

Modèles hybrides de sources pour la quantification de l'aléa sismique

Livret de rapports

Ibrahim SEYDI

October 20, 2025

Contents

1 Ier Quadrimestre – Mai, Juin, Juillet, Aout 2025	3
2 IIème Quadrimestre – Septembre, Octobre, Novembre, Décembre 2025	18

Ier Quadrimestre

CHAPTER 1

1er Quadrimestre – Mai, Juin, Juillet, Aout 2025

Rapport de la semaine du 12/05/2025

Travaux/Tâches effectués

- Récupération ordinateur au service IT
- Réunion Adel
- Progression inscription ADUM
- Préparation et pitch thèse + séminaire technique risque sismique
- Création Github privé (+ permet également d'archiver les articles, accessible par tous)
- Lecture article PSHA classique
- Début de rédaction de résumé critique
- Recherche bibliographique sur les méthodes zoneless

Problèmes rencontrés / Questions en suspens

- Copie de la convention de cotutelle SIGNÉE
- Copie de l'attestation d'assurance responsabilité civile
- Connexion à distance avec ordi EDF
- Config Authenticator

Idées surgies

- Piste axe de recherche sur "amélioration des modèles de récurrence", tentative de modélisation bayésienne (priors informatifs, ...) ; Etude de la loi de GR sur la magnitude (le paramètre qui semble supplanter les autres)
- Piste méthodes zoneless (cf recherche biblio)

Lectures et ressources

- Titre : *A probabilistic seismic hazard map for the metropolitan France*
Type : Article
Lien ou référence : Dans Github
Notes : Article de référence sur l'approche standard PSHA ; Gloria Senfaute co-autrice

Étapes suivantes

- Finaliser inscription ADUM
- Finir analyse critique écrit
- Fixer axe de recherche à suivre
- Remplir docs pour forfait jour et envoyer à Bruno pour signature
- Continuer à suivre travail de Adel sur processus de Hawkes
- Se faire une culture sismo (docs Gloria)
- Formaliser mathématiquement (chap 2/3 du papier de l'approche classique PSHA),
- Deadline(s) à venir : ...

Notes libres

- Manque de productivité vendredi, un peu de fatigue, quelques lacunes sur le sommeil (à améliorer)
- ...

Rapport de la semaine du 19/05/2025

Travaux/Tâches effectués

- Réunion Adel
- Réunion avec Pierre ARROUCAU et Guillaume DANIEL pour discuter de l'orientation de la thèse et incorporation de la thèse dans SIGMA 3/WP1 (livrables à fournir et réunions avec eux tous les 2/3 mois)
- Accès à SoftID
- Premier jour à APT, réunion avec Sophie et Julien : formalisme du problème + mini cours BNP
- Formalisme modèle
- Test du KDE sur un exemple simple (carré unitaire), sans recherche d'optimisation
- Recherche bibliographique (dossier avec différents articles, thèses, cours)
- Lecture document écrit par étudiant (cf mail Sophie 'Ressources')
- Pseudo code Stick Breacking

Problèmes rencontrés / Questions en suspens

- Pas de connexion dans bureau APT donc partage compliqué + pas étudiant donc pas de compte Eduroam

Idées à explorer

- Pistes pour faire de l'inférence : Chinese Restaurant Process + Gibbs ; VI ?
- Eduroam coté EDF

Lectures et ressources

- Titre : *Density Estimation Using Nonparametric Bayesian Methods*
Type : Article
Lien ou référence : https://openscholarship.wustl.edu/art_sci_etds/1507/
Notes : Rapport réalisé par un étudiant
- Titre : *Bayesian Nonparametrics*
Type : Livre
Lien ou référence : ...
Notes : Prêter par Sophie

Étapes suivantes

- Améliorer le KDE (choix de la fenêtre)
- Implémenter Stick Breacking (soit proposer une évaluation du choix de K, soit fixer un seuil pour la taille du bâton, à partir duquel on sort de la boucle (utiliser plutôt un while))
- Rédaction fiche Bayésien NP
- Clarifier et poser sur papier, pistes pour faire de l'inférence
- Lire biblio
- Finir analyse papier PSHA

