere-Jones ([106], 1970) et Kagan ([47], 1973) ont été parmi les premiers à modéliser l'activité sismique grâce aux processus ponctuels. Ogata ([67], 1999) a présenté quelques modèles utiles dans l'analyse de la sismicité et résumé les méthodes basées sur la vraisemblance pour l'analyse des processus ponctuels appliqués à des données sismiques.

5.3 Étude des convergences

Nous présentons quelques propriétés asymptotiques de l'intensité conditionnelle et de la log-vraisemblance développées par Ogata ([64], 1978) en ce qui concerne la convergence en probabilité, ainsi que celles présentées par Puri et Tuan ([74], 1986) qui eux traitent de convergence presque-sûre.

5.3.1 Propriétés asymptotiques de l'intensité stochastique

Propriété 5.3.1 (Ogata, 1978) *Soit N un processus ponctuel stationnaire ergodique. Si son processus intensité $\lambda_\theta(t)$ est stationnaire et prévisible avec un moment d'ordre 2 fini, alors :*

$$\frac{1}{T} \int_0^T \lambda_\theta(t) dt \xrightarrow[T \rightarrow +\infty]{p} E(\lambda_\theta(0)).$$

Preuve 5.3.1 Pour tout t , $\lambda_\theta(t)$ est une fonction mesurable du processus ponctuel N , alors par stationnarité, λ satisfait le théorème ergodique de Birkhoff-Khinchin.

Propriété 5.3.2 (Ogata, 1978) Soit $\psi_\theta(t)$ une fonction de $\lambda_\theta(t)$, $\lambda_\theta^{(i)}(t)$, $i = 1, 2$, telle que $\psi_\theta(t)$ soit intégrable, alors :

$$\frac{1}{T} \int_0^T \psi_\theta(t) dt \xrightarrow[T \rightarrow +\infty]{p} E(\psi_\theta(0)).$$

Propriété 5.3.3 Considérons un processus ponctuel défini par l'équation (5.4), supposons que la limite de g en l'infini est 0, κ est borné et $N(0) < +\infty$, alors $\lambda_\theta(t) - \tilde{\lambda}_\theta(t) \xrightarrow[t \rightarrow +\infty]{p.s.} 0$.

Preuve 5.3.2

$$\lambda_\theta(t) - \tilde{\lambda}_\theta(t) = \int_{-\infty}^0 \kappa(M)g(t-s)dN(s)$$

g est une densité de probabilité et κ est bornée par B , ainsi

$$\begin{aligned} \forall \epsilon > 0, \exists T^* > 0, \forall t > T^* &\implies g(t-s) < \epsilon \\ &\implies \lambda_\theta(t) - \tilde{\lambda}_\theta(t) < \epsilon \int_{-\infty}^0 \kappa(M)dN(s) \\ &< \epsilon B \int_{-\infty}^0 dN(s). \end{aligned}$$

Ce qui permet d'aboutir au résultat.

Remarque 5.3.1 Ce résultat signifie qu'en dépit du fait que les observations débutent à une date initiale (par convention $t = 0$), nous pouvons utiliser l'approximation $\tilde{\lambda}_\theta$ de λ_θ quand la longueur de la fenêtre d'observation T est assez large. Le processus est considéré comme ayant atteint son état de stabilité.

5.3.2 Propriétés asymptotiques de la log-vraisemblance

Notons $\phi_\theta(t) = \log \lambda_\theta(t)$.

Propriété 5.3.4 (Convergence en presque sûre de $\tilde{L}_T(\theta)$; Puri et Tuan, 1986)
Supposons que λ_θ définie par l'équation 5.1 satisfait les propositions 5.3.1-5.3.2-5.3.3 et λ_θ , ϕ_θ sont des variables bornées, alors elles sont de carrés intégrables. Par conséquent,

nous obtenons presque sûrement la convergence de $\tilde{L}_T(\theta)$:

$$\frac{1}{T} \tilde{L}_T(\theta) \xrightarrow[T \rightarrow +\infty]{p.s.} E[\phi_\theta(0)\lambda(0) - \lambda_\theta(0)] = \Psi(\theta). \quad (5.7)$$

Théorème 5.3.1 (Convergence asymptotique de $\tilde{L}_T(\theta)$; Ogata, 1978)

Supposons que :

1. N est un processus ponctuel stationnaire, ergodique et ordonné ;
2. $\forall \theta \in \Theta, \lambda_\theta$ est prévisible ;
3. $\lambda_\theta(t)$ est continue en θ et $\lambda_\theta(t) > 0$ p.s. pour tout $\theta \in \Theta$;
4. $E\left(\sup_{0 < h \leq 1} \frac{1}{h} N([0, h])^2\right) < +\infty$;
5. $\frac{\partial \phi_\theta}{\partial \theta_i}, \frac{\partial^2 \phi_\theta}{\partial \theta_i \partial \theta_j}, \frac{\partial^3 \phi_\theta}{\partial \theta_i \partial \theta_j \partial \theta_k}$ existent et sont continus en $\theta, \forall t \in \mathbb{R}_+$;
6. $\frac{\partial \lambda_\theta}{\partial \theta_i}, \frac{\partial^2 \lambda_\theta}{\partial \theta_i \partial \theta_j}$ ont un moment fini d'ordre 2 pour tout $\theta \in \Theta$;
7. pour chaque $\theta \in \Theta$, l'information de Fisher $I(\theta) = \{I_{ij}(\theta)\}_{i,j=1,\dots,d}$ et $I_{ij}(\theta) = E\left(\frac{1}{\lambda_\theta} \frac{\partial \lambda_\theta}{\partial \theta_i} \frac{\partial \lambda_\theta}{\partial \theta_j}\right)$ est non-singulière et chaque élément $\frac{1}{\lambda_\theta} \frac{\partial \lambda_\theta}{\partial \theta_i} \frac{\partial \lambda_\theta}{\partial \theta_j}$ a un moment fini d'ordre 2 ;
8. pour tout $\theta \in \Theta$:

$$E\left(\frac{1}{\sqrt{T}} \int_0^T \left|\frac{\partial \lambda_\theta}{\partial \theta_i} - \frac{\partial \tilde{\lambda}_\theta}{\partial \theta_i}\right| dt\right) \xrightarrow[T \rightarrow +\infty]{} 0$$
 et

$$E\left(\frac{1}{\sqrt{T}} \int_0^T \left|\lambda_\theta - \tilde{\lambda}_\theta\right| \frac{1}{\tilde{\lambda}_\theta} \left|\frac{\partial \tilde{\lambda}_\theta}{\partial \theta_i}\right| dt\right) \xrightarrow[T \rightarrow +\infty]{} 0.$$

alors,

$$\frac{1}{\sqrt{T}} \tilde{L}_T(\theta^*) \xrightarrow[T \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, I(\theta^*)). \quad (5.8)$$

5.3.3 Propriétés asymptotiques des estimateurs de maximum de vraisemblance

Les estimateurs du maximum de vraisemblance θ_T satisfont l'équation $\frac{\partial L_T(\theta)}{\partial \theta} = 0$. Dans ce qui suit, U_θ désigne un voisinage de θ .

Propriété 5.3.5 L'intensité conditionnelle définie par l'équation (5.4) possède les propriétés suivantes :

1. Pour tout $t \in \mathbb{R}_+$, $\lambda_\theta(t)$ est prévisible $\forall \theta \in \Theta$.
2. $\forall t \in \mathbb{R}_+$, $\lambda_\theta(t) \geq 0$ p. s. pour tout $\theta \in \Theta$.
3. $\forall t \in \mathbb{R}_+$, $\lambda_\theta(t)$ est p.s. continue en θ .

Preuve 5.3.3

1. Pour $t \geq 0$, $\lambda_\theta(t)$ est \mathcal{H}_{t-} -mesurable. De plus, λ_θ est continue à gauche sur \mathbb{R}_+^* , ceci nous mène au résultat.
2. $\forall t \in \mathbb{R}_+$, $\mu > 0$ et $\int_{-\infty}^{t-} \kappa(M)g(t-s)dN(s) \geq 0$ car g est une densité de probabilité et κ est une fonction positive.
3. $\lambda_\theta(t)$ est une fonction linéaire de θ donc λ_θ est continue en θ .

Remarque 5.3.2 De même, l'intensité conditionnelle définie par l'équation(5.1) possède les propriétés présentés ci-dessus.

Propriété 5.3.6 Soit θ_1 et θ_2 les paramètres respectivement de λ_{θ_1} et λ_{θ_2} . $\lambda_{\theta_1}(0) = \lambda_{\theta_2}(0)$ p. s. si et seulement si $\theta_1 = \theta_2$.

Propriété 5.3.7 (convergence en probabilité de θ_T ; Ogata1978) Considérant un processus ponctuel N stationnaire et ergodique défini l'équation 5.1 et les propriétés 5.3.5 et 5.3.6, de plus, supposons que pour tout $\theta \in \Theta$, il existe un voisinage U_θ tel que :

- i. $\sup_{\theta' \in U_\theta} |\lambda_{\theta'}(t) - \tilde{\lambda}_{\theta'}(t)| \xrightarrow[t \rightarrow +\infty]{P} 0$,
- ii. $\sup_{\theta' \in U_\theta} |\log \tilde{\lambda}_{\theta'}(t)| \leq a$, pour $\alpha > 0$, un moment fini d'ordre $(2 + \alpha)$ borné par rapport à t ,
- iii. $\forall \theta' \in U_\theta |\lambda_{\theta'}(0)| \leq v_0$ et $|\phi_{\theta'}(0)| \leq v_1$ avec v_0 et v_1 : variables aléatoires avec un moment fini d'ordre 2;

Alors, nous obtenons le résultat suivant,

$$\theta_T \xrightarrow[T \rightarrow +\infty]{P} \theta^*. \quad (5.9)$$

Théorème 5.3.2 (normalité asymptotique de θ_T ; Ogata, 1978)

Supposons que l'EMV θ_T satisfait l'équation $\frac{\partial \tilde{L}_T(\theta)}{\partial \theta} = 0$, les conditions du théorème 5.3.1 et les conditions suivantes :

Pour tout $\theta \in \Theta$, il existe un voisinage $U(\theta)$ de θ tel que si

$$\sup_{\theta' \in U(\theta)} |\lambda_{\theta'}^{(3)}(t)| = H(t) \text{ et } \sup_{\theta' \in U(\theta)} |\phi_{\theta'}^{(3)}(t)| = G(t),$$

alors $E[H(0)] < +\infty$ et $E[\lambda_{\theta^*}(0)^2 G(0)^2] < +\infty$.

Ainsi,

$$\sqrt{T}(\theta_T - \theta^*) \xrightarrow[T \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, I(\theta^*)^{-1}) \quad (5.10)$$

$$2 \left(\tilde{L}_T(\theta_T) - \tilde{L}_T(\theta^*) \right) \xrightarrow[T \rightarrow +\infty]{\mathcal{L}} \chi_d^2. \quad (5.11)$$

5.4 Modèle ETAS et processus de Dirichlet

Dans le modèle spatio-temporel défini par l'intensité conditionnelle (5.3), l'intensité d'arrière-plan est une fonction de la position spatiale. Plusieurs cas peuvent être considérés concernant l'expression de cette intensité d'arrière-plan. D'ailleurs, Ogata ([66], 1998), Ogata et Zhuang ([70], 2006) se sont intéressés aux situations suivantes :

- l'intensité d'arrière-plan $\mu(x)$ est une constante, ce qui correspond à un processus de Poisson homogène pour les séismes principaux ;
- l'intensité d'arrière-plan $\mu(x)$ s'écrit de la forme $\mu u(x)$ où μ est une constante et la fonction u est une surface bi-cubique B-splines.

Au lieu de cette approche déterministe, une approche semi-paramétrique permet de considérer cette intensité d'arrière-plan de la façon suivante :

$$(\mu(x))_{x \in X} \sim DP(\alpha, G_0) \quad (5.12)$$

où G_0 est un champ log-gaussien stationnaire isotropique (Brix et Diggle [13], 2001) .

Soient $(S_{ij})_{(i,j) \in \llbracket 1, n \rrbracket \times \llbracket 1, m \rrbracket}$ une partition de X et N^* , le processus ponctuel de Cox correspondant aux séismes principaux au cours de la période $[0, T]$. Conditionnellement à $(\mu(x))_{x \in S_{ij}}$, $N^*(S_{ij})$ suit une loi de Poisson de paramètre $T \int_{S_{ij}} \mu(x) dx$, ceci traduit l'indépendance conditionnelle des séismes principaux par rapport à l'environnement.

La formulation de ce modèle est proche de celle donnée par Kottas et al. ([51], 2007), mais ces derniers s'intéressant à la répartition spatiale d'une maladie faisaient intervenir la population à risque dans chacun des S_{ij} . Notons que le fait d'observer les nombres de séismes entre plusieurs dates $t_1 < t_2 < \dots < t_k = T$ sur chaque S_{ij} fournit des données longitudinales dont nous pouvons tirer partie pour mettre en œuvre l'inférence

statistique (Kottas et al. [51], 2007).

5.5 Étude de cas : l'arc des Petites Antilles

Si au départ, l'intérêt portait uniquement sur l'aspect temporel, c'est-à-dire que l'on s'intéressait à la succession chronologique des occurrences, dans un deuxième temps, les chercheurs ont pu aborder l'aspect spatio-temporel (représentation de la distribution des épacentres dans l'espace et le temps), puis l'aspect spatio-temporel avec marques (représentation de la distribution des épacentres dans l'espace et le temps en prenant la magnitude comme marque).

5.5.1 Présentation des données

Nous considérons un catalogue constitué des événements sismiques ayant eu lieu dans l'Arc des Petites Antilles entre les latitudes $10^\circ - 20^\circ\text{N}$ et longitudes $65^\circ - 58^\circ\text{O}$ durant la période de Janvier 1999 à Octobre 2004. Un total de 7095 séismes a été recensé dans le catalogue du CDSA (Centre de Données Sismiques des Antilles) et 2263 d'entre eux ont une magnitude au dessus de la magnitude de complétude fixée environ à $M_c = 3$ dans cette région (Bengoubou-Valérius et al. [7], 2008). La magnitude de référence utilisée dans ce catalogue est la magnitude M_d .

Un second catalogue était disponible et constitué des événements sismiques reportés dans le catalogue de l'ANSS (Advanced National Seismic System) du 8 Avril 1963 au 31 Décembre 2009 dans la même région (latitude 10 à 20°N et longitude 65 à 58°O). Néanmoins, nous avons été confrontés à des données incomplètes au début des années 60, de plus, ce catalogue contient plusieurs types de magnitude. Pour cette raison, nous utiliserons uniquement les données du CDSA afin de mener notre étude.

La figure 5.1 montre les localisations spatiales des séismes avec une attention particulière à ceux de magnitude supérieure à 4,1. Les figures 5.2 et 5.3 représentent les distributions temporelles par rapport à la magnitude.

