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

Bayesian modeling of earthquakes

Ibrahim SEYDI

30 juin 2025

Restricted case

Spatial data without aftershocks or foreshocks

Let observed seismic events be :

$$x_1, \ldots, x_n \in \mathbb{R}^2.$$

We assume that each observation $x_i = (x_i^{(1)}, x_i^{(2)})$ is a spatial coordinate drawn independently from a density f which is unknown.

We set that :

$$\begin{aligned} x_i \mid \theta_i &\sim \mathcal{N}(x_i \mid \mu_i, \Sigma_i), \quad i = 1, \ldots, n \\ \theta_i = (\mu_i, \Sigma_i) \mid G &\sim G \\ G \mid \alpha, G_0 &\sim \mathrm{DP}(\alpha, G_0) \\ G_0 \mid m_0, \Lambda_0, \psi_0, \nu_0 &= \mathcal{NIW}(m_0, \Lambda_0, \psi_0, \nu_0) \end{aligned}$$

where :

- $\mathcal{N}(\cdot \mid \mu_i, \Sigma_i)$ is a bivariate normal distribution
- $\theta_i = (\mu_i, \Sigma_i) \in \mathbb{R}^2 \times \mathcal{S}_+^2$, with $\mathcal{S}_+^2$ the set of symmetric positive definite 2×2 covariance matrices
- G is a random probability measure over the parameter space θ, drawn from a **Dirichlet process**
- G_0 is the **base measure** following a normal-inverse-Wishart on θ_i.

Liste de papiers pour méthodes zoneless (à filtrer) :

- Woessner et al (2015) The 2013 European Seismic hazard model : key components and results.
- Petersen MD, Harmsen SC, Jaiswal KS, Rukstales KS, Luco N, Haller KM, Mueller CS, Shumway AM (2018) Seismic hazard, risk, and design for south America.
- Helmstetter A, Werner MJ (2012) Adaptive spatiotemporal smoothing of seismicity for long-term earthquake forecasts in California.
- Woo G (1996) Kernel estimation methods for seismic hazard area source modeling.
- S. Molina, C. Lindholm, H. Bungum (2001) Probabilistic seismic hazard analysis : zoning free versus zoning methodology.
- Chethanamba Kempanna Ramanna, G. Dodagoudar (2012) Probabilistic seismic hazard analysis using kernel density estimation technique for Chennai, India.
- S. Lasocki (2021) Kernel Density Estimation in Seismology
- M. Danese, M. Lazzari, B. Murgante (2008) Kernel Density Estimation Methods for a Geostatistical Approach in Seismic Risk Analysis : The Case Study of Potenza Hilltop Town (Southern Italy)
- C. Stock, Euan Smith (2002) Adaptive Kernel Estimation and Continuous Probability Representation of Historical Earthquake Catalogs
- C. Stock, Euan Smith (2002) Comparison of Seismicity Models Generated by Different Kernel Estimations
- G. Estévez-Pérez, H. L. Cimadevila, A. Quintela-del-Río (2002) Nonparametric analysis of the time structure of seismicity in a geographic region
- M. Crespo, F. Martínez, J. Martí (2014) Seismic hazard of the Iberian Peninsula : evaluation with kernel functions
- Francis Tong, Stanisław Lasocki, Beata Orlecka-Sikora (2025) Non-parametric kernel density estimation of magnitude distribution for the analysis of seismic hazard posed by anthropogenic seismicity
- Karaburun, A. ; Demirci, A. (2016) Spatio-temporal cluster analysis of the earthquake epicenters in Turkey and its surrounding area between 1900 and 2014
- Kernel Density Estimation for the Interpretation of Seismic Big Data in Tectonics Using QGIS : The Türkiye–Syria Earthquakes (2023)
-

Simulation de processus de Dirichlet

Simulation par Stick-Breacking

Générer une (approximation) de la densité sous la forme :

$$f(x) = \sum_{k=1}^{K} w_k \, \delta_{\theta_k}(x)$$

où : $\theta_k \sim G_0$.

Tronquer le modèle à un nb K fixé de composantes du mélange.