Notes pour moi-même

- Avancer progressivement, ne pas hésiter à commencer doucement avec cas jouet ou exemples simples (à l'instar de TS)
 - Echanger avec collaborateurs dès que doute ou besoin (à l'instar de TS)
 - Jeudi 29 férié
-

Rapport de la semaine du 26/05/2025

Travaux/Tâches effectués

- Re-travail du KDE spatial (gaussiennes dans les quatre coins + CV pour avoir fenêtre optimal)
- Lecture articles
- Début rédaction fiche Bayésien NP
- Ecriture Stick-Breacking avec seuil τ et début implémentation DP par Stick-Breacking

Problèmes rencontrés / Questions en suspens

- Eclaircissement sur le pk de l'étirement du KDE quand les localisations des points étaient générées aléatoirement ?
- Cross Valid avec OpenTURNS ?
- Pb avec fonction Beta de OpenTURNS

Idées à explorer

- Modèles de diffusion/SBM (cf article envoyé par Merlin, essayer de voir dans quelle mesure c'est exploitable (idée pour un peu plus tard encore))

Lectures et ressources

- Titre : *Bayesian nonparametrics and the probabilistic approach to modelling*
Type : Article
Auteur : Zoubin Ghahramani
Notes : Auteur a pas mal bossé sur Bayésien non paramétrique
- Titre : *dirichletprocess: An R Package for Fitting Complex Bayesian Nonparametric Models*
Type : Article
Lien ou référence : ...
Auteur : Gordon J. Ross
Notes : Auteur qui a pas mal étudié sur Hawkes et DP
- Titre : *Dirichlet Process*
Type : Cours
Auteur :
Notes :

- Titre : *Lecture 23: Bayesian Nonparametrics: Dirichlet Processes*
Type : Article
Auteur : Eric P. Xing
Notes : Auteur a pas mal bossé sur DP
- Titre : *Nonparametric Bayesian methods: a gentle introduction and overview*
Type : Article
Auteur : Steven N. MacEachern
Notes : Reprendre partie 2.4 et 2.6 (et p-e 3) pour fiche BNP
- Titre : *Non-parametric Bayesian Methods*
Type : Cours
Auteur : Zoubin Ghahramani
Notes : Excellent !

Étapes suivantes

- Finir implémentation Stick-Breaking (avec NIW)
- Finir Fiche Bayésien NP (partie)
- Lire articles sur Inférence et commencer à faire
- Rédiger proprement inférence dans Fiche Bayésien NP
- Régler ptits pbs sur KDE

Notes à moi-même

- Venue si possible à Agro mercredi

Rapport des semaines 02/06/2025, 09/06/2025 et 16/06/2025

Travaux/Tâches effectués

- Régler ptits pbs sur KDE/SB
- Avancement fiche Bayésien NP
- Lecture articles
- Travail sur première approche pour faire de l'inférence (Neal 2000)
- Travail avec Sophie sur comment intégrer un prior informatif (Formalisme d'une piste)
- Séminaire journeé utilisateurs OT
- Code DPMM avec prior informatif
- Avancement analyse critique du papier standard PSHA
- Rdv avec Sismologues Pierre ARROUCAU et Guillaume DANIEL
- Étude des caractéristiques d'un prior NIW

Problèmes rencontrés / Questions en suspens

- Lien entre densité d'une distribution d'évènements et intensité d'un processus Poissonien ? (cf mail Sophie)

Lectures et ressources

- Titre : *Markov Chain Sampling Methods for Dirichlet Process Mixture Models*
Type : Article
Auteur : Radford M. Neal
Notes : Article fondamental pour inférence (Gibbs)
- Titre : *Nonparametric Bayesian methods : a gentle introduction and overview*
Type : Article
Auteur : Steven N. MacEachern
Notes : pour fiche BNP