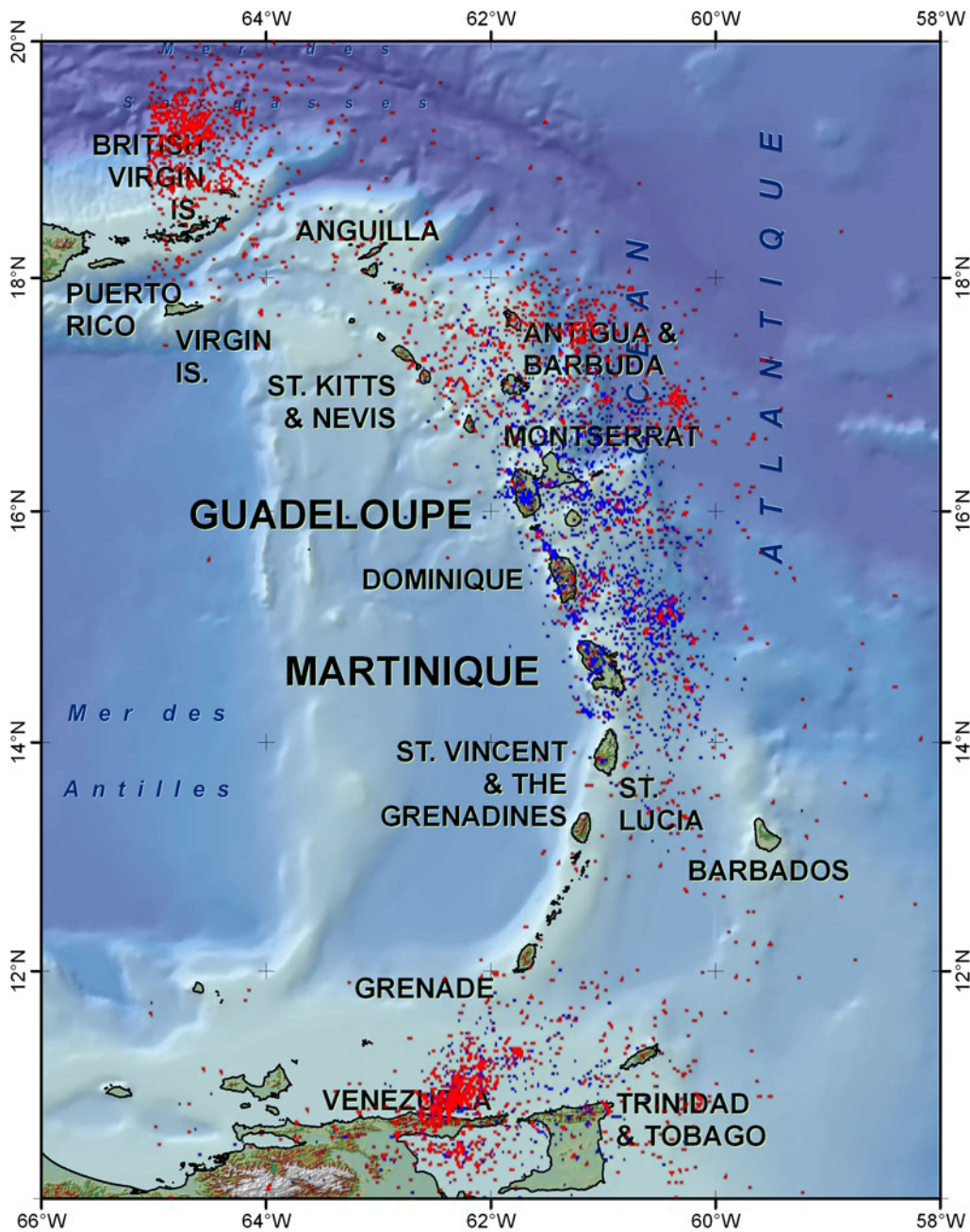


FIGURE 5.1 – Répartition spatiale des séismes des Petites Antilles de 1999 à 2004. Les points rouges correspondent aux magnitudes supérieures à 4,1 et les bleus ceux de magnitudes inférieures.

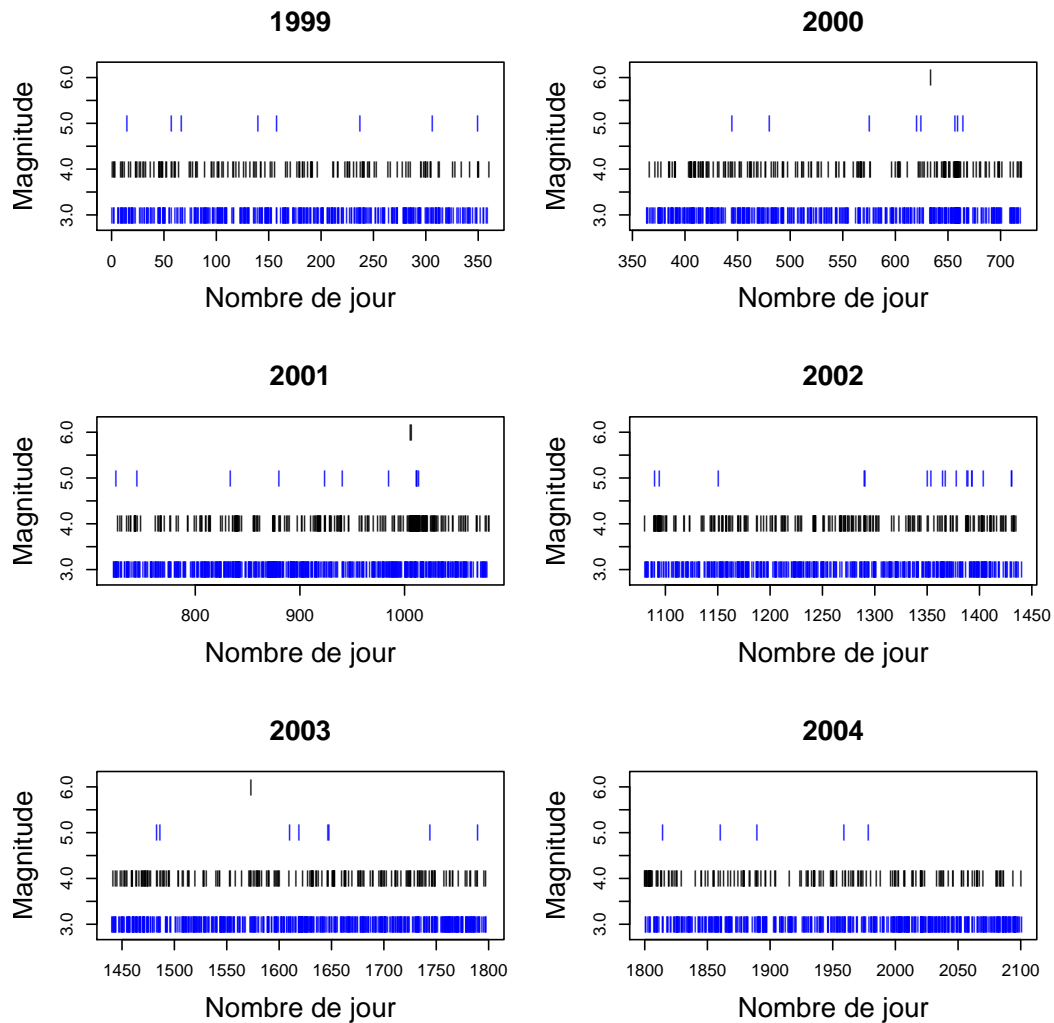


FIGURE 5.2 – Distribution temporelle des séismes des Petites Antilles de 1999 à 2004 de magnitude supérieure à 3 et arrondie à l'entier supérieur. Le jour numéro 1 correspond au premier Janvier 1999.

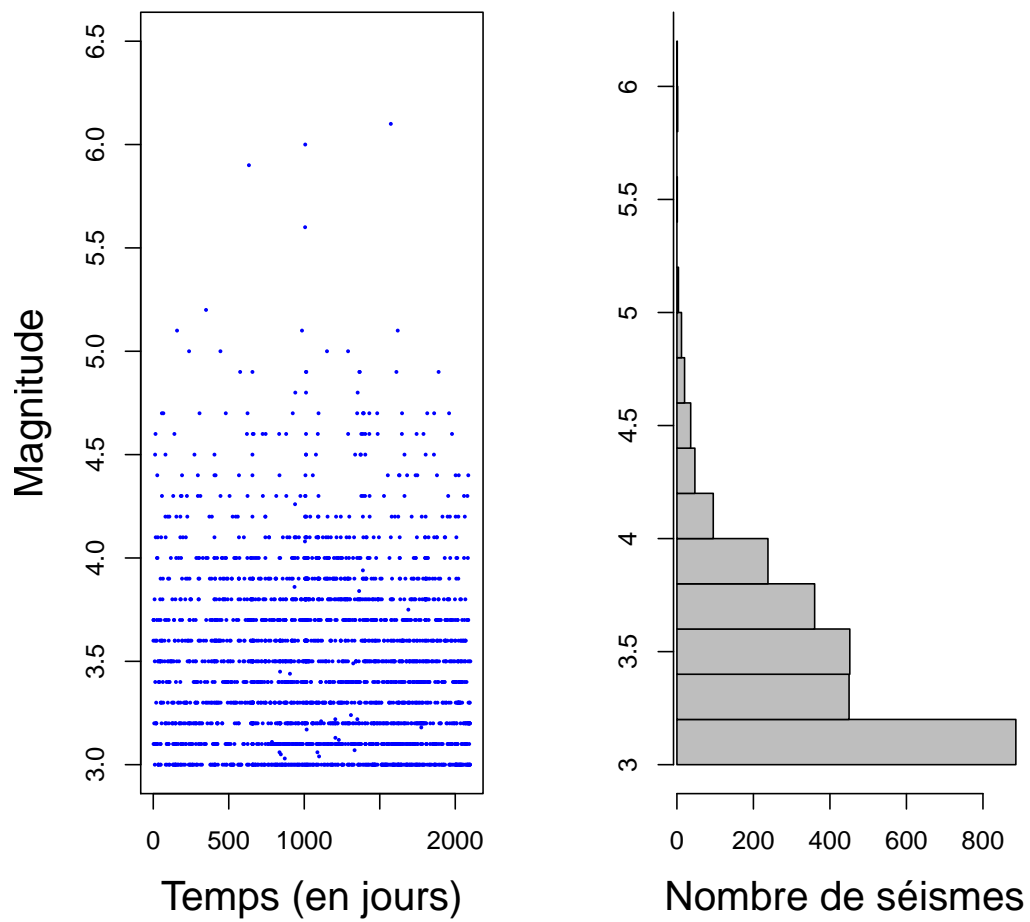


FIGURE 5.3 – Distribution temporelle des séismes par rapport à leur magnitude. Le jour numéro 1 correspond au premier Janvier 1999.

5.5.2 Méthode de thinning

Dans certains cas, il est nécessaire de construire un catalogue sismique approprié constitué des événements principaux. L'idée est de soustraire des groupes au catalogue complet, ces groupes doivent donc être identifiés et extraits. Knopoff ([50], 2000) a fourni un algorithme pour l'identification des répliques, des événements précurseurs et des clusters. Une récente description des techniques de declustering peut être trouvée en ligne (van Stiphout et al. [105], 2012). Cependant, puisque nous étudions l'évolution spatio-temporelle et temporelle des séismes, les répliques sont prises en compte comme décrit par Zhuang et al. ([114], 2005). Ces auteurs ont développé une méthode stochastique de declustering basée sur l'amincissement du processus ponctuel telle que la totalité des événements sismiques soit partagée en un processus d'arrière-plan et des processus déclencheurs. Ceci a été mené en calculant la probabilité pour chaque événement étant déclenché par un événement précédent. Comparée à la méthode conventionnelle de declustering, la méthode stochastique ne contient pas de paramètre arbitraire (taille de la fenêtre des répliques par exemple). Ainsi, les procédures conventionnelles donnent un catalogue différent pour chaque choix de ces paramètres et un choix optimal difficile des paramètres s'impose. Dans la méthode stochastique de declustering, les probabilités pour chaque événement généré sont obtenues par une méthode itérative et sont la clé pour identifier les événements d'arrière-plan des clusters.

5.5.3 Étude du modèle ETAS

Le modèle considéré correspond à l'expression 5.4 avec $\theta = (\mu, A, \alpha, c, p)$, plus précisément :

$$\lambda_{\theta}(t|H_{t-}) = \mu + \sum_{t_i < t} A e^{\alpha(M_i - M_c)} g_{c,p}(t - t_i) \quad (5.13)$$

Nous avons estimé les paramètres μ, A, α, c, p du modèle ETAS pour quatre fonctions déclenchantes différentes : les distributions Gamma (paramètre de taux c et paramètre de forme p), Weibull (paramètre d'échelle c et paramètre de forme p) et Log-normal (moyenne et déviation standard égales à c et p respectivement), la loi d'Omori modifiée (paramètre d'échelle c et un paramètre de forme p).

Remarque 5.5.1 *Selon la loi d'Omori modifiée, la fonction déclenchante a l'expression suivante : $g_{c,p}(t) = \frac{p-1}{c} \left(1 + \frac{t}{c}\right)^{-p}$ avec $c > 0$ et $p > 1$.*

$g_{c,p}(t) =$ avec	Distribution Gamma $\frac{t^{p-1}}{\Gamma(p)c^p}e^{-t/c}$ $c > 0$ et $p > 0$	Distribution Weibull $\frac{pt^{p-1}}{c^p}e^{-(t/c)^p}$ $c > 0$ et $p > 0$	Distribution Lognormal $\frac{1}{tp\sqrt{2\pi}}e^{-\frac{(\ln t - c)^2}{2p^2}}$ $c \in \mathbb{R}$ et $p > 0$	Loi d'Omori modifiée $\frac{p-1}{c} \left(1 + \frac{t}{c}\right)^{-p}$ $c > 0$ et $p > 1$
μ	1, 539	1, 389	0, 858	1, 367
A	0, 486	0, 525	0, 692	0, 544
a	0, 520	0, 532	0, 614	0, 470
c	0, 628	0, 935	0, 730	0, 20
p	0, 519	0, 556	3, 523	1, 628
L	2509, 886	2514, 050	2504, 625	2419, 543
AIC	-5009, 772	-5018, 100	-4999, 250	-4829, 086
χ^2	14, 980	14, 683	25, 843	44, 071
p -value	0, 183	0, 197	0, 007	0, 000

TABLE 5.1 – Estimation des paramètres du modèle ETAS, résultats de la log-vraisemblance et du test du Chi-2 pour trois fonctions déclenchantes.