Input

- Nombre de composantes K
- Param de concentration $\alpha > 0$

Étapes :

1. Initialisation : Créer liste vide **poids** = **[]** et le reste du bâton : $r = 1.0$; Créer liste vide $\boldsymbol{\theta}$
2. Pour $k = 1$ à $K - 1$:
   - Tirer $v_k \sim \mathrm{Beta}(1, \alpha)$
   - Calculer $w_k = v_k \cdot r$
   - Ajouter w_k à **poids**
   - Mise à jour du baton $r = r \cdot (1 - v_k)$
   - Simuler $\theta_k \sim G_0$
   - Ajouter θ_k à $\boldsymbol{\theta}$
3. Ajouter $w_K = r$ à **poids**
4. Simuler $\theta_K \sim G_0$ et l'ajouter à $\boldsymbol{\theta}$

Output : Approximation d'un DP avec : $\mathcal{P} = \sum_{k=1}^{K} w_k \delta_{\theta_k}$

Simulation par Stick-Breaking (avec seuil τ)

Générer une approximation de la densité sous la forme :

$$f(x) = \sum_{k=1}^{\infty} w_k \, \delta_{\theta_k}(x)$$

où $\theta_k \sim G_0$, et les poids sont générés par le procédé de Stick-Breaking.

Input :

- Param de concentration $\alpha > 0$
- Seuil $\tau > 0$

Étapes :

1. Initialisation :
   - Liste vide des poids : `poids = [ ]`
   - Liste vide des paramètres : $\boldsymbol{\theta} = []$
   - Reste du bâton : $r \leftarrow 1.0$
   - Indice : $k \leftarrow 1$
2. Tant que $r > \tau$, faire :
   - Tirer $v_k \sim \text{Beta}(1, \alpha)$
   - Calculer $w_k = v_k \cdot r$
   - Ajouter w_k à `poids`
   - Mettre à jour : $r \leftarrow r \cdot (1 - v_k)$
   - Simuler $\theta_k \sim G_0$
   - Ajouter θ_k à $\boldsymbol{\theta}$
   - Incrément $k \leftarrow k + 1$

Output : Approximation d'un DP avec :

$$\mathcal{P} = \sum_{k=1}^{K} w_k \, \delta_{\theta_k}(x), \quad \text{où } K \text{ est déterminé en fonction de } \tau$$

Intégrer l'information des zonages sismotectoniques sous forme de prior informatif

Nous avons accès à un nombre n de positions de séismes sur un lieu Ω :

$$x_1, \ldots, x_n \quad \sim \quad f \quad (f \text{ densité})$$

où $x_i = (x_i^{(1)}, x_i^{(2)})$ pour tout $i \in [1, n]$.
Notre approche : <u>Estimation bayésienne non paramétrique de f</u>

$$\begin{cases} f(x) \ = \ \int \ \mathcal{N}(\mu, \Sigma) \ \ dG(\mu, \Sigma) \\ \qquad G \ \sim \ \text{DP}(\alpha, G_0) \end{cases}$$

Autre formulation :

$$\begin{aligned} f(x) &= \sum_{k=1}^{\infty} w_k \, \mathcal{N}(\mu_k, \Sigma_k) \\ (w_k)_k &\sim \mathrm{SB}(\alpha) \\ (\mu_k, \Sigma_k)_k &\sim G_0 = \mathcal{N}(\mu_k \mid \mu_0, \frac{\Sigma_k}{\lambda_0}) \, \mathcal{IW}(\Sigma_k \mid \psi_0, \nu_0) \end{aligned}$$

But : Intégrer prior informatif du zonage sismotectonique

On note f_0 la densité de la distribution du zonage. on a :

$$f_0(x) \;=\; \frac{\sum_{j=1}^{J} w_{0,j} \; \mathbb{1}_{S_{0,j}}(x)}{\sum_{j=1}^{J} w_{0,j} \; A_{0,j}}$$

où $S_{0,1}, \ldots, S_{0,J}$ est une partition de Ω et représente les zones d'un zonage sismotectonique et chaque $A_{0,j}$ est la surface de $S_{0,j}$.