Étapes suivantes

- Finir analyse critique sur papier approche standard
- Faire apparaître des lignes de niveaux sur simu
- Subplots pour comparaisons entre lignes de niveaux et f_0 (catalogue)
- Calcul distances entre \tilde{f}_0 et f et entre f et f_0 (en fonction de α , et des hyperparamètres du NIW)
- Simu avec prior non informatif et comparer en fonction de α
- Colormap/Heatmap de f , f_0 , \tilde{f}_0
- Extension de la piste actuelle à un cadre plus complexe
- Deadline à venir : 18/06/25 - Réunion avec P.ARROUCAU et G.DANIEL
- Deadline à venir : 27/06/25 - Abstract travaux pour Luis et Sismologues

Notes pour moi-même

- Inclure Anne Dutfoy dans boucle des personnes intéressées par la thèse
-

Rapport de la semaine du 23/06/2025

Travaux/Tâches effectués

- Faire apparaître des lignes de niveaux sur simu DPMM
- Subplots pour comparaisons entre lignes de niveaux et f_0 (catalogue)
- Calcul distances entre \tilde{f}_0 et f et entre f et f_0 (en fonction de α)
- Colormap/Heatmap de f , f_0 , \tilde{f}_0
- Simu avec prior non informatif et comparer en fonction de α
- Biblio pour vérif que notre piste n'existe pas déjà dans la nature

Problèmes rencontrés / Questions en suspens / Idées à explorer

- Mauvais domaine de définition de la densité du DPMM. Deux pistes de solutions :
 1. Réécriture de la densité avec constante de normalisation (qui est compliquée). Constante peut être approchée ? Estimateur de MOnte-Carlo ? Constante de normalisation dépend des paramètres il me semble donc si utilisation d'une méthode itérative ça risque de poser pbs ? Méthodes SMC peut aider ici (permet de bosser avec des densités non normalisées, et peut donner des estimations de la constante de normalisation) ? Méthode pour tirer dans un domaine bornée ?
 2. Considérer la densité a posteriori traditionnelle et contraindre les paramètres du mélange (centres et covariances des gaussiennes) pour qu'ils soient dans une zone bornée. Dans inférence : pénaliser la vraisemblance du modèle pour être dans support d'intérêt ? ; Sorte de rejet lorsque l'on est pas dans support d'intérêt ?
2/ -> garde intacte la forme du posterior (conjugaison éventuelle)

Étapes suivantes

- Mettre au propre le code
- Modifier code pour avoir DPMM avec renormalisation des gaussiennes tronquées
- Regarder les marginales
- Finir analyse critique sur papier approche standard
- Écrire lien entre densité et intensité d'un processus (cf mail Sophie)
- Finir calcul de la conjugaison de la NIW avec Normale

- Extension de la piste actuelle à un cadre plus complexe
- Deadline à venir : 02/07/2025 - RDV avec toute la team rapprochée

Notes pour moi-même

- Biblio pour vérif que notre piste n'existe pas déjà dans la nature
-

Rapport des semaines 30/06/2025, 07/07/2025

Travaux/Tâches effectués

- Mise au propre du code
- Réunion avec la team rapprochée : Mise en clair des problématiques, enjeux, objectifs, pistes exploratoires
- Restructuration du code/github
- Avancement sur analyse papier PSHA standard
- Analyse du code de Julien
- Upgrade au plan premium d'Overleaf
- Création d'une carte stratégique mentale

Problèmes rencontrés / Questions en suspens / Idées à explorer

- Renormalisation du DPMM informatif sur $[0, 2]^2$, puis suppression de cette renormalisation (on s'en fiche pour l'instant)

Étapes suivantes

- Finir restructuration codes (faire des classes, ...)
- Faire Docstring à l'image de codes phebus
- Réaliser tout ce que je fais avec des zones irrégulières (from code phebus)