À cet effet, nous avons écrit un programme dans l'environnement \mathbf{R} puisque le package *PtProcess* (Harte [39], 2010) ne prenait pas en compte trois de nos fonctions déclenchantes. Pour chaque distribution, la table 5.1 résume les résultats obtenus : estimations des paramètres, log-vraisemblance, AIC et test du Chi-2 pour statuer sur l'ajustement du modèle proposé aux données par le biais des dates rééchelonnés. La figure 5.4 illustre les quatre fonctions déclenchantes considérées avec la valeur des paramètres c et p obtenus dans la table 5.1 sur l'intervalle de temps $[0, 1]$. Le modèle ayant l'AIC le plus bas est celui prenant en compte la distribution de Weibull, il s'ajuste bien aux données avec une p -value égale à 0, 197 pour le test du Chi-2 ainsi qu'un processus résiduel et des temps d'attentes rééchelonnés proches de ceux attendus quand le modèle est vrai (figures 5.5).

5.5.4 Discussion

Les séismes peuvent être considérés comme des réalisations de processus ponctuels, ainsi, une modélisation temporelle et spatiale peut être faite. La théorie de la vraisemblance pour de tels modèles fournit des outils particulièrement utiles pour l'analyse

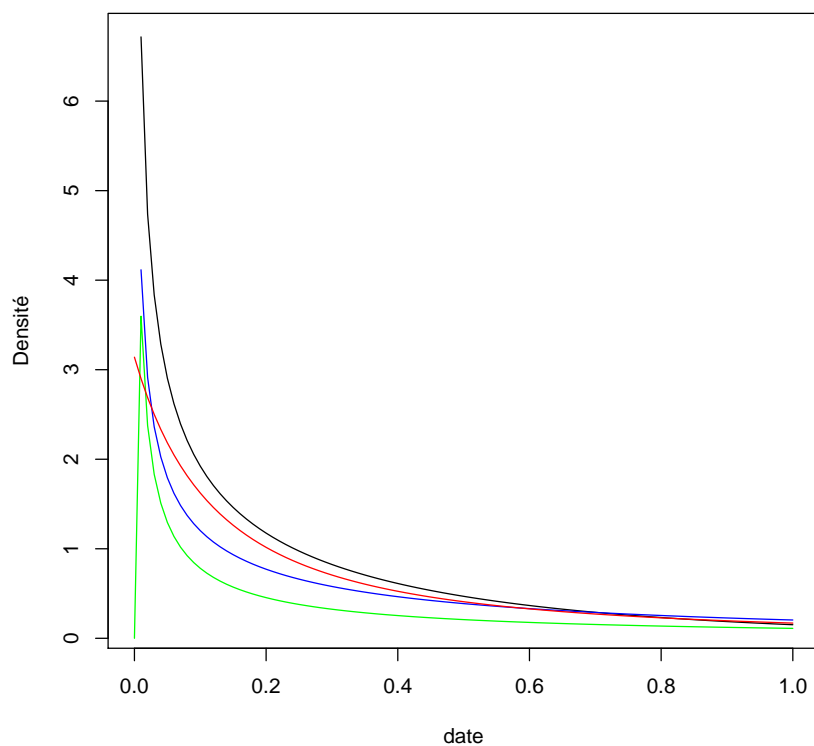
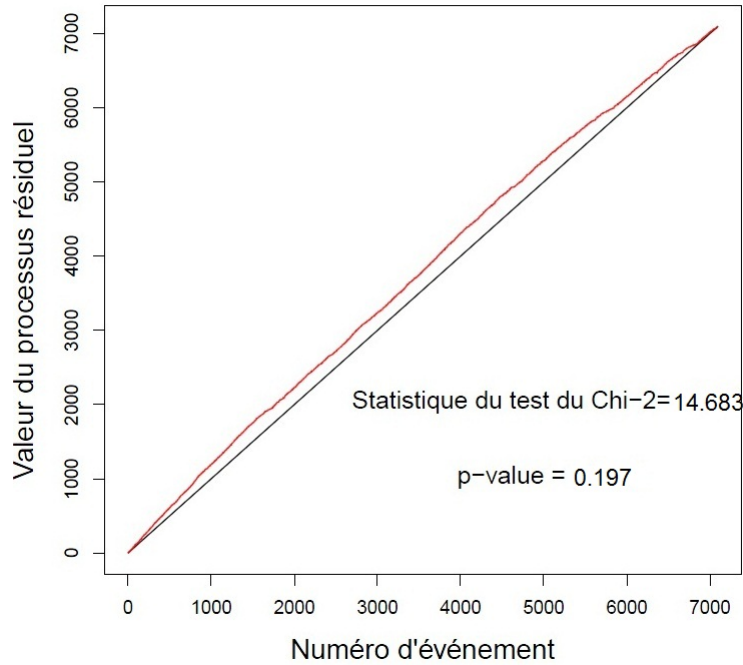
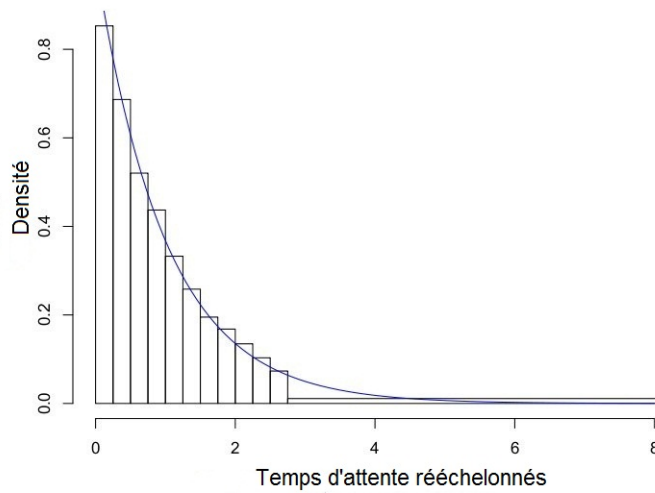


FIGURE 5.4 – Densité des distributions Gamma (noir), Weibull (bleu), Lognormal (vert) et loi d’Omori modifiée (rouge) sur l’intervalle $[0, 1]$ (valeurs de c et p issues de la table 5.1).



(a)



(b)

FIGURE 5.5 – Analyse résiduelle des événements sismiques avec comme fonction déclenchante la densité de Weibull : (a) Courbe de l'intensité cumulée en rouge, droite d'équation $y=x$ en noir ; (b) Histogramme des temps d'attente rééchelonnés, densité exponentielle de paramètre 1 en bleu.

statistique des activités sismiques. Des techniques de vraisemblance appropriées permettent d'estimer les paramètres du modèle considéré tels que l'intensité d'arrière-plan, la fonction déclenchante, et de tester l'ajustement au modèle. Une étude des variations spatiale et temporelle de taux de sismicité a été menée dans le cas des Petites Antilles à l'aide du catalogue sismique du CDSA contenant 7095 événements de Janvier 1999 à Octobre 2004. Considérant un modèle ETAS, différentes distributions pour la fonction déclenchante, incluant la loi d'Omori modifiée, ont été comparées. A cet effet, nous avons écrit un programme dans l'environnement **R**. Ceci nous a mené à conclure que la distribution de Weibull est la meilleure, dans la mesure où l'AIC est le moins élevé et les temps d'attente rééchelonnés sont proches de ceux attendus. Ainsi, la loi d'Omori modifiée adoptée par un grand nombre d'auteurs depuis l'article d'Ogata ([65], 1988) pour le modèle ETAS ne semble pas adaptée du tout aux données de sismicité de la région des Petites Antilles. La figure 5.4 indique la différence de comportement au voisinage de zéro pour les quatre fonctions proposées. La meilleure adéquation de la densité de Weibull par rapport à la loi d'Omori modifiée semble indiquer une plus forte agrégation des événements secondaires (répliques) à la suite d'un événement sismique principal.

Ces travaux sur les données sismiques de l'arc des Petites Antilles ont fait l'objet de plusieurs communications dans un congrès national et des congrès internationaux (Valmy [97], 2011; Valmy et Vaillant [98, 100, 99], 2011) et d'un article sous presse (Valmy et Vaillant [104], 2012).

Conclusions et perspectives

Divers domaines scientifiques nécessitent de prendre en considération des situations pour lesquelles nous observons des occurrences d'événements. Les positions et les dates de ces occurrences sont d'un intérêt capital si nous voulons décrire et prévoir de façon efficace le phénomène étudié. De telles situations se retrouvent par exemple en épidémiologie, sismologie et neurophysiologie. Telle est la motivation de notre travail de thèse pour lequel nous avons étudié d'une part les processus de Cox dirigé par un processus caché associé à un processus de Dirichlet, d'autre part, des processus ponctuels à mémoire de type Epidemic Type Aftershock Sequence (ETAS).

Dans un premier temps, des mélanges de processus ponctuels sont considérés pour prendre en compte la variabilité de l'environnement par une approche bayésienne hiérarchique. Nous nous sommes focalisés sur des données consistant en des positions d'occurrences spatiales avec une information incomplète sur les dates d'occurrences. Ensuite, nous avons considéré des données de type comptages dans des unités d'échantillonnage spatiales. Nous avons pris en compte des lois *a priori* de type stick-breaking pour une approche bayésienne semi-paramétrique avec l'utilisation du processus de Dirichlet et avec une intensité stochastique shot-noise généralisé. Nous nous sommes intéressés à des situations où des événements cachés ou non observables se produisent et ont une influence sur le nombre d'occurrences pour le phénomène observé. Nous avons proposé un modèle de processus de Cox ayant un processus intensité de type shot noise généralisé associé à un processus de Dirichlet. Les propriétés distributionnelles de ce modèle ont été déterminées ainsi que les lois limites relatives à certains paramètres, par exemple, la bande passante spatiale. Il permet de modéliser un phénomène consistant en des occurrences à des positions spatiales et des dates provoquées par des événements cachés se produisant selon un processus de Poisson. La contribution de chacun de ces événements suit un processus de Dirichlet centré sur une loi positive, telle la loi *Gamma*. L'une des motivations premières est l'étude d'infection observée dans le domaine végé-

tale par rapport à des maladies vectorielles pour lesquelles les infestations du vecteur ne sont pas observables. Un algorithme MCMC hybride Gibbs-Metropolis-Hastings a été proposé. Il est à sauts réversibles et permet de prendre en compte la variabilité de la dimension de l'espace des paramètres. Il fournit alors la loi *a posteriori* des paramètres suivants : nombre attendu d'événements cachés par unité d'espace et de temps, contribution espérée par événement caché, degré d'influence spatiale d'événements cachés, degré d'influence temporelle d'événements cachés et degré de corrélation des contributions des événements cachés. L'application à des données artificielles et à des données réelles montre le potentiel de la méthode.

Diverses perspectives s'ouvrent au travail précédemment décrit. Tout d'abord, nous comptons appliquer notre modèle à des phénomènes de santé publique (facteurs favorisant la résurgence des cas de dengue observés), d'épidémiologie végétale (étude la propagation de la cercosporiose aux Antilles françaises avec le CIRAD de Guadeloupe) et d'écologie (régulation des populations invasives de la fourmi manioc en Guadeloupe avec l'Université Paris-Est-Créteil, Laboratoire BIOEMCO, UMR 7618). Nos objectifs sont de développer des nouveaux outils en mettant à disposition un modèle stochastique susceptible de décrire et de caractériser l'évolution de certaines maladies vectorielles ou la propagation spatio-temporelle d'ennemis des cultures. Ensuite, nous comptons également améliorer la convergence des méthodes MCMC utilisées par rapport à l'efficacité de l'exploration de l'espace des paramètres.

Nous avons abordé dans un second temps l'intérêt des processus ponctuels en sismologie. En effet, une répartition spatiale et/ou temporelle d'occurrences de tremblement de terre dans une zone pendant une période donnée constitue une réalisation de processus ponctuel. La vraisemblance de processus ponctuel fournit un outil particulièrement utile pour l'analyse statistique des activités sismiques. Nous avons donc utilisé des techniques de vraisemblance appropriées afin d'estimer les paramètres du modèle considéré tels que l'intensité d'arrière-plan, la fonction déclenchante, puis de tester l'ajustement au modèle. Utilisant le catalogue sismique du CDSA de Janvier 1999 à Octobre 2004, une étude des variations spatiale et temporelle du taux de sismicité a été menée dans le cas des Petites Antilles. Par ailleurs, en considérant un modèle Epidemic Type Aftershock Sequence (ETAS), différentes distributions pour la fonction déclenchante, incluant la loi d'Omori modifiée, ont été comparées. Nous avons montré que la loi d'Omori modifiée ne s'ajuste pas aux données sismiques des Petites Antilles et la fonction déclenchante la plus adaptée est le modèle de Weibull. Cela implique que le temps d'attente entre répliques dans la zone des Petites Antilles est plus faible que celui des régions à sismicité

décrite par la loi d'Omori modifiée. À cet effet, nous avons écrit différents programmes dans l'environnement **R**. Nous suggérons également de considérer le modèle d'intensité conditionnelle spatio-temporelle telle que décrite par Ogata ([68], 2011) en y ajoutant une intensité d'arrière-plan, non plus déterministe, mais de loi un processus de Dirichlet centré sur un processus spatial log-gaussien.

Des contacts avec le BRGM de Guadeloupe ont été initiés au cours de cette thèse et une collaboration est envisagée sur des modèles de sismicité pour l'arc des Petites Antilles.

VOCABULAIRE RENCONTRÉ EN SISMOLOGIE

- Aléa sismique : probabilité pour une région de subir des secousses sismiques.
- Catalogue sismique : les catalogues de sismicité instrumentale rassemblent les données décrivant les séismes détectés par les sismomètres des réseaux de surveillance sismique. Les sismomètres mesurent, en un lieu géographique donné, la vitesse du déplacement selon les trois composantes de l'espace suite à une sollicitation sismique. Ils restituent cette information sur un sismogramme dont l'analyse par le sismologue permet de définir certaines caractéristiques du séisme (localisation, magnitude, mécanisme au foyer, ...).
- Données sismiques : ensemble des informations (localisation, date, heure, magnitude, ...) d'un séisme. Quand toutes ces informations sont disponibles pour un séisme, les données sont dites complètes.
- Échelle de Richter : échelle de référence qui évalue l'énergie des séismes par la valeur de la magnitude.
- Épicentre microsismique (ou épicentre réel) : point de la surface terrestre situé à la verticale du foyer (ou hypocentre) d'un séisme.
- Événements sismiques : occurrences lors desquelles de l'énergie est libérée dans l'écorce terrestre, ce qui a entraîné une série d'ondes sismiques.
- Foyer (ou hypocentre) : point de départ de la rupture des roches.
- Intensité (d'un séisme) : mesure des secousses du sol évaluée à partir des dégâts subis par les constructions, les modifications de la surface du sol et les impressions des témoins.
- Latitude : la distance angulaire, généralement exprimée en degré, séparant ce point de l'équateur. Les latitudes se comptent de -90° à $+90^\circ$ et la latitude de l'équateur est 0° .
- Longitude : la distance angulaire, généralement exprimée en degré, séparant ce point du méridien de Greenwich. La longitude se compte de 0° à 360° ou de -180° à $+180^\circ$ et la longitude de Greenwich est 0° .
- Magnitude : unité de grandeur d'un séisme comprise entre 0 et 9 qui traduit l'énergie libérée par le séisme au niveau du foyer : on utilise l'échelle de Richter. Ils existent plusieurs types de magnitude, dont certaines sont détaillées dans le tableau ci-dessous :

Type de magnitude	Domaine d'application	Rayon d'application
Durée (Md)	< 4	0 – 400 km
Local (Ml)	2 – 6	0 – 400
Onde de surface (Ms)	5 – 8	20 – 180 °
Moment (Mw)	> 3,5	
Ondes de volume (Mb)	4 – 7	16 – 200 °

Commentaires

La magnitude :

- de durée Md est calculée pour des petits séismes, elle rend compte de la durée du signal entre l'arrivée de la première onde et l'instant où le signal retrouve le niveau du bruit de fond sismique initial.
- locale Ml est définie pour des séismes locaux (courte distance), elle dépend d'une relation entre la distance d'enregistrement et l'amplitude maximale observée.
- des ondes de surface Ms est valable pour toutes les distances, elle est calculée à partir des ondes de surface de Rayleigh et caractérise des séismes importants à faible profondeur.
- de moment Mw est directement reliée au moment sismique du séisme, elle est utilisée pour les gros séismes pour lesquels les Ms et Ml sous-estiment les énergies libérées.
- des ondes de volume Mb est calculée à partir des ondes de compression P et elle est généralement utilisée pour les séismes profonds qui ne génèrent pas d'ondes de surface.