Une idée serait d'utiliser des gaussiennes pour approcher les découpes du zonage avec $\mu_{0,j}$ des centroides des zones $S_{0.j}$ et $\Sigma_{0,j}$ des diamètres d'ellipses (?). On aurait :

$$\tilde{f}_0(x) \;=\; \frac{\sum_{j=1}^{J} w_{0,j} \; \mathcal{N}(\mu_{0,j}, \Sigma_{0,j})}{\sum_{j=1}^{J} w_{0,j} \; A_{0,j}}$$

Ainsi, on aurait la mesure de base a priori informative suivante :

$$G_0^{\text{inf}}(\cdot) \;=\; \sum_{j=1}^{J} w_{0,j} \; \mathcal{N}(\cdot \mid \mu_{0,j}, \frac{\Sigma_j}{\lambda_0}) \; \mathcal{IW}(\cdot \mid \Sigma_{0,j}, \nu_0)$$

Première étape : Évaluation de la qualité de la version informative

---

On cherche à construire une densité spatiale sur la carte de France Ω à partir d'un zonage sismo, c'est-à-dire :

$$\int_{\text{France}} f_0(x, y) \, dx \, dy \quad = \quad 1$$

où $f_0(x, y)$ est constante sur chaque zone $S_{0,j}$.

Soit $\{S_{0,1}, \ldots, S_{0,J}\}$ un zonage sismotectonique de la France Ω, tel que $\Omega = \bigcup_{j=1}^{J} S_{0,j}$, avec $S_{0,j} \cap S_{0,i} = \emptyset$ si $i \neq j$. Chaque $S_{0,j}$ a :

- une surface $A_{0,j} = \mathrm{Surf}(S_{0,j})$
- une poids associé $w_{0,j} \geq 0$, avec : $\sum_{j=1}^{J} w_{0,j} \ = \ 1$

On peut définir :

$$f_0(x, y) = \sum_{j=1}^{J} \frac{w_{0,j}}{A_{0,j}} \cdot \mathbb{1}_{S_{0,j}}(x, y)$$

On a bien une densité spatiale sur la France.
Si on a une catégorisation de chaque zone selon un niveau de sismicité, on peut attribuer un facteur $\lambda_{0,j}$ proportionnel à chaque caté pour obtenir un poids normalisé comme suit :

$$w_{0,j} = \frac{\lambda_{0,j}}{\sum_j \lambda_{0,j}}$$

On a :

$$\begin{aligned}
\int_\Omega f_0(x, y)\ dx\, dy \ &= \ \int_\Omega \sum_{j=1}^{J} \frac{w_{0,j}}{A_{0,j}} \cdot \mathbb{1}_{S_{0,j}}(x, y)\ dx\, dy \ = \ \sum_{j=1}^{J} \frac{w_{0,j}}{A_{0,j}} \cdot \int_\Omega \mathbb{1}_{S_{0,j}}(x, y)\ dx\, dy \\
&= \ \sum_{j=1}^{J} \frac{w_{0,j}}{A_{0,j}} \cdot A_{0,j} \ = \ \sum_{j=1}^{J} w_{0,j} \ = \ 1
\end{aligned}$$

Donc $f_0(x, y)$ est bien une densité sur Ω.

La loi associée $\mathbb{P}_{X,Y} : \mathcal{B}(\mathbb{R}^2) \to [0, 1]$ associée à cette densité serait donnée par :

$$\begin{aligned}
\mathbb{P}_{X,Y}(B) \ &= \ \mathbb{P}((X, Y) \in B) \ = \ \int_B f_0(x, y)\ dx\, dy \\
&= \ \sum_{j=1}^{J} \frac{w_{0,j}}{A_{0,j}} \cdot \int_{B \cap S_{0,j}} dx\, dy \qquad \text{pour tout borélien } B
\end{aligned}$$

La loi Normale-Inverse Wishart est une loi jointe sur : la moyenne d'une loi normale multivariée $\boldsymbol{\mu}$ et la matrice de covariance à cette loi normale multivariée $\boldsymbol{\Sigma}$. Cette loi est caractérisée par quatre hyperparamètres :

- $\boldsymbol{\mu_0}$ (vect de dim d) : la moyenne prior sur $\boldsymbol{\mu}$
- λ_0 (scalaire positif) : un facteur d'échelle sur la précision de la moyenne