Notes pour moi-même

- Continuer à faire de la biblio pour vérif que piste n'existe pas dans la nature

Rapport des semaines 14/07/2025, 21/07/2025

Travaux/Tâches effectués

- Fin restructuration codes
- EM pour approximation du mélange d'uniformes par mélange de gaussiennes
- Écriture de la méthodo réalisée jusqu'ici
- Études contre model Merlin, discussions perspectives thèses avec Sophie
- Docstrings codes

Problèmes rencontrés / Questions en suspens / Idées à explorer

- ADUM (ENCORE ET TJR)

Étapes suivantes

- Appliquer méthode avec des zones irrégulières (from code phebus)
- Commencer inférence (Gibbs et VI)
- Lire autour de HDP et DDPM
- Comparaison KDE, DPMM standard, DPMM informatif
- Lire autour de GP pour estimation log densité (Logistic Gaussian Process/Cox)
- Finir analyse papier standard PSHA
- Modifier et finir Carte mentale
- Deadline(s) à venir : CA - 14/08/2025

Notes pour moi-même

- Continuer à faire de la biblio pour vérif que piste n'existe pas dans la nature

Rapport des semaines 28/07/2025, 04/08/2025, 08/09/2025

Travaux/Tâches effectués

- Application DPMM informatif sur exemple avec des zones irrégulières
- Re-travail du code
- Lecture autour de HDP
- Points avec Adel, lecture rapport de stage, commentaires sur rapport
- Avancement Fiche Bayésien NP
- Travail sur inférence du modèle informatif, ébauche -> inspiration de l'algo 2 de Neal 2000
- Avancement ADUM
- Fin de stage d'Adel : rapport + slides
- Formation OpenTurns (08/09/10 septembre)

Problèmes rencontrés / Questions en suspens / Idées à explorer

- ADUM (ENCORE ET TJR)
- Questionner Adel sur différents méthodes de simu des Hawkes (existence d'une méthode de simu horizontale (conceptuellement parlant) ?)

Lectures et ressources

- Titre : *Bayesian nonparametrics and the probabilistic approach to modelling*
Auteur : Zoubin Ghahramani
Type : Article
Lien ou référence : Dans Github
Notes : Cool (à revoir pour relever points intéressants pour la suite)
- Titre : *Hierarchical Dirichlet Processes*
Type : Article
Lien ou référence : <https://people.eecs.berkeley.edu/~jordan/papers/hierarchical-dp.pdf>
Notes : Papier fondateur HDP
- Titre : *Hierarchical Dirichlet Processes*
Type : Article
Lien ou référence : https://www.stat.berkeley.edu/~aldous/206-Exch/Papers/hierarchical_dirichlet.pdf
Notes : Papier technique (approche développée) fondateur HDP

- Titre : *Markov Chain Sampling Methods for Dirichlet Process Mixture Models*
Auteur : Radford M. Neal
Type : Article
Lien ou référence : Dans Github
Notes : Semble être référence imtemporelle pour Gibbs pour DPMM ; Pillier sur lequel repose Gibbs pour DPMM informatif

Étapes suivantes

- Mettre de l'ordre dans l'organisation du github
- Implémentation inférence Gibbs pour DPMM informatif
- Modifier distance L^2 -> Utiliser méthode de calcul d'intégrale
- Lire autour du choix du paramètre de concentration
- Comparaison KDE, DPMM standard, DPMM informatif sur données simulées
- Lecture autour de DDPM
- Deadline(s) à venir : ... - ...