Les sismologues indiquent la taille d'un tremblement de terre en magnitude. Il existe différentes façons de mesurer la magnitude à partir des sismomètres, car chaque méthode œuvre sur un champ limité de magnitude. Certaines méthodes sont basées sur les ondes de volume, qui se propagent profondément selon la structure des roches, d'autres sur les ondes de surface, qui se propagent dans les couches supérieures.

- Magnitude de complétude d'un catalogue : magnitude au delà de laquelle les données sismiques sont complètes.
- Précurseurs (séismes) : petits séismes précédant le séisme principal dans une série d'événements localisés dans une zone proche.
- Répliques sismiques (ou événements secondaires) : séismes de moindre importance pouvant succéder au séisme principal.
- Risque sismique : conséquences de l'aléa sismique sur le plan humain, matériel et économique.
- Séisme principal (ou majeur ou choc principal) : le plus grand tremblement de terre dans une série.
- Séisme (ou tremblement de terre ou secousse tellurique) : mouvement sur une faille qui engendre des secousses plus ou moins violentes à la surface du sol.
- Sismicité (ou séismicité) : distribution géographique des séismes en fonction du temps.
- Sismologie : science qui étudie les tremblements de terre naturels ou artificiels, et d'une manière générale la propagation des ondes sismiques à travers la Terre.

PROGRAMMES INFORMATIQUES

COMPLÉMENTAIRES SOUS *R*

B.1 Modèle de Cox associé à un Dirichlet

B.1.1 Cas de données de positions spatio-temporelles

```

MCMCPPSN <- fonction(h, alpha0, b0, mu0, sigma1_0, sigma2_0, N, positions,
dateobs, L1, L2, MAXM) {
# MCMC pour le processus ponctuel shot noise généralisé sur le rectangle [0, L1] × [0, L2]
# et l'intervalle [0, T] où T est la dernière date d'observations
# N est le nombre d'itérations de la chaîne
# alpha0, b0, mu0, sigma1_0, sigma2_0 sont les valeurs initiales
# h est le vecteur (de longueur 5) des paramètres de dilatation de la loi instrumentale
halpha = h[1]; hb = h[2]; hmu = h[3]; hsigma1 = h[4]; hsigma2 = h[5]
# positions est la matrice à 3 colonnes des positions spatiales et des dates de première
# détection des occurrences
# positions a donc en colonne 1 les abscisses, en colonne 2 les ordonnées, et en colonne
# 3 les dates de 1ère détection
# dateobs est le vecteur des dates d'observation (dans l'ordre)
# la matrice (Nx5) en sortie "param" correspond aux N transitions, aux 5 paramètres
# library(actuar) installe le package "actuar" afin d'utiliser la fonction "dinvgamma"
# donnant la densité de la loi inverse-gamma
tt = dateobs[length(dateobs)] # tt est la dernière date d'observations
param = matrix(0, nrow = N, ncol = 5) # Itération 1
logalpha = log(alpha0) + halpha * rnorm(1)
logb = log(b0) + hb * rnorm(1)
logmu = log(mu0) + hmu * rnorm(1)
logsigma1 = log(sigma1_0) + hsigma1 * rnorm(1)
logsigma2 = log(sigma2_0) + hsigma2 * rnorm(1)

```



```

b = exp(logb); mu = exp(logmu); alpha = exp(logalpha)
M = max(1, rpois(1, lambda = mu * L1 * L2 * tt))
print(M)
a = rgamma(1, shape = b, scale = 1)
if(M > 1) { for(i in 2 : M) {
u = runif(1); L = (u < alpha / (alpha + i - 1)); j = sample(1 : (i - 1), 1)
a = c(a, L * rgamma(1, shape = b, scale = 1) + (!L) * a[j]) } }
ys = cbind(matrix(L1 * runif(M), ncol = 1), matrix(L2 * runif(M), ncol = 1), matrix(tt * runif(M),
ncol = 1))
param[1, ] = c(logalpha, logb, logmu, logsigma1, logsigma2)
#Itérations pour N >= t > 1
for(tin 2 : N) { if(0 == t%%10) { print(t) }
param[t, 1] = movealpha(param[t - 1, 1], a, param[t - 1, 2], M, halpha)
param[t, 2] = moveb(param[t, 1], a, param[t - 1, 2], M, hb)
param[t, 3] = movemu(param[t - 1, 3], M, hmu)
param[t, 4] = movesigma1(param[t - 1, 4], param[t - 1, 5], M, a, ys, positions
, dateobs, L1, L2, hsigma1)
param[t, 5] = movesigma2(param[t, 4], param[t - 1, 5], M, a, ys, positions,
dateobs, L1, L2, hsigma2)
Mays = moveMays(param[t, 1], param[t, 2], param[t, 3], param[t, 4],
param[t, 5], M, a, ys, positions, dateobs, L1, L2, MAXM)
M = Mays[[1]]; a = Mays[[2]]; ys = Mays[[3]] }
param }

```

```

calintcum < -function(a, ys, s1, s2, x, y, t) {
#Calcul de l'intensité cumulée en un point d'abscisse x, d'ordonnée y à la date t
#a est le vecteur des contributions à l'intensité
#ys est la matrice à 3 colonnes des coordonnées spatio-temporelles du shot noise
#s1 est la bande passante spatiale, s2 est la bande passante temporelle
z = sum(a * dnorm(x - ys[, 1], sd = s1) * dnorm(y - ys[, 2], sd = s1) * pexp(t - ys[, 3], rate = 1/s2))
z }

```

```

simuPPSN < -function(alpha, b, mu, s1, s2, T, L1, L2) {
#simulation d'un processus de Cox d'intensité shot noise généralisé
#sur l'intervalle [0, T] sur le rectangle [0, L1] × [0, L2]
#a est un vecteur de longueur M
M = max(1, rpois(n = 1, lambda = mu))
#En fait, mu doit être choisi suffisamment grand de sorte que M=0 négligeable
a = rgamma(1, scale = 1, shape = b)
if(M > 1) { for(i in 2 : M) {
u = runif(1); L = u < alpha / (alpha + i - 1); j = sample(1 : (i - 1), 1)
a = c(a, L * rgamma(1, shape = b) + (!L) * a[j]) } }

```

```

ys = cbind(matrix(L1*runif(M), ncol = 1), matrix(L2*runif(M), ncol = 1), matrix(T*runif(M),
ncol = 1))
maxintens = max(sum(a/(2 * pi * s2 * (s1)^2))) #calcul du max de l'intensité
posidate = poihomspatem(maxintens, T, L1, L2, FALSE)
#matrice à 3 colonnes (abscisse, ordonnée, date)
n = dim(posidate)[1]
print(n)
u = runif(n); z = NULL
for(i in 1 : n) {
x = posidate[i, 1]; y = posidate[i, 2]; t = posidate[i, 3];
z = c(z, sum(a*dnorm(x-ys[, 1], sd = s1)*dnorm(y-ys[, 2], sd = s1)*dexp(t-ys[, 3], rate = 1/s2))) }
L = u < (z/maxintens)
resu = posidate[L, ]
resu }

traceintens < -function(mu, s1, s2, b, alpha) {
#simulation et tracé de l'intensité cumulée en t=1 d'un PP shot noise généralisé
#sur un carré unité dans l'intervalle [0,1]
x = y = seq(0, 1, length = 100); t = 1
M < -rpois(1, mu)
a = rgamma(1, scale = 1, shape = b)
if(M > 1) { for(i in 2 : M) {
u = runif(1); L = u < alpha; j = sample(1 : (i - 1), 1)
a = c(a, (L * rgamma(1, shape = b) + (!L) * a[j])) } }
ys = matrix(runif(3 * M), ncol = 3) }

intens < -function(x, y) {
#Calcul de l'intensité cumulée en (x,y), la date t étant fixée
jva = sum(a * dnorm(x - ys[, 1], sd = s1) * dnorm(y - ys[, 2], sd = s1) * pexp(t - ys[, 3], rate = 1/s2))
jva }

poihomspatem < -function(intens, T, L1, L2, jva) {
#nombre de points dans S×[0,T]
n = rpois(1, intens * T * L1 * L2)
#positions et dates des occurrences
posidates = matrix(c(L1 * runif(n), L2 * runif(n), sort(T * runif(n))), ncol = 3)
if(jva) { }
posidates }

movealpha < -function(logalpha, a, logb, M, halpha) {
logalphaneu = logalpha + halpha * rnorm(1)#instrumentale gaussienne
alpha = exp(logalpha)

```

```

alphanew = exp(logalphanew)
Ralpha = dgamma(alphanew, shape = 9, scale = 2, log = TRUE) - dgamma(alpha, shape = 9,
scale = 2, log = TRUE) + calprDP(a, logalphanew, logb, M) - calprDP(a, logalpha, logb, M)
if(Ralpha > log(runif(1))) { resu = logalphanew }
else { resu = logalpha }
resu }

moveb < -function(logalpha, a, logb, M, hb) {
logbnew = logb + hb * rnorm(1) #instrumentale log-gaussienne
Rb = logb - 0.5*logb^2 - logbnew + 0.5*logbnew^2 + calprDP(a, logalpha, logbnew, M) - calprDP(a, logalpha, logb, M)
if(Rb > log(runif(1))) { resu = logbnew }
else { resu = logb }
resu }

movemu < -function(logmu, M, hmu) {
logmunew = logmu + hmu * rnorm(1) #instrumentale log-gaussienne
munew = exp(logmunew); mu = exp(logmu)
Rmu = dlnorm(munew, meanlog = 1, sdlog = 1, log = TRUE) - dlnorm(mu, meanlog = 1, sdlog =
1, log = TRUE) + dpois(M, lambda = munew, log = TRUE) - dpois(M, lambda = mu, log = TRUE)
if(Rmu > log(runif(1))) { resu = logmunew }
else { resu = logmu }
resu }

movesigma1 < -function(logsigma1, logsigma2, M, a, ys, positions, dateobs, L1, L2, hsigma1) {
logsigma1new = logsigma1 + hsigma1 * rnorm(1) #instrumentale log-gaussienne
sigma1new = exp(logsigma1new); sigma1 = exp(logsigma1); sigma2 = exp(logsigma2)
#La ligne précédente est faite pour utiliser à la ligne suivante "callogvrais"
Rsigma1 = callogvrais(a, ys, sigma1new, sigma2, positions, dateobs, L1, L2)
- callogvrais(a, ys, sigma1, sigma2, positions, dateobs, L1, L2)
if(Rsigma1 > log(runif(1))) { resu = logsigma1new }
else { resu = logsigma1 }
resu }

movesigma2 < -function(logsigma1, logsigma2, M, a, ys, positions, dateobs, L1, L2, hsigma2) {
logsigma2new = logsigma2 + hsigma2 * rnorm(1) #instrumentale log-gaussienne
sigma2new = exp(logsigma2new); sigma2 = exp(logsigma2); sigma1 = exp(logsigma1)
#La ligne précédente est faite pour utiliser à la ligne suivante "callogvrais"
Rsigma2 = sigma2new - sigma2 + callogvrais(a, ys, sigma1,
sigma2new, positions, dateobs, L1, L2) - callogvrais(a, ys, sigma1, sigma2, positions, dateobs, L1, L2)
if(Rsigma2 > log(runif(1))) { resu = logsigma2new }
else { resu = logsigma2 }
resu }

```

```

moveMays <- fonction(logalpha, logb, logmu, logsigma1, logsigma2, M, a,
ys, positions, dateobs, L1, L2, MAXM) {
alpha = exp(logalpha); b = exp(logb); mu = exp(logmu);
sigma1 = exp(logsigma1); sigma2 = exp(logsigma2)
LL = as.numeric(runif(1) < 0.5); LL = 2 * LL - 1
Mnew = min(max(2, M + LL), MAXM)
# M >= 2, Mnew vaut M+1 avec une proba 1/2 ou M-1 avec une proba 1/2
if(Mnew <= M) {
choix = sample.int(M, Mnew)
anew = a[choix]
ysnew = ys[choix,]
else {
u = runif(1); L = u < alpha/(alpha + M); j = sample(1 : M, 1)
anew = c(a, (L * rgamma(1, shape = b, scale = 1) + (!L) * a[j]))
ysnew = rbind(ys, c(runif(1, max = L1), runif(1, max = L2), runif(1, max = dateobs[length(dateobs)])))
}
RMays = dpois(Mnew, lambda = mu, log = TRUE) - dpois(M, lambda = mu, log = TRUE)
+ calprDP(anew, logalpha, logb, Mnew) - calprDP(a, logalpha, logb, M)
RMays = RMays + callogvrais(anew, ysnew, sigma1, sigma2, positions, dateobs, L1, L2)
- callogvrais(a, ys, sigma1, sigma2, positions,
dateobs, L1, L2)
if(RMays > log(runif(1))) { resu = list(Mnew, anew, ysnew) }
else { resu = list(M, a, ys) }
resu }

```

```

calprDP <- fonction(a, logalpha, logb, M) {
# Calcul de la log-proba des contributions "a" selon une loi Gamma
# M doit être la longueur du vecteur a
alpha = exp(logalpha); b = exp(logb)
z = dgamma(a[1], shape = b, scale = 1, log = TRUE)
if(M > 1) {
for(jin2 : M) {
z = z + log((1/(alpha + j - 1)) * (alpha * dgamma(a[j], shape = b, scale = 1) + sum(a[j] == a[1 :
(j - 1)]))) } }
z }