- $\mathbf{\Psi_0}$ (matrice d x d, sym et def pos) : un paramètre d'échelle pour la matrice de covariance
- ν_0 (degré de liberté $>$ d - 1) : un paramètre qui contrôle la concentration de la loi Inverse Wishart sur $\mathbf{\Sigma}$

La NIW est donnée ainsi :

$$\mathbf{\Sigma} \sim \mathcal{IW}(\Psi_0, \nu_0), \quad \boldsymbol{\mu} \mid \mathbf{\Sigma} \sim \mathcal{N}(\boldsymbol{\mu_0}, \frac{\mathbf{\Sigma}}{\lambda_0})$$

où $\mathcal{IW}$ est la loi inverse Wishart et $\boldsymbol{\mu}$ suit une normale multivariée avec covariance $\mathbf{\Sigma}/\lambda_0$.

La deensité de l'Inverse Wishart est :

$$f(\mathbf{\Sigma} \mid \mathbf{\Psi_0}, \nu_0) = \frac{|\mathbf{\Psi_0}|^{\frac{\nu_0}{2}}}{2^{\frac{\nu_0 d}{2}} \Gamma_d\left(\frac{\nu_0}{2}\right)} |\mathbf{\Sigma}|^{-\frac{\nu_0+d+1}{2}} \exp\left(-\frac{1}{2}\mathrm{tr}(\mathbf{\Psi_0}\mathbf{\Sigma}^{-1})\right)$$

avec $\Gamma_d(\cdot)$ la fonction gamma multivariée dim d, $|\cdot|$ le déterminant et tr la trace.

La densité de la loi normale conditionnelleest :

$$f\left(\boldsymbol{\mu} \mid \boldsymbol{\mu_0}, \frac{\mathbf{\Sigma}}{\lambda_0}\right) = \frac{1}{(2\pi)^{\frac{d}{2}} |\mathbf{\Sigma}/\lambda_0|^{\frac{1}{2}}} \exp\left(-\frac{\lambda_0}{2}(\boldsymbol{\mu} - \boldsymbol{\mu_0})^T \mathbf{\Sigma}^{-1} (\boldsymbol{\mu} - \boldsymbol{\mu_0})\right)$$

Donc la densité NIW est :

$$\boxed{\begin{aligned} f(\boldsymbol{\mu}, \mathbf{\Sigma}) &= f(\mathbf{\Sigma} \mid \mathbf{\Psi_0}, \nu_0) \cdot f\left(\boldsymbol{\mu} \mid \boldsymbol{\mu_0}, \frac{\mathbf{\Sigma}}{\lambda_0}\right) \\ &= \frac{|\mathbf{\Psi_0}|^{\frac{\nu_0}{2}} \lambda_0^{\frac{d}{2}}}{(2\pi)^{\frac{d}{2}} 2^{\frac{\nu_0 d}{2}} \Gamma_d\left(\frac{\nu_0}{2}\right)} |\mathbf{\Sigma}|^{-\frac{\nu_0+d+2}{2}} \\ &\quad \times \exp\left(-\frac{1}{2}\mathrm{tr}(\mathbf{\Psi_0}\mathbf{\Sigma}^{-1}) - \frac{\lambda_0}{2}(\boldsymbol{\mu} - \boldsymbol{\mu_0})^T \mathbf{\Sigma}^{-1} (\boldsymbol{\mu} - \boldsymbol{\mu_0})\right) \end{aligned}}$$

$\Gamma_d(\cdot)$ est la fonction gamma multivariée dim d, définie par :

$$\Gamma_d(a) = \pi^{\frac{d(d-1)}{4}} \prod_{i=1}^{d} \Gamma\left(a + \frac{1-i}{2}\right)$$