Notes pour moi-même

- Continuer à faire de la biblio pour vérif que piste n'existe pas dans la nature
 - Relire approche de discréétisation pour comprendre comment erreur a pu arriver
 - Suivi cours de Vincent Guigne, Deep Learning avec Pytorch
-

IIème Quadrimestre

CHAPTER 2

IIème Quadrimestre – Septembre, Octobre, Novembre, Décembre 2025

Rapport des semaines 15/09/2025 et 22/09/2025

Travaux/Tâches effectués

- Lecture article sur inférence du paramètre de concentration
- Début implémentation Gibbs pour DPMM informatif
- Ré-organisation du github
- Finalisation inscription 2024/2025 et avancement inscription 2025/2026
- Rdv avec ED
- Rédaction maths de échantillonnage des paramètres mu et Sigma dans cas mélange NIW
- Fin implémentation Gibbs
- Soutenance stage Adel
- Rdv Pierre et Guillaume Sismo -> Possibilité d'avoir un catalogue suffisamment précis pour ETAS ; Pas de poids donné par experts dans mélange d'uniformes ; Séparation données historiques/instrumentales pas spécialement pertinent d'un point de vue sismo ; Le alpha (de par sa fonction) dans DPMM semble être élément qui suscite le plus l'intérêt des sismos

Problèmes rencontrés / Questions en suspens / Idées à explorer

- Information des poids du mélange d'uniformes sans les données ?
- Piste du GP, envisageable sans poids
- Écrire densité cible comme une combinaison convexe :

$$f(x) = (1 - \alpha) f_{prior}(x) + \alpha f_{res}$$

où f_{prior} est mélange d'uniformes à poids inconnus et f_{res} est résidu lisse tq :

$$f_{res}(x) = \frac{\exp(g(x))}{\int_{\mathcal{D}} \exp(g(x)) dx}$$

avec $g \sim \mathcal{GP}$.

Lectures et ressources

- Titre : *Prior selection for the precision parameter of Dirichlet Process Mixtures*

Type : Article

Lien ou référence : <https://arxiv.org/pdf/2502.00864.pdf> ; Sur Github

Notes : Choix du α dans DPMM, paramètre de concentration ; Lecture à poursuivre ; Fait l'étagage des approches existantes avant proposition d'une nouvelle approche

Étapes suivantes

- Corrections code si besoin, faire classe Prior NIW
- Diagnostics de convergence (ou de non convergence) -> cf mail Merlin
- Lire autour de VI pour DPMM
- Application DPMM standard à données réelles
- Lire autour de GP pour estimation log densité (Logistic Gaussian Process/Cox)
- Finir rédaction sur HDP (CRF)
- Remettre nez dans KDE -> Mettre au propre si besoin + optim estimation
- Deadline à venir : Rdv Véronique Gerauld - 30/09/2025 à 16h40
- Événement à venir : Soutenance Antoine Von Biesbroeck - 01/10/2025 à Polytechnique
- Deadline à venir : Rdv Pierre Barbillon - 02/10/2025 à 9h
- Événement à venir : ETICS 2025 - Du 06 au 10 octobre -> Pas de point la semaine pro

Notes pour moi-même

- Faire rappel à Julien concernant le livre de Gelman
 - Suivre demande de justificatif de formation OpenTurns
 - Continuer à faire de la biblio pour vérif que piste n'existe pas dans la nature
 - Dès que certif de sco 2025/2026, envoyer à Morgane + aller voir Thierry SLV pour remboursement CVEC
-

Rapport des semaines 29/09/2025 et 06/10/2025

Travaux/Tâches effectués

- Lecture article, rapport Adel
- Rdv Pierre Barbillon / CSI - 02/10/2025
- Rédaction formalisme du modèle ETAS
- Lecture GP-ETAS
- ETICS 2025 - Toute la semaine du 06 oct

Problèmes rencontrés / Questions en suspens / Idées à explorer

- ...