```

```

callogvrais <- fonction(a, ys, s1, s2, positions, dateobs, L1, L2) {
# Calcul de la log-vraisemblance pour des observations sur le rectangle [0, L1] x [0, L2]
# positions est la matrice à 3 colonnes des positions spatiales et des dates de 1ère détection
# des occurrences
# dateobs est le vecteur des dates d'observation
# a est le vecteur des contributions à l'intensité
# ys est la matrice à 3 colonnes des coordonnées spatio-temporelles du shot noise

```

```

‡ s1 est la bande passante spatiale, s2 est la bande passante temporelle
I = length(dateobs) ‡ I est le nombre de cartes observées
n = dim(positions)[1] ‡ n est le nombre de sites infectés
intcumtot = sum(a * (pnorm(L1 - ys[, 1], sd = s1) - pnorm(0 - ys[, 1], sd = s1)) * (pnorm(L2 -
ys[, 2], sd = s1) - pnorm(0 - ys[, 2], sd = s1))
* pexp(dateobs[I] - ys[, 3], rate = 1/s2))
z = 0
Ldate = (positions[, 3] <= dateobs[1]) ‡ vecteur logique des occurrences jusqu'à la 1ère date d'observation
posx = positions[Ldate, 1] ‡ abscisses des occurrences jusqu'à la 1ère date d'observation
posy = positions[Ldate, 2] ‡ ordonnées des occurrences jusqu'à la 1ère date d'observation
for(k in 1 : sum(Ldate)) { z = z + log(calintcum(a, ys, s1, s2, posx[k], posy[k], dateobs[1])) }
if(I > 1) {
for(i in 2 : I) { ‡ itérations sur les intervalles ]t_{i-1}, t_i] d'observations
Ldate = (positions[, 3] <= dateobs[i]) & (positions[, 3] > dateobs[i-1])
‡ vecteur logique des occurrences dans ]t_{i-1}, t_i]
posx = positions[Ldate, 1]; posy = positions[Ldate, 2]
for(k in 1 : sum(Ldate)) {
z = z + log(calintcum(a, ys, s1, s2, posx[k], posy[k], dateobs[i])
- calintcum(a, ys, s1, s2, posx[k], posy[k], dateobs[i-1])) } } }
z = z - intcumtot
z }

```

B.1.2 Cas de données de comptages spatiales

```

MCMCPPSN <- function(h, alpha0, b0, mu0, sigma0, N, comptages,
sommets, L1, L2, MAXM) {
‡ MCMC pour le processus ponctuel shot noise généralisé avec données sous forme de
‡ comptage spatial dans le rectangle [0, L1] × [0, L2]
‡ N est le nombre d'itérations de la chaîne
‡ alpha0, b0, mu0, sigma0 sont les valeurs initiales
‡ h est le vecteur (de longueur 5) des paramètres de dilatation de la loi instrumentale
halpha = h[1]; hb = h[2]; hmu = h[3]; hsigma = h[4]
‡ comptages est le vecteur des comptages dans des unités spatiales rectangulaires
‡ sommets est une matrice à 4 colonnes (coordonnées du point en bas à gauche, longx, longy)
‡ permettant de calculer l'intégrale sur chaque unité spatiale rectangulaire
‡ la matrice (N×4) en sortie "param" correspond aux N transitions, aux 4 paramètres
‡ library(actuar) installe le package "actuar" afin d'utiliser la fonction "dinvgamma"
‡ donnant la densité de la loi inverse-gamma
param = matrix(0, nrow = N, ncol = 4)
‡ Itération 1
logalpha = log(alpha0) + halpha * rnorm(1); logb = log(b0) + hb * rnorm(1)
logmu = log(mu0) + hmu * rnorm(1); logsigma = log(sigma0) + hsigma * rnorm(1)
b = exp(logb); mu = exp(logmu); alpha = exp(logalpha)
M = max(1, rpois(1, lambda = mu * L1 * L2))
print(M)
a = rgamma(1, shape = b, scale = 1)

```

```

if(M > 1) {
for(i in 2 : M) {
u = runif(1); L = (u < alpha/(alpha + i - 1)); j = sample(1 : (i - 1), 1)
a = c(a, L * rgamma(1, shape = b, scale = 1) + (!L) * a[j]) } }
ys = cbind(matrix(L1 * runif(M), ncol = 1), matrix(L2 * runif(M), ncol = 1))
param[1,] = c(logalpha, logb, logmu, logsigma)
# Itérations pour N >= t > 1
for(tin2 : N) {
if(0 == t%%10) { print(t) }
param[t, 1] = movealpha(param[t - 1, 1], a, param[t - 1, 2], M, halpha)
param[t, 2] = moveb(param[t, 1], a, param[t - 1, 2], M, hb)
param[t, 3] = movemu(param[t - 1, 3], M, hmu)
param[t, 4] = movesigma(param[t - 1, 4], M, a, ys, comptages, sommets
, L1, L2, hsigma)
Mays = moveMays(param[t, 1], param[t, 2], param[t, 3], param[t, 4], M,
a, ys, comptages, sommets, L1, L2, MAXM)
M = Mays[[1]]; a = Mays[[2]]; ys = Mays[[3]]; }
param }

simuPPSN <- function(alpha, b, mu, s1, s2, T, L1, L2) {
#simulation d'un processus de Cox d'intensité shot noise généralisé
#sur l'intervalle [0, T] sur le rectangle [0, L1] x [0, L2]
#a est un vecteur de longueur M
M = max(1, rpois(n = 1, lambda = mu)) #mu doit être choisi assez grand de sorte que M=0 négligeable
a = rgamma(1, scale = 1, shape = b)
if(M > 1) {
for(i in 2 : M) {
u = runif(1); L = u < alpha/(alpha + i - 1); j = sample(1 : (i - 1), 1)
a = c(a, L * rgamma(1, shape = b) + (!L) * a[j]) } }
ys = cbind(matrix(L1 * runif(M), ncol = 1), matrix(L2 * runif(M), ncol = 1), matrix(T * runif(M),
ncol = 1))
maxintens = max(sum(a/(2 * pi * s2 * (s1)^2))) #calcul du max de l'intensité
posidate = poiHomspatem(maxintens, T, L1, L2, FALSE) #matrice à 3 colonnes (abscisse, ordonnée, date)
n = dim(posidate)[1]
print("n = "); print(n)
u = runif(n)
z = NULL
for(i in 1 : n) {
x = posidate[i, 1]; y = posidate[i, 2]; t = posidate[i, 3];
z = c(z, sum(a * dnorm(x - ys[, 1], sd = s1) * dnorm(y - ys[, 2], sd = s1) * dexp(t - ys[, 3], rate = 1/s2))) }
L = u < (z/maxintens)
resu = posidate[L, ]

```

```
resu }
```

```
traceintens <- fonction(mu, s1, s2, b, alpha) {
#simulation et tracé de l'intensité cumulée en t=1 d'un processus ponctuelshot noise généralisé
#sur un carré unité dans l'intervalle [0,1]
x = y = seq(0, 1, length = 100); t = 1
M <- rpois(1, mu)
a = rgamma(1, scale = 1, shape = b)
if(M > 1) {
for(i in 2 : M) {
u = runif(1); L = u < alpha; j = sample(1 : (i - 1), 1)
a = c(a, (L * rgamma(1, shape = b) + (!L) * a[j])) } }
ys = matrix(runif(3 * M), ncol = 3) }
```

```
intens <- fonction(x, y) {
#Calcul de l'intense cumulée en (x,y), la date t étant fixée
lva = sum(a * dnorm(x - ys[, 1], sd = s1) * dnorm(y - ys[, 2], sd = s1) * pexp(t - ys[, 3], rate = 1/s2))
lva }
```

```
poihomspatem <- fonction(intens, T, L1, L2, jva) {
#nombre de points dans S×[0,T]
n = rpois(1, intens * T * L1 * L2)
#positions et dates des occurrences
posidates = matrix(c(L1 * runif(n), L2 * runif(n), sort(T * runif(n))), ncol = 3)
if(jva) { }
posidates }
```

```
movealpha <- fonction(logalpha, a, logb, M, halpha) {
logalphanew = logalpha + halpha * rnorm(1)#instrumentale gaussienne
alpha = exp(logalpha); alphanew = exp(logalphanew)
Ralpha = dgamma(alphanew, shape = 9, rate = .5, log = T) - dgamma(alpha, shape = 9,
rate = .5, log = T) + calprDP(a, logalphanew, logb, M) - calprDP(a, logalpha, logb, M)
if(Ralpha > log(runif(1))) { resu = logalphanew }
else { resu = logalpha }
resu }
```

```
moveb <- fonction(logalpha, a, logb, M, hb) {
logbnew = logb + hb * rnorm(1)#instrumentale log-gaussienne
Rb = logb - 0.5 * logb2 - logbnew + 0.5 * logbnew2 + calprDP(a, logalpha, logbnew, M)
- calprDP(a, logalpha, logb, M)
if(Rb > log(runif(1))) { resu = logbnew }
else { resu = logb }
```

```
resu }
```

```
movemu <- fonction(logmu, M, hmu) {
logmunew = logmu + hmu * rnorm(1) #instrumentale log-gaussienne
munew = exp(logmunew); mu = exp(logmu)
Rmu = dlnorm(munew, meanlog = 1, sdlog = 1, log = T) - dlnorm(mu, meanlog = 1,
sdlog = 1, log = T) + dpois(M, lambda = munew, log = T) - dpois(M, lambda = mu, log = T)
if(Rmu > log(runif(1))) { resu = logmunew }
else { resu = logmu }
resu }
```

```
movesigma <- fonction(logsigma, M, a, ys, comptages, sommets, L1, L2, hsigma) {
logsigmanew = logsigma + hsigma * rnorm(1) #instrumentale log-gaussienne
sigmanew = exp(logsigmanew); sigma = exp(logsigma);
#La ligne précédente est faite pour utiliser à la ligne suivante "callogvrais"
Rsigma = callogvrais(a, ys, sigmanew, comptages, sommets) - callogvrais(a, ys, sigma, comptages, sommets)
if(Rsigma > log(runif(1))) { resu = logsigmanew }
else { resu = logsigma }
resu }
```

```
moveMays <- fonction(logalpha, logb, logmu, logsigma, M, a, ys, comptages, sommets, L1, L2, MAXM) {
alpha = exp(logalpha); b = exp(logb); mu = exp(logmu); sigma = exp(logsigma);
LL = as.numeric(runif(1) < 0.5); LL = 2 * LL - 1
Mnew = min(max(2, M + LL), MAXM)
#M >= 2, Mnew vaut M+1 avec une proba 1/2 ou M-1 avec une proba 1/2
if(Mnew <= M) {
choix = sample.int(M, Mnew); anew = a[choix]; ysnew = ys[choix, ] }
else {
u = runif(1); L = u < alpha / (alpha + M); j = sample(1 : M, 1)
anew = c(a, (L * rgamma(1, shape = b, scale = 1) + (!L) * a[j]))
ysnew = rbind(ys, c(runif(1, max = L1), runif(1, max = L2))) }
RMays = dpois(Mnew, lambda = mu, log = TRUE) - dpois(M, lambda = mu, log = TRUE)
+ calprDP(anew, logalpha, logb, Mnew) - calprDP(a, logalpha, logb, M)
RMays = RMays + callogvrais(anew, ysnew, sigma, comptages, sommets)
- callogvrais(a, ys, sigma, comptages, sommets)
if(RMays > log(runif(1))) { resu = list(Mnew, anew, ysnew) }
else { resu = list(M, a, ys) }
resu }
```

```
calprDP <- fonction(a, logalpha, logb, M) {
# Calcul de la log-proba du vecteur des contributions "a" selon une loi Gamma
# M doit être la longueur du vecteur a
```



```

alpha = exp(logalpha); b = exp(logb); z = dgamma(a[1], shape = b, scale = 1, log = TRUE)
if(M > 1) {
  for(j in 2 : M) {
    z = z + log((1/(alpha + j - 1)) * (alpha * dgamma(a[j], shape = b, scale = 1)
+ sum(a[j] == a[1 : (j - 1)]))) } }
z }

callogvrais <- fonction(a, ys, sigma, comptages, sommets) {
# calcul de la log-vraisemblance pour des comptages spatiaux
# comptages est le vecteur des comptages dans des unités spatiales rectangulaires
# sommets est une matrice à 4 colonnes (coordonnées du point en bas à gauche, longx, longy)
# permettant de calculer l'intégrale sur chaque unité spatiale rectangulaire
# a est le vecteur des contributions à l'intensité
# ys est la matrice à 2 colonnes des coordonnées spatiales du shot noise
# sigma est la bande passante spatiale
n = length(comptages); intcums = rep(0, n)
for(i in 1 : n) { intcums[i] = sum(a * (pnorm(sommets[i, 1] + sommets[i, 3] - ys[, 1], sd = sigma)
- pnorm(sommets[i, 1] - ys[, 1], sd = sigma)) * (pnorm(sommets[i, 2] + sommets[i, 4] - ys[, 2], sd =
sigma)
- pnorm(sommets[i, 2] - ys[, 2], sd = sigma))) }
z = sum(dpois(comptages, intcums, log = T))
z }

```

B.2 Calcul du log-rapport de vraisemblance

Considérant le modèle de Cox associé à un Dirichlet dans le cas de données de comptages spatiales, les scripts suivants permettent de calculer le log rapport de vraisemblance dans les cas où α est quelconque, α tend vers 0 et α tend vers $+\infty$.