Paramètre	**Effet/Interprétation**
$\boldsymbol{\mu_0}$	• Moyenne de la loi de $\boldsymbol{\mu}$ • Plus λ_0 est grand, plus $\boldsymbol{\mu}$ est concentré autour de $\boldsymbol{\mu_0}$ • Plus λ_0 est faible, $\boldsymbol{\mu}$ s'écarte de $\boldsymbol{\mu_0}$
λ_0	• C'est un facteur d'échelle sur la variance de $\boldsymbol{\mu}$ • Contrôle l'incertitude a priori sur la moyenne $\boldsymbol{\mu}$ • $\lambda_0 \to 0$ signifie une incertitude infinie sur $\boldsymbol{\mu}$
$\boldsymbol{\Psi_0}$	• Contrôle la taille et l'orientation moyennes des matrices de covariance $\boldsymbol{\Sigma}$ • Plus les valeurs propres de $\boldsymbol{\Psi_0}$ sont grandes, plus les réalisations de $\boldsymbol{\Sigma}$ sont grandes (variances plus larges) et/ou des corrélations plus marquées
ν_0	• Contrôle la concentration de la loi de $\boldsymbol{\Sigma}$ • Doit être strictement supérieur à $d-1$ pour que la moyenne existe • Plus ν_0 est grand, plus $\boldsymbol{\Sigma}$ est concentré autour de sa moyenne $\boldsymbol{\Psi_0}/(\nu_0 - d - 1)$ • Si ν_0 est faible (proche de d), la dispersion des matrices $\boldsymbol{\Sigma}$ est grande (forte incertitude)

On cherche à approximer la distance L^2 entre deux densités f et g de $\mathbb{R}^2$, définie par :

$$\|f - g\|_{L^2} = \left(\int_\Omega (f(x,y) - g(x,y))^2 \, dx \, dy \right)^{1/2}$$

où Ω est domaine d'étude.

On peut utiliser l'approche des somme discrète pour approximer l'intégrale. On utilise la formule d'approximation :

$$\int_\Omega (f(x,y) - g(x,y))^2 \, dx \, dy \quad \approx \quad \sum_{i,j} (f(x_i,y_j) - g(x_i,y_j))^2 \cdot \Delta x \cdot \Delta y$$

où Δx, Δy sont les pas de grille.

On cherche à approximer la densité moyenne d'un DPMM avec un prior informatif. L'approximation se fait par moyennage Monte Carlo sur N densités générées. Chaque composante k du mélange est tirée sous une loi normale-inverse-Wishart :

$$(\mu_k, \Sigma_k) \sim \mathrm{NIW}(\mu_0^{(j)}, \lambda_0, \Psi_0, \nu_0) \quad \text{où} \quad \mu_0^{(j)} \in \{[0.5, 0.5], [1.5, 0.5], [0.5, 1.5], [1.5, 1.5]\},$$

$$\Psi_0 = \begin{pmatrix} 0.26 & 0 \\ 0 & 0.26 \end{pmatrix}, \quad \lambda_0 = 50.0, \quad \text{et} \quad \nu_0 = 4.$$

Les poids du mélange sont générés via stick-breaking avec paramètre de concentration α et seuil de troncature τ :

$$v_k \sim \mathrm{Beta}(1, \alpha), \qquad w_k = v_k \prod_{i=1}^{k-1} (1 - v_i)$$

$$\text{On arrête quand } \prod_{i=1}^{k} (1 - v_i) < \tau$$

Les poids sont normalisés ensuite : $\sum_{k=1}^{K} w_k = 1$

Chaque densité générée s'écrit comme un mélange de normales :

$$f^{(i)}(\cdot) = \sum_{k=1}^{K^{(i)}} w_k^{(i)} \cdot \mathcal{N}(\cdot \mid \mu_k^{(i)}, \Sigma_k^{(i)})$$

avec les composantes $(\mu_k^{(i)}, \Sigma_k^{(i)}) \sim \mathrm{NIW}(\mu_0^{(j)}, \lambda_0, \Psi_0, \nu_0)$

Enfin, on calcule la moyenne empirique des densités :

$$\bar{f}_N(\cdot) = \frac{1}{N} \sum_{i=1}^{N} f^{(i)}(\cdot) = \frac{1}{N} \sum_{i=1}^{N} \sum_{k=1}^{K^{(i)}} w_k^{(i)} \cdot \mathcal{N}(\cdot \mid \mu_k^{(i)}, \Sigma_k^{(i)})$$