Lectures et ressources

- Titre : *Processus de Hawkes - Modélisation, Simulation et Inférence*
Auteur : Loubna Ben Allal
Type : Rapport Master
Lien ou référence : https://inria.hal.science/hal-02947341/file/Processus_de_hawkes1.pdf ; Sur Github
Notes : Vraiment intéressant, bonne base pour débuter
- Titre : *The Effect of Seismic Sequences in Probabilistic Seismic Hazard Analysis*
Auteur : Nevena Šipčić et al.
Type : Article
Lien ou référence : Sur Github
Notes : Papier référence ETAS pour stage d'Adel
- Titre : *Processus de Hawkes pour l'aléa sismique*
Auteur : Adel EL HABAZI
Type : Rapport de stage
Lien ou référence : Sur Overleaf
- Titre : *Prévoir les séismes : une vanité à l'ère de l'IA ?*
Auteurs : Clara Carlier et Lucas De Lara
Type : Projet de ML
Lien ou référence : https://www.imo.universite-paris-saclay.fr/~yannig.goude/Materials/ProjetMLF/Rapports2020/Rapport_CARLIER_DELARA.pdf
Notes : Quelques éléments pour ETAS

- Titre : *ETAS: An R Package for Fitting the Space-Time ETAS Model to Earthquake Data*
Auteur : Abdollah Jalilian
Type : Article
Lien ou référence : <https://www.jstatsoft.org/article/view/v088c01> ; Sur Github
Notes : Quelques éléments pour spécification modèle ETAS, mais à lire complètement pcq pas mal de choses intéressantes, notamment la notion de 'Stochastic declustering'
- Titre : *GP-ETAS: semiparametric Bayesian inference for the spatio-temporal epidemic type aftershock sequence model*
Auteurs : Christian Molkenthin et al.
Type : Article
Lien ou référence : Sur Github
Notes : J'aime bien

Étapes suivantes

- Lire autour de GP pour estimation log densité (Logistic Gaussian Process/Cox)
- Finir lecture papier GP-ETAS et noter grandes lignes
- Suivre biblio ETAS (arxiv) (notamment regarder qui cite Molkenthin)
- Simuler GP avec carte en dessous (avec Jospeh)
- Demander Molkenthin codes GP-ETAS et faire tourner GP-ETAS avec fonction d'activation à 0
- Deadline à venir :
- Événement à venir : Journée des doctorants de l'EDMH - 17/10/2025 -> Préparer présentation gong

Notes pour moi-même

- Faire rappel à Julien concernant le livre de Gelman
 - Dès que certif de sco 2025/2026, envoyer à Morgane + aller voir Thierry SLV pour remboursement CVEC
-

Rapport de la semaine 13/10/2025

Travaux/Tâches effectués

- Synthèse GP-ETAS
- Premier cours de DeepL
- Début de code GP
- Début lecture article sur l'augmentation de données par processus de Poisson latentes
- Journée de rentrée des doctorants de l'EDMH (journée du vendredi)
- Correspondance avec Molkenthin

Problèmes rencontrés / Questions en suspens / Idées à explorer

- Idée pour plus tard : incorporation du zonage dans le noyau spatial de la fonction d'auto-excitation du ETAS

Lectures et ressources

- Titre : *GP-ETAS: semiparametric Bayesian inference for the spatio-temporal epidemic type aftershock sequence model*
Auteurs : Christian Molkenthin et al.
Type : Article
Lien ou référence : Sur Github
Notes : J'aime bien
- Titre : *Efficient Bayesian Inference of Sigmoidal Gaussian Cox Processes*
Auteurs : Christian Donner et Manfred Opper
Type : Article
Lien ou référence : Sur Github
Notes : Base pour les stratégies de data augmentation par processus de Poisson et par Pólya-Gamma

Étapes suivantes

- Faire fiche marathon doctorants (à faire assez vite)
- Bosser codes DL
- Code GP (avec prior informatif)

- Finir fiche synthèse GP-ETAS
- Développement mathématique parties floues
- Finir lecture article Donner et Opper 2018
- Reprendre code ETAS d'Adel
- Configurer Zotero et partager
-
- Deadline à venir : ... - ...
- Événement à venir : ... - ...

Notes pour moi-même

- Faire rappel à Julien concernant le livre de Gelman
 - Dès que certif de sco 2025/2026, envoyer à Morgane + aller voir Thierry SLV pour remboursement CVEC
 - Mettre à jour mdp edf
-