B.2.1 Calcul de la probabilité sous un processus de Dirichlet

```

logprobaDP <- fonction(a, alpha, b) {
#a est un vecteur dont on calcule la probabilité sous un DP centré sur une loi Gamma de paramètre
#de forme b et de paramètre d'échelle 1.
#alpha est le paramètre de concentration de la DP
n = length(a)
resu = dgamma(a[1], shape = b, scale = 1, log = T)
if(n > 1) {
  for(i in 2 : n) {
    resu = resu + log((alpha/(alpha + i - 1)) * dgamma(a[i], shape = b, scale = 1)
+ sum(a[1 : (i - 1)] == a[i]/(alpha + i - 1))) } }
resu }

```

B.2.2 Cas α quelconque

```

logposter <- fonction(alpha, b, sigma, mu, comptages, sommets, L1, L2, Nsim) {
#calcul de la log-vraisemblance pour des comptages spatiaux dans le rectangle L1*L2
#comptages est le vecteur des comptages dans des unités spatiales rectangulaire
#sommets est une matrice à 4 colonnes (coordonnées du point en bas à gauche, longx, longy)
#permettant de calculer l'intégrale sur chaque unité spatiale rectangulaire

```

```

#a est le vecteur des contributions à l'intensité
#sigma est la bande passante spatiale
#Nsim est le nombre de simulations
resu1 = dgamma(alpha, shape = 9, scale = 0.5, log = T) + dlnorm(b, meanlog = 0, sdlog = 1, log =
T) + dlnorm(mu, meanlog = 1, sdlog = 1, log = T) + duniif(sigma, min = 0, max = 2, log = T)
n = length(comptages)
resu = 0
resu2 = rep(0, Nsim)
for(kin1 : Nsim) {
M = max(1, rpois(1, lambda = mu * L1 * L2))
print(M)
a = rgamma(1, shape = b, scale = 1)
if(M > 1) {
for(i in2 : M) {
u = runif(1); L = (u < alpha/(alpha + i - 1)); j = sample(1 : (i - 1), 1)
a = c(a, L * rgamma(1, shape = b, scale = 1) + (!L) * a[j]) } }
ys = cbind(matrix(L1 * runif(M), ncol = 1), matrix(L2 * runif(M), ncol = 1))
intcums = rep(0, n)
for(i in1 : n) {
intcums[i] = sum(a * (pnorm(sommets[i, 1] + sommets[i, 3] - ys[, 1], sd = sigma)
- pnorm(sommets[i, 1] - ys[, 1], sd = sigma)) * (pnorm(sommets[i, 2] + sommets[i, 4] - ys[, 2], sd =
sigma)
- pnorm(sommets[i, 2] - ys[, 2], sd = sigma))) }
resu = sum(dpois(comptages, intcums, log = T))
resu = resu + dpois(M, mu * L1 * L2, log = T) + logprobaDP(a, alpha, b)
resu2[k] = resu - sum(dpois(comptages, L1 * L2/n, log = T)) }
z = list(resu1, resu2)
z }

```

B.2.3 Cas α tendant vers 0

```

logposter_alpha0 < -function(alpha, b, sigma, mu, comptages, sommets, L1, L2, Nsim) {
#alpha prend la valeur 0
#calcul de la log-vraisemblance pour des comptages spatiaux dans le rectangle L1*L2
#comptages est le vecteur des comptages dans des unités spatiales rectangulaire
#sommets est une matrice à 4 colonnes (coordonnées du point en bas à gauche, longx, longy)
# permettant de calculer l'intégrale sur chaque unité spatiale rectangulaire
#a est le vecteur des contributions à l'intensité
#sigma est la bande passante spatiale
#Nsim est le nombre de simulations
resu1 = dgamma(alpha, shape = 9, scale = 0.5, log = T) + dlnorm(b, meanlog = 0, sdlog = 1, log =
T) + dlnorm(mu, meanlog = 1, sdlog = 1, log = T) + duniif(sigma, min = 0, max = 2, log = T)
n = length(comptages)
resu = 0
resu2 = rep(0, Nsim)
for(k in1 : Nsim) {
M = max(1, rpois(1, lambda = mu * L1 * L2))
print(M)
a = rgamma(1, shape = b, scale = 1); a = rep(a, M)
ys = cbind(matrix(L1 * runif(M), ncol = 1), matrix(L2 * runif(M), ncol = 1))
intcums = rep(0, n)
for(i in1 : n) {
intcums[i] = sum(a * (pnorm(sommets[i, 1] + sommets[i, 3] - ys[, 1], sd = sigma)
- pnorm(sommets[i, 1] - ys[, 1], sd = sigma)) * (pnorm(sommets[i, 2] + sommets[i, 4] - ys[, 2], sd =
sigma)
- pnorm(sommets[i, 2] - ys[, 2], sd = sigma))) }

```

```

resu = sum(dpois(comptages, intcums, log = T))
resu = resu + dpois(M, mu * L1 * L2, log = T) + logprobaDP(a, alpha, b)
resu2[k] = resu - sum(dpois(comptages, L1 * L2/n, log = T)) }
z = list(resu1, resu2)
z }

```

B.2.4 Cas α tendant vers $+\infty$

```

logposter_inf <- function(b, sigma, mu, comptages, sommets, L1, L2, Nsim) {
#calcul de la log-vraisemblance pour des comptages spatiaux dans le rectangle L1*L2
#comptages est le vecteur des comptages dans des unités spatiales rectangulaire
#sommets est une matrice à 4 colonnes (coordonnées du point en bas à gauche, longx, longy)
#permettant de calculer l'intégrale sur chaque unité spatiale rectangulaire
#a est le vecteur des contributions à l'intensité
#sigma est la bande passante spatiale
#Nsim est le nombre de simulations
resu1 = dgamma(alpha, shape = 9, scale = 0.5, log = T) + dlnorm(b, meanlog = 0, sdlog = 1, log =
T) + dlnorm(mu, meanlog = 1, sdlog = 1, log = T) + duniif(sigma, min = 0, max = 2, log = T)
n = length(comptages)
resu = 0
resu2 = rep(0, Nsim)
for(kin1 : Nsim) {
M = max(1, rpois(1, lambda = mu * L1 * L2))
print(M)
a = rgamma(1, shape = b, scale = 1)
ys = cbind(matrix(L1 * runif(M), ncol = 1), matrix(L2 * runif(M), ncol = 1))
intcums = rep(0, n)
for(i in 1 : n) {
intcums[i] = sum(a * (pnorm(sommets[i, 1] + sommets[i, 3] - ys[, 1], sd = sigma)
- pnorm(sommets[i, 1] - ys[, 1], sd = sigma)) * (pnorm(sommets[i, 2] + sommets[i, 4] - ys[, 2], sd =
sigma)
- pnorm(sommets[i, 2] - ys[, 2], sd = sigma))) }
resu = sum(dpois(comptages, intcums, log = T))
resu = resu + dpois(M, mu * L1 * L2, log = T) + dgamma(a, log = T, shape = b, scale = 1)
resu2[k] = resu - sum(dpois(comptages, L1 * L2/n, log = T)) }
z = list(resu1, resu2)
z }

```

B.3 Modèle ETAS

B.3.1 Estimations de paramètres

```

#Calcul de la densité en x de la loi d'Omori modifiée
#scale>0 est le paramètre d'échelle, shape>1 est celui de forme
domori <- function(t, scale = c, shape = p) {
res = ((shape - 1)/scale) * ((1 + t/scale)^(-shape))
res }
#Calcul de la fonction de répartition en x de la loi d'Omori modifiée
#scale>0 est le paramètre d'échelle, shape>1 est celui de forme
pomori <- function(t, scale = c, shape = p) {
res = 1 - (1 + t/scale)^(1 - shape)
res }

```

```

# Calcul de l'intensité d'un ETAS temporel à la date t
intensite <- fonction(t, mu, A, alpha, c, p, obs, Mc, densas == "gamma") {
# mu est l'intensité d'arrière plan
# A est le nombre attendu d'après-chocs générés par un choc de magnitude Mc
# alpha est le taux relatif d'après-chocs générés par choc de magnitude >= Mc
# c et p sont les paramètres de la loi de probabilité de date d'après-choc
# obs est la matrice des observations à 4 colonnes :
# correspondant à : date, longitude, latitude, magnitude
# Mc est la magnitude de complétude
# densas est l'option précisant la loi de probabilité des dates d'après-choc
# On choisira entre les 3 options "gamma", "weibull" et "lognormal"
mag = obs[,4]; time = obs[,5]
lambda = mu
if(densas == "gamma" &t > min(time)) {
lambda = mu + sum(A * exp(alpha * pmax(mag[time < t] - Mc, 0))
* dgamma(t - time[time < t], rate = c, shape = p)) }
if(densas == "weibull" &t > min(time)) {
lambda = mu + sum(A * exp(alpha * pmax(mag[time < t] - Mc, 0))
* dweibull(t - time[time < t], scale = c, shape = p)) }
if(densas == "lognormal" &t > min(time)) {
lambda = mu + sum(A * exp(alpha * pmax(mag[time < t] - Mc, 0))
* dlnorm(t - time[time < t], meanlog = c, sdlog = p)) }
if(densas == "omori" &t > min(time)) {
lambda = mu + sum(A * exp(alpha * pmax(mag[time < t] - Mc, 0))
* domori(t - time[time < t], scale = c, shape = p)) }
lambda }

```

```

intenscum <- fonction(t, mu, A, alpha, c, p, obs, Mc, densas == "gamma") {
# Calcul de l'intensité cumulée d'un ETAS temporel sur l'intervalle [0,t]
# mu est l'intensité d'arrière plan
# A est le nombre attendu d'après-chocs générés par un choc de magnitude Mc
# alpha est le taux relatif d'après-chocs générés par choc de magnitude >= Mc
# c et p sont les paramètres de la loi de probabilité de date d'après-choc
# obs est la matrice des observations a 4 colonnes :
# correspondant a : date, longitude, latitude, magnitude
# Mc est la magnitude de complétude.
# densas est l'option précisant la loi de probabilité des dates d'après-choc
# On choisira entre les 3 options "gamma", "weibull" et "lognormal"
mag = obs[,4]; time = obs[,5]
cumintens = mu * t
if(densas == "gamma" &t > min(time)) {
cumintens = mu * t + sum(A * exp(alpha * pmax(mag[time < t] - Mc, 0))
* pgamma(t - time[time < t], rate = c, shape = p)) }
if(densas == "weibull" &t > min(time)) {
cumintens = mu * t + sum(A * exp(alpha * pmax(mag[time < t] - Mc, 0))
* pweibull(t - time[time < t], scale = c, shape = p)) }
if(densas == "lognormal" &t > min(time)) {
cumintens = mu * t + sum(A * exp(alpha * pmax(mag[time < t] - Mc, 0))
* plnorm(t - time[time < t], meanlog = c, sdlog = p)) }
if(densas == "omori" &t > min(time)) {
cumintens = mu * t + sum(A * exp(alpha * pmax(mag[time < t] - Mc, 0))
* pomori(t - time[time < t], scale = c, shape = p)) }
cumintens }

```

B.3.2 Calcul de vraisemblance

```

lvn_PPgamma <- fonction(theta, obs) {
# Calcul de l'opposé de la log-vraisemblance d'un processus ponctuel ETAS temporel
# doit être en mémoire : Mc
# theta est le vecteur des paramètres car premier argument (requis par "nlminb")
# theta=(mu,A,alpha,c,p)
# En deuxième argument, "obs" est la matrice des observations (5 colonnes).
mu = theta[1]; A = theta[2]; alpha = theta[3]; c = theta[4]; p = theta[5];
time = obs[, 5]
resu = -log(mu)
for(i in 2 : nrow(obs)) {
resu = resu - log(intensite(time[i], mu, A, alpha, c, p, obs, Mc, densas = "gamma")) }
resu + intencum(time[nrow(obs)], mu, A, alpha, c, p, obs, Mc, densas = "gamma") }

```

```

lvn_PPweibull <- fonction(theta, obs) {
# Calcul de l'opposé de la log-vraisemblance d'un processus ponctuel ETAS
# doit être en mémoire : Mc
# theta est le vecteur des paramètres car premier argument (requis par "nlminb")
# theta=(mu,A,alpha,c,p)
# En deuxième argument, "obs" est la matrice des observations (5 colonnes).
mu = theta[1]; A = theta[2]; alpha = theta[3]; c = theta[4]; p = theta[5];
time = obs[, 5]
resu = -log(mu)
for(i in 2 : nrow(obs)) {
resu = resu - log(intensite(time[i], mu, A, alpha, c, p, obs, Mc, densas = "weibull")) }
resu + intencum(time[nrow(obs)], mu, A, alpha, c, p, obs, Mc, densas = "weibull") }

```

```

lvn_PPlognormal <- fonction(theta, obs) {
# Calcul de l'opposé de la log-vraisemblance d'un processus ponctuel ETAS
# doit être en mémoire : Mc
# theta est le vecteur des paramètres car premier argument (requis par "nlminb")
# theta=(mu,A,alpha,c,p)
# En deuxième argument, "obs" est la matrice des observations (5 colonnes).
mu = theta[1]; A = theta[2]; alpha = theta[3]; c = theta[4]; p = theta[5];
time = obs[, 5]
resu = -log(mu)
for(i in 2 : nrow(obs)) {
resu = resu - log(intensite(time[i], mu, A, alpha, c, p, obs, Mc, densas = "lognormal")) }
resu + intencum(time[nrow(obs)], mu, A, alpha, c, p, obs, Mc, densas = "lognormal") }

```

```

lvn_PPomori <- fonction(theta, obs) {
# Calcul de l'opposé de la log-vraisemblance d'un processus ponctuel ETAS temporel
# doit être en mémoire : Mc
# theta est le vecteur des paramètres car premier argument (requis par "nlminb")
# theta=(mu,A,alpha,c,p)
# En deuxième argument, "obs" est la matrice des observations (5 colonnes).
mu = theta[1]; A = theta[2]; alpha = theta[3]; c = theta[4]; p = theta[5];
time = obs[, 5]
resu = -log(mu)

```

```

for(i in 2 : nrow(obs)) {
  resu = resu - log(intensite(time[i], mu, A, alpha, c, p, obs, Mc, densas = "omori")) }
resu + intencum(time[nrow(obs)], mu, A, alpha, c, p, obs, Mc, densas = "omori") }

```

```

  calfc1 <- function(t, mu, A, alpha, c, p, obs, Mc, densas = "gamma") {
    mag = obs[, 4]; time = obs[, 5]
    mag2 = max(mag[time < t] - Mc, 0)
    if(densas == "gamma") {
      resu = sum(mag2 * exp(alpha * mag2) * dgamma(t - time[time < t], rate = c, shape = p)) }
    if(densas == "weibull") {
      resu = sum(mag2 * exp(alpha * mag2) * dweibull(t - time[time < t], scale = c, shape = p)) }
    if(densas == "lognormal") {
      resu = sum(mag2 * exp(alpha * mag2) * dlnorm(t - time[time < t], meanlog = c, sdlog = p)) }
    resu }

```

```

  calfc2 <- function(T, mu, A, alpha, c, p, obs, Mc, densas = "gamma") {
    mag = obs[, 4]; time = obs[, 5]; mag2 = max(mag[time < t] - Mc, 0)
    if(densas == "gamma") {
      resu = sum(mag2 * exp(alpha * mag2) * pgamma(T - time[time < T], rate = c, shape = p)) }
    if(densas == "weibull") {
      resu = sum(mag2 * exp(alpha * mag2) * pweibull(T - time[time < T], scale = c, shape = p)) }
    if(densas == "lognormal") {
      resu = sum(mag2 * exp(alpha * mag2) * plnorm(T - time[time < T], meanlog = c, sdlog = p)) }
    resu }

```

```

  calfc3 <- function(x, y, t, mu, A, alpha, c, p, obs, Mc, densas = "gamma") {
    lati = obs[, 1]; longi = obs[, 2]; depth = obs[, 3]; mag = obs[, 4]; time = obs[, 5]
    mag = max(mag[time < t] - Mc, 0)
    if(densas == "gamma") {
      resu = sum((1/siglongi^3) * ((x - longi[time < t])^2 - siglongi^2) * exp(alpha * mag)
      * dgamma(t - time[time < t], rate = c, shape = p) * dnorm(x - longi[time < t], sd = siglongi)
      * dnorm(y - lati[time < t], sd = siglati)) }
    if(densas == "weibull") {
      resu = sum((1/siglongi^3) * ((x - longi[time < t])^2 - siglongi^2) * exp(alpha * mag)
      * dweibull(t - time[time < t], scale = c, shape = p) * dnorm(x - longi[time < t], sd = siglongi)
      * dnorm(y - lati[time < t], sd = siglati)) }
    if(densas == "lognormal") {
      resu = sum((1/siglongi^3) * ((x - longi[time < t])^2 - siglongi^2) * exp(alpha * mag)
      * dlnorm(t - time[time < t], meanlog = c, sdlog = p) * dnorm(x - longi[time < t], sd = siglongi)
      * dnorm(y - lati[time < t], sd = siglati)) }
    resu }

```

```

  lvn_PP2gamma <- function(theta, obs) {
  # Calcul du gradient de l'opposé de la log-vraisemblance d'un processus ponctuel ETAS temporel
  mu = theta[1]; A = theta[2]; alpha = theta[3]; c = theta[4]; p = theta[5]
  mag = obs[, 4]; time = obs[, 5]
  resu = rep(0, 5)

```

```

for(i in 1 : nrow(obs)) {
resu[1] = resu[1] - 1/intensite(longi[i], lati[i], time[i], mu, A, alpha, c, p, obs, Mc, densas = "gamma")
resu[3] = resu[3] - calfc1(time[i], mu, A, alpha, c, p, obs, Mc, densas = "gamma")
/intensite(time[i], mu, A, alpha, c, p, obs, Mc, densas = "gamma") }
resu[1] = T + resu[1]
resu[2] = (-mu*resu[1]+intenscum(T, mu, A, alpha, c, p, obs, Mc, densas = "gamma")-nrow(obs))/A
resu[3] = A * (resu[3] + calfc2(T, mu, A, alpha, c, p, obs, Mc, densas = "gamma"))
resu[4] = 0; resu[5] = 0 }

```

```

lvn_PP2weibull <- function(theta, obs) {
# Calcul du gradient de l'opposé de la log-vraisemblance d'un processus ponctuel ETAS temporel
mu = theta[1]; A = theta[2]; alpha = theta[3]; c = theta[4]; p = theta[5]
mag = obs[, 4]; time = obs[, 5]
resu = rep(0, 5)
for(i in 1 : nrow(obs)) {
resu[1] = resu[1] - 1/intensite(longi[i], lati[i], time[i], mu, A, alpha, c, p, obs, Mc, densas = "weibull")
resu[3] = resu[3] - calfc1(time[i], mu, A, alpha, c, p, obs, Mc, densas = "weibull")
/intensite(time[i], mu, A, alpha, c, p, obs,
Mc, densas = "weibull")
}
resu[1] = T + resu[1]
resu[2] = (-mu*resu[1]+intenscum(T, mu, A, alpha, c, p, obs, Mc, densas = "weibull")-nrow(obs))/A
resu[3] = A * (resu[3] + calfc2(T, mu, A, alpha, c, p, obs, Mc, densas = "weibull"))
resu[4] = 0
resu[5] = 0 }

```

```

lvn_PP2lognormal <- function(theta, obs) {
# Calcul du gradient de l'opposé de la log-vraisemblance d'un processus ponctuel ETAS temporel
mu = theta[1]; A = theta[2]; alpha = theta[3]; c = theta[4]; p = theta[5]
mag = obs[, 4]; time = obs[, 5]
resu = rep(0, 5)
for(i in 1 : nrow(obs)) {
resu[1] = resu[1] - 1/intensite(longi[i], lati[i], time[i], mu, A, alpha, c, p, obs,
Mc, densas = "lognormal")
resu[3] = resu[3] - calfc1(time[i], mu, A, alpha, c, p, obs, Mc, densas = "lognormal")
/intensite(time[i], mu, A, alpha, c, p, obs, Mc, densas = "lognormal") }
resu[1] = T + resu[1]
resu[2] = (-mu * resu[1] + intenscum(T, mu, A, alpha, c, p, obs, Mc, densas =
"lognormal") - nrow(obs))/A
resu[3] = A * (resu[3] + calfc2(T, mu, A, alpha, c, p, obs, Mc, densas = "lognormal"))
resu[4] = 0; resu[5] = 0 }

```

```

EMV_ETAS <- function(obs, theta0, T, Mc, densas = "gamma") {
# Calcul des estimateurs de maximum de vraisemblance d'un processus ponctuel ETAS temporel
# theta0 est le vecteur initial
# "nlminb" fait partie du package "stats"
# la dernière option ici "obs=" correspond au nom de l'argument de la fonction objectif qui est
# fixe en cours d'optimisation (on l'a appelé en effet "obs" dans "lvn_gamma2").
# A droite de cette option, on a mis "obs" qui est le nom du vecteur des observations
resu = nlminb(start = theta0, lower = rep(0, 5), objective = lvn_PPgamma,

```

```
obs = obs)
resu }
```

```
SC_gamma <- fonction(theta, obs) {
# Calcul de la somme des carrés de résidus d'un processus ponctuel ETAS temporel
# (cas Gamma fonction déclenchante)
# afin d'estimer les paramètres avec la fonction "nlminb"
mu = theta[1]; A = theta[2]; alpha = theta[3]; c = theta[4]; p = theta[5];
time = obs[,5]
resu = NULL; n = nrow(obs)
for(i in 1 : n) {
resu = c(resu, intenscum(time[i], mu, A, alpha, c, p, obs, Mc, densas = "gamma")) }
resu = resu - c(0, resu[-n]); resu = sum((1 - resu)^2)
resu }
```

```
SC_weibull <- fonction(theta, obs) {
# Calcul de la somme des carrés de résidus d'un processus ponctuel ETAS temporel
# (cas Weibull fonction déclenchante)
# afin d'estimer les paramètres avec la fonction "nlminb"
mu = theta[1]; A = theta[2]; alpha = theta[3]; c = theta[4]; p = theta[5];
time = obs[,5]
resu = NULL; n = nrow(obs)
for(i in 1 : n) {
resu = c(resu, intenscum(time[i], mu, A, alpha, c, p, obs, Mc, densas = "weibull")) }
resu = resu - c(0, resu[-n])
resu = sum((1 - resu)^2)
resu }
```

```
SC_lognormal <- fonction(theta, obs) {
# Calcul de la somme des carrés de résidus d'un processus ponctuel
# ETAS temporel
# (cas log-normal fonction déclenchante)
# afin d'estimer les paramètres avec la fonction "nlminb"
mu = theta[1]; A = theta[2]; alpha = theta[3]; c = theta[4]; p = theta[5];
time = obs[,5]
resu = NULL; n = nrow(obs)
for(i in 1 : n) {
resu = c(resu, intenscum(time[i], mu, A, alpha, c, p, obs, Mc, densas = "lognormal")) }
resu = resu - c(0, resu[-n]); resu = sum((1 - resu)^2)
resu }
```

```
traceresidu <- fonction(mu, A, alpha, c, p, obs, densas = "gamma") {
# Analyse résiduelle d'un processus ponctuel ETAS temporel
time = obs[,5]
resu = NULL
n = nrow(obs)
if(densas == "gamma") {
for(i in 1 : n) {
```


160 ANNEXE B. PROGRAMMES INFORMATIQUES COMPLÉMENTAIRES SOUS R

```

resu = c(resu, intenscum(time[i], mu, A, alpha, c, p, obs, Mc, "gamma")) } }
if(densas == "weibull") {
for(i in 1 : n) {
resu = c(resu, intenscum(time[i], mu, A, alpha, c, p, obs, Mc, "weibull")) } }
if(densas == "lognormal") {
for(i in 1 : n) {
resu = c(resu, intenscum(time[i], mu, A, alpha, c, p, obs, Mc, "lognormal")) } }
if(densas == "omori") {
for(i in 1 : n) {
resu = c(resu, intenscum(time[i], mu, A, alpha, c, p, obs, Mc, "omori")) } }
plot(1 : n, 1 : n, type = "l", xlab = "Event number", ylab =
"Residual process value", ylim = c(0, max(n, resu)), cex.lab = 1.5)
lines(1 : n, resu, type = "l", col = "red")
resu = sort(resu)
‡ car dans le fichier obs, les dates ne sont pas forcément ordonnées croissantes
U = resu - c(0, resu[-n])
lala = hist(U, breaks = c(seq(0, 2.75, .25), max(3, U)), plot = F)
ef fobs = lala$counts
prob = pexp(seq(0.25, 2.75, .25)); prob = c(prob, 1)
effesp = n * (prob - c(0, prob[-length(prob)]))
statsq = sum(((ef fobs - effesp)^2)/effesp)
p = 1 - pchisq(statsq, df = 11)
‡ test du khi-2 d'ajustement a la loi exponentielle de paramètre 1
text(n*5/7, n*2/7, paste("Chi-squaredteststatistic = ", as.character(round(statsq, 3))), cex = 1.3)
text(n*5/7, n/7, paste("p - value = ", as.character(round(p, 3))), cex = 1.3)
list(residus = resu, stat = statsq, pvalue = p) }

```

Bibliographie

- [1] J. ALBERT, *Bayesian Computation with R*, Springer, New York, 2nd ed., 2009.
- [2] Q. AN, C. WANG, I. SHETERVEV, E. WANG, D. DUNSON, AND L. CARIN, *Hierarchical kernel stick-breaking process for multi-task image analysis*, in Proceedings of the 25th International Conference on Machine Learning, 2008.
- [3] P. K. ANDERSEN, O. BORGAN, R. D. GILL, AND N. KEIDING, *Statistical Models Based on Counting Processes*, Springer series in statistics, Springer, 1993.
- [4] A. BADDELEY, R. TURNER, J. MØLLER, AND M. HAZELTON, *Residual analysis for spatial point processes*, Journal of the Royal Statistical Society. Series B (Statistical Methodology), 67 (2005), pp. 617–666.
- [5] V. BENEŠ, K. BODLAK, J. MØLLER, AND R. WAAGEPETERSEN, *A case study on point process modelling in disease mapping*, Image Analysis & Stereology, 24 (2005), pp. 159–168.
- [6] V. BENEŠ AND B. Z. FRCALOVÁ, *Modelling and simulation of a neurophysiological experiment by spatio-temporal point processes.*, Image Analysis & Stereology, 27 (2008), pp. 47–52.
- [7] M. BENGBOU-VALÉRIUS, S. BAZIN, D. BERTIL, F. BEAUDUCEL, AND A. BOSSON, *Cdsa : a new seismological data center for the french lesser antilles*, Seismological Research Letters, 79 (2008), pp. 90–102.
- [8] D. BLACKWELL AND J. MACQUEEN, *Ferguson distributions via polya urn schemes*, Annals of Statistics, (1973), pp. 353–355.
- [9] P. BRÉMAUD, *Point Processes and Queues, Martingales Dynamics*, Springer Verlag, 1981.

- [10] P. BRÉMAUD AND L. MASSOULIÉ, *Hawkes branching point processes without ancestors*, Journal of Applied Probability, 38 (2001), pp. 122–135.
- [11] A. BRIX, *Generalized gamma measures and shot-noise cox processes*, Advanced in Applied Probability, 31 (1999), pp. 929–953.
- [12] A. BRIX AND J. CHADŒUF, *Spatio-temporal modeling of weeds and shotnoise cox processes*, Biometrical Journal, 44 (2002), pp. 83–99.
- [13] A. BRIX AND P. DIGGLE, *Spatio-temporal prediction for log-gaussian cox processes*, Journal of the Royal Statistical Society B, 63 (2001), pp. 823–841.
- [14] F. CHIMARD, *Mélanges de processus ponctuels spatio-temporels et approche bayésienne semi-paramétrique*, PhD thesis, Université des Antilles et de la Guyane, 2010.
- [15] F. CHIMARD, L. VALMY, AND J. VAILLANT, *Statistique d'un processus de cox associé à un processus de dirichlet*, Journal de la Société Statistique de France, accepté (2012).
- [16] A. D. CLIFF AND J. K. ORD, *Spatial processes : models and applications*, Pion Ltd, London, UK, 1981.
- [17] C. COMAS, J. MATEU, AND P. DELICADO, *On tree intensity estimation for forest inventories : Some statistical issues*, Biometrical Journal, 53 (2011), pp. 994–1010.
- [18] P. COWPERTWAITA, P. O'CONNELLA, A. METCALFEB, AND J. MAWDSLEY, *Stochastic point process modelling of rainfall. 1. single-site fitting and validation*, Journal of Hydrology, 175 (1996), pp. 17–46.
- [19] D. R. COX AND V. ISHAM, *Point processes*, Chapman & Hall, London, 1980. Monographs on Applied Probability and Statistics.
- [20] N. CRESSIE, *Statistics for Spatial Data*, John Wiley and Sons Inc., New York, 1993.
- [21] D. DALEY AND D. VERE-JONES, *An introduction to the theory of point processes*, Springer-Verlag, New York, 1988.
- [22] D. DALEY AND D. VERE-JONES, *An Introduction to the Theory of Point Processes*, Springer, 2nd Ed., New York, 2003.
- [23] P. DIACONIS AND D. FREEDMAN, *De Finetti's theorem for Markov chains*, The Annals of Probability, 8 (1980), pp. 115–130.

- [24] P. DIGGLE, *Statistical Analysis of Spatial Point Patterns*, Academic Press, London, 1983.
- [25] P. DIGGLE, *Spatio-temporal point processes : Methods and applications*, working paper 78, Department of Biostatistics, Johns Hopkins University, 2005.
- [26] P. DIGGLE, B. ROWLINGSON, AND T. SU, *Point process methodology for on-line spatio-temporal disease surveillance*, *Environmetrics*, 16 (2005), pp. 423–434.
- [27] P. J. DIGGLE AND R. K. MILNE, *Negative binomial quadrat counts and point processes*, *Scandinavian Journal of Statistics. Theory and Applications*, 10 (1983), pp. 257–267.
- [28] J. DROESBEKE, J. FINE, AND G. SAPORTA, *Méthodes Bayésiennes en Statistique*, Technip, 2002.
- [29] J. A. DUAN, M. GUINDANI, AND A. E. GELFAND, *Generalized spatial dirichlet process models*, *Biometrika*, 94 (2007), pp. 809–825.
- [30] D. B. DUNSON AND J.-H. PARK, *Kernel stick-breaking processes*, *Biometrika*, 95 (2008), pp. 307–323.
- [31] R. S. EHLERS AND S. P. BROOKS, *Adaptative proposal construction for reversible jumps mcmc*, *Scandinavian Journal*, 35 (2008), pp. 677–690.
- [32] T. FERGUSON, *A bayesian analysis of some nonparametric problems*, *The Annals of Statistics*, 1 (1973), pp. 209–230.
- [33] ———, *Prior distributions on spaces of probability measures*, *The Annals of Statistics*, 2 (1974), pp. 615–629.
- [34] R. M. FEWSTER, J. L. LAAKE, AND S. T. BUCKLAND, *Line transect sampling in small and large regions*, *Biometrics*, 61 (2005), pp. 856–859.
- [35] M. GARCÍA-DA-ROSA, C. OLIVARO, M. P. CERDEIRAS, AND A. VÁZQUEZ, *Antibacterial activity of *ibicella lutea* glycosides*, *Revista Latinoamericana de Química*, 38 (2010), pp. 98–102.
- [36] C. GOURIÉROUX AND A. MONFORT, *Séries temporelles et modèles dynamiques, Second Edition*, Collection Économie et Statistiques Avancées. Série École Nationale de la Statistique et de l'Administration et du Centre d'Études des Programmes Économiques, Economica, 1995.
- [37] P. GREEN, *Reversible jump markov chain monte carlo computation and bayesian model determination*, *Biometrika*, 82 (1995), pp. 711–732.

- [38] P. GREEN AND S. RICHARDSON, *Hidden markov models and disease mapping*, American Statistical Association, 97 (2002), pp. 1055–1070.
- [39] D. HARTE, *Ptprocess : An r package for modelling marked point processes indexed by time*, Journal of Statistical Software, 35 (2010), pp. 1–32.
- [40] A. G. HAWKES, *Spectra of some self-exciting and mutually exciting point processes*, Biometrika, 58 (1971), pp. 83–90.
- [41] L. HOLDEN, S. SANNAN, AND H. BUNGUM, *A stochastic marked point process model for earthquakes*, Natural Hazards and Earth System Sciences, 3 (2003), pp. 95–101.
- [42] M. HUTCHINGS, *Standing crop and pattern in pure stands of mercurialis perennis and rubus fruticosus in mixed deciduous woodland*, Oikos, 31 (1978), pp. 351–357.
- [43] H. ISHWARAN AND L. F. JAMES, *Gibbs sampling methods for stick-breaking priors*, American Statistical Association, 96 (2001), pp. 161–173.
- [44] J. L. JENSEN AND N. V. PETERSEN, *Asymptotic normality of the maximum likelihood estimator in state space models*, Annals of Statistics, 27 (1999), pp. 514–535.
- [45] C. JI, D. MERL, T. B. KEPLER, AND M. WEST, *Spatial mixture modelling for unobserved point processes : Examples in immunofluorescence histology*, Bayesian Analysis, 4 (2009), pp. 297–316.
- [46] M. JORDAN, *Hierarchical models, nested models and completely random measures*, tech. report, University of California, Berkeley, 2010.
- [47] Y. KAGAN, *A probabilistic description of the seismic regime*, Izv. Acad. Sci. U.S.S.R., Phys. Solid Earth, (1973), pp. 213–219.
- [48] ———, *Short-term properties of earthquake catalogs and models of earthquake source*, Bulletin of the Seismological Society of America, 94 (2004), pp. 1207–1228.
- [49] A. F. KARR, *Point Processes and Their Statistical Inference*, Probability : pure and applied ; 2, New-York, 2nd edition ed., 1991.
- [50] L. KNOPOFF, *The magnitude distribution of declustered earthquakes in southern california*, Proceedings of the National Academy of Sciences of the United States of America, 97 (2000), pp. 11880–11884.

- [51] A. KOTTAS, J. A. DUAN, AND A. E. GELFAND, *Modeling disease incidence data with spatial and spatio-temporal dirichlet process mixtures*, Biometrical Journal, 49 (2007), pp. 1–14.
- [52] A. KOTTAS AND B. SANSÓ, *Bayesian mixture modeling for spatial poisson process intensities, with applications to extreme value analysis*, Statistical Planning and Inference, 137 (2007), pp. 3151–3163.
- [53] Y. KUTOYANTS, *Statistical Inference for Spatial Poisson Processes*, vol. 134 of Lecture Notes in Statistics, Springer-Verlag, New York, 1998.
- [54] P. LANSKY AND J. VAILLANT, *Stochastic model of the overdispersion in the place cell discharge*, Biosystems, 58 (2000), pp. 27–32.
- [55] A. LAWSON, *Bayesian disease mapping : hierarchical modeling in spatial epidemiology*, Interdisciplinary Statistics, CRC Press, 2009.
- [56] B. MATERN, *Doubly Stochastic Poisson Processes in the Plane*, vol. 1 of Statistical Ecology, Patil, G., penn state univ press ed., 1971, pp. 195–213.
- [57] G. J. MELVILLE AND A. H. WELSH, *Line transect sampling in small regions*, Biometrics, 57 (2001), pp. 1130–1137.
- [58] J. MØLLER, *Shot noise cox processes*, Advanced in Applied Probability, 35 (2003), pp. 614–640.
- [59] J. MØLLER AND C. DIAZ-AVALOS, *Structured spatio-temporal shot-noise cox point process models, with a view to modelling forest fires*, Scandinavian Journal of Statistics, 37 (2010), pp. 2–25.
- [60] J. MØLLER AND F. P. SCHENBERG, *Thinning spatial point processes into poisson processes*, Advanced in Applied Probability, 2 (2010), pp. 347–358.
- [61] J. MØLLER AND G. TORRISI, *Generalised shot noise cox processes*, Advanced in Applied Probability, 37 (2005), pp. 48–74.
- [62] J. MØLLER AND R. WAAGEPETERSEN, *Modern statistics for spatial point processes*, Scandinavian Journal of Statistics, 34 (2007), pp. 643–684.
- [63] R. NEAL, *Markov chain sampling methods for dirichlet process mixture models*, Journal of Computational and Graphical Statistics, 9 (2000), pp. 249–265.
- [64] Y. OGATA, *The asymptotic behaviour of maximum likelihood estimators for stationary point processes*, Annals of the Institute of Statistical Mathematics, 30 (1978), pp. 243–261.

- [65] —, *Statistical models for earthquake occurrences and residual analysis for point processes*, Journal of American Statistical Association, 83 (1988), pp. 9–27.
- [66] —, *Space-time point-process models for earthquake occurrences*, Annals of the Institute of Statistical Mathematics, 50 (1998), pp. 379–402.
- [67] Y. OGATA, *Seismicity analysis through point process modeling : a review*, Pure and Applied Geophysics, 155 (1999), pp. 471–507.
- [68] Y. OGATA, *Significant improvements of the space-time etas model for forecasting of accurate baseline seismicity*, Earth Planets Space, 63 (2011), pp. 217–229.
- [69] Y. OGATA, K. KATSURA, AND M. TANEMURA, *Modelling heterogeneous space-time occurrences of earthquakes and its residual analysis*, Journal of the Royal Statistical Society. Series C (Applied Statistics), 52 (2003), pp. 499–509.
- [70] Y. OGATA AND J. ZHUANG, *Space-time etas models and an improved extension*, Tectonophysics, 413 (2006), pp. 13–23.
- [71] D. P. PADILLA, A. GONZÁLEZ-CASTRO, AND M. NOGALES, *Significance and extent of secondary seed dispersal by predatory birds on oceanic islands : the case of the canary archipelago*, Journal of Ecology, 100 (2012), pp. 416–427.
- [72] L. PANINSKI, *Maximum likelihood estimation of cascade point-process neural encoding models*, Computation in Neural Systems, 15 (2004), pp. 243–262.
- [73] G. PERRY, B. MILLER, AND N. ENRIGHT, *A comparison of methods for the statistical analysis of spatial point patterns in plant ecology*, Plant Ecology, 187 (2006), pp. 59–82.
- [74] L. PURI AND P. TUAN, *Maximum likelihood estimation for stationary point processes*, Proceedings of the National Academy of Sciences, 83 (1986), pp. 541–545.
- [75] V. RAO AND Y. W. TEH, *Spatial normalized gamma processes*, in Advances in Neural Information Processing Systems, 2009.
- [76] S. RATHBUN, *Asymptotic properties of the maximum likelihood estimator for spatio-temporal point processes*, Journal of Statistical Planning and Inference, 51 (1996), pp. 55–74.
- [77] B. RIPLEY, *Spatial Statistics*, John Wiley and Sons Inc., New York, 1981.
- [78] C. ROBERT, *Méthodes de Monte Carlo par chaînes de Markov*, Economica, Paris, 1996.
- [79] —, *Le choix bayésien : Principes et pratique*, Springer, Paris, 2006.

- [80] A. ROGERS, *Statistical Analysis of Spatial Dispersion the quadrat method*, Pion Limited., London, 1974.
- [81] R. ROSYCHUK, C. HUSTON, AND N. PRASAD, *Spatial event cluster detection using a compound poisson distribution*, *Biometrics*, 62 (2006), pp. 465–470.
- [82] F. P. SCHENBERG, *Multidimensional residual analysis of point process models for earthquake occurrences*, *Journal of the American Statistical Association*, 98 (2003), pp. 789–795.
- [83] J. SETHURAMAN, *A constructive definition of dirichlet priors*, *Statistica Sinica*, 4 (1994), pp. 639–650.
- [84] D. J. SPIEGELHALTER, N. G. BEST, B. P. CARLIN, AND A. VAN DER LINDE, *Bayesian measures of model complexity and fit*, *Journal of the Royal Statistical Society, Series B*, 4 (2002).
- [85] D. STOYAN AND A. PENTTINEN, *Recent applications of point process methods in forestry statistics*, *Statistical Science*, 15 (2000), pp. 61–78.
- [86] J. SULLIVAN, P. WILLIAMS, AND S. TIMMINS, *Effects of Pinus Radiata Plantations on Environmental Weed Invasion Into Adjacent Native Forest Reserves*, Wellington, N.-Z., 2006.
- [87] M. TANNER, *Tools for Statistical Inference : Methods for the Exploration of Posterior Distributions and Likelihood Functions*, New-York, 1996.
- [88] L. TAYLOR, *Assessing and interpreting the spatial distributions of insect populations*, *Annual Review of Entomology*, 29 (1984), pp. 321–357.
- [89] Y. W. TEH, M. I. JORDAN, M. J. BEAL, AND D. M. BLEI, *Hierarchical dirichlet processes*, *Journal of the American Statistical Association*, 101 (2006), pp. 1566–1581.
- [90] W. TRUCCOLO, U. EDEN, J. DONOGHUE, AND E. BROWN, *A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects*, *Journal of Neurophysiology*, 93 (2005), pp. 1074–1089.
- [91] J. VAILLANT, *Negative binomial distributions of individuals and spatio-temporal cox processes*, *Scandinavian Journal of Statistics*, 18 (1991), pp. 235–248.
- [92] —, *échantillonnage et étude statistique de populations en milieu hétérogène*, *Revue de Statistique Appliquée*, 40 (1992), pp. 15–26.

- [93] —, *Testing hypotheses about occupancy rates for clustered populations*, Biometrical Journal, 34 (1992), pp. 735–744.
- [94] —, *Tests of occupancy rates for clustered populations : Case where the cluster sizes are known*, Biometrical Journal, 36 (1994), pp. 33–42.
- [95] J. VAILLANT AND S. DERRIDJ, *Estimation of european corn borer egg masses density by sampling of runs*, Researches on Population Ecology, 31 (1989).
- [96] J. VAILLANT, G. PUGGIONI, L. WALLER, AND J.-H. DAUGROIS, *A spatio-temporal analysis of the spread of sugar cane yellow leaf virus*, Journal of Time Series Analysis, 32 (2011), pp. 396–406.
- [97] L. A. VALMY, *Statistique des processus ponctuels - applications à la sismologie*, in 4^{ème} Rencontre des Jeunes Statisticiens, Aussois (France), Septembre 2011, SFDS.
- [98] L. A. VALMY AND J. VAILLANT, *Étude comparative de modèles stochastiques pour l'activité sismique*, in 19^{ème} Conférence Géologique de la Caraïbe, Gosier (Guadeloupe), Mars 2011, CGC.
- [99] —, *Étude statistique d'événements sismiques caribéens*, in 39^{ème} Congrès de la Société Statistique du Canada, Wolfville (Canada), Juin 2011, SSC.
- [100] —, *Statistique asymptotique de processus auto-excitatifs spatio-temporels*, in 43^{èmes} Journées de la Statistique, Tunis (Tunisie), Mai 2011, SFDS.
- [101] —, *Bayesian inference on a cox process associated with a dirichlet process*, en cours d'achèvement (2012).
- [102] —, *Échantillonnage spatial et inférence sur un processus de dirichlet caché*, in 44^{èmes} Journées de la Statistique, Bruxelles (Belgique), Mai 2012, SFDS.
- [103] —, *Statistical inference for point processes associated with a dirichlet process*, in 40^{ème} Congrès de la Société Statistique du Canada, Guelph (Canada), Juin 2012, SSC.
- [104] —, *Statistical models in seismology : Lesser antilles arc case*, Bulletin de la Société Géologique de France, 184 (1) (2013).
- [105] T. VAN STIPHOUT, J. ZHUANG, AND D. MARSAN, *Seismicity declustering*, Community Online Resource for Statistical Seismicity Analysis, (2012).
- [106] D. VERE-JONES, *Stochastic models for earthquake occurrence (with discussion)*, Journal of the Royal Statistical Society Series B, 32 (1970), pp. 1–62.

- [107] —, *Forecasting earthquakes and earthquake risk*, International Journal of Forecasting, 11 (1995), pp. 503–538.
- [108] D. VERE-JONES, R. ROBINSON, AND W. YANG, *Remarks on the accelerated moment release model : Problems of model formulation, simulation and estimation*, Geophysical Journal International, 144 (2001), pp. 517–531.
- [109] L. WALLER, *Bayesian thinking in spatial statistics*, Handbook of Statistics Handbook of Statistics, 25 (2005), pp. 589–622.
- [110] R. WOLPERT AND K. ICKSTADT, *Poisson/gamma random field models for spatial statistics*, Biometrika, 85 (1998), pp. 251–267.
- [111] T. ZHANG AND S. C. KOU, *Nonparametric inference of doubly stochastic poisson process data via the kernel method*, Journal of Applied Statistics, 4 (2010), pp. 1913–1941.
- [112] J. ZHUANG, *Second-order residual analysis of spatio-temporal point processes and applications in model evaluation*, Journal of the Royal Statistical Society, 64 (2006), pp. 635–653.
- [113] J. ZHUANG, A. CHRISTOPHERSEN, M. SAVAGE, D. VERE-JONES, Y. OGATA, AND D. JACKSON, *Differences between spontaneous and triggered earthquakes : Their influences on foreshock probabilities*, Journal of Geophysical Research, 113 (2008), pp. 1–9.
- [114] J. ZHUANG, Y. OGATA, AND D. VERE-JONES, *Diagnostic Analysis of Space-time Branching Processes for Earthquakes*, A. Baddeley, P. Gregori, J. Mateu, R. Stoica and D. Stoyan, 2005, ch. Chapitre 15, pp. pages 275–290.

