

# Construction de lois a priori objectives pour l'estimation bayésienne de courbes de fragilité sismique

Antoine Van Biesbroeck,  
Clément Gauchy, Cyril Feau, Josselin Garnier

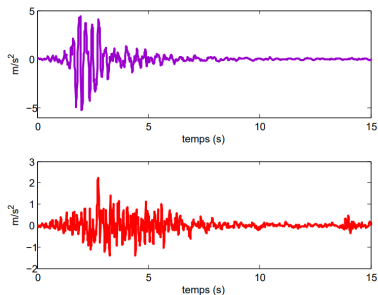
Commissariat à l'Energie Atomique et aux Energies Alternatives

10 juin 2022



- 1 Introduction
  - Indicateur de mesure sismique et de rupture
  - Courbes de fragilité
- 2 Information mutuelle et prior objectif
  - Prior de référence non asymptotique
  - Prior de référence asymptotique
- 3 Applications numériques
- 4 Caractère propre de Jeffreys
- 5 Conclusion

# Indicateur sismique et rupture

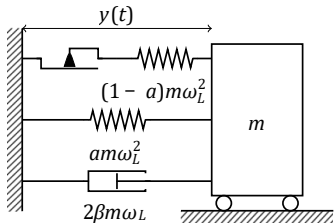


Signaux d'accélération sismique

Catégoriser le signal  $s$  par un indicateur sismique :

- 1  $PGA = \max_{t \in [0, T]} |s(t)|;$
- 2  $PGV = \max_{t \in [0, T]} \left| \int_0^t s(\tau) d\tau \right|;$
- 3  $E = \int_0^T s^2(\tau) d\tau;$
- 4  $SA = 4\pi^2 f_L^2 \max_{t \in [0, T]} |\tilde{y}(t)|.$

Un modèle mécanique traduit ensuite la réponse de la structure au signal  $s$ .



$$y''(t) + 2\beta 2\pi f_L y'(t) + f^{nl}(t) = -s(t)$$

modèle mécanique

On définit:

- 1 Le déplacement maximal :  $DM = \max_{t \in [0, T]} |y(t)|$ .
- 2 La rupture :  $\mathbf{1}_{DM > C}$

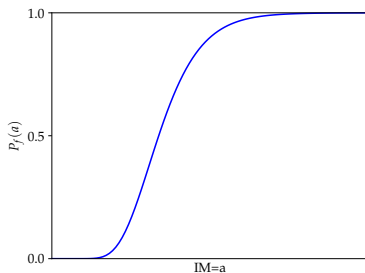


Figure: Exemple de courbe de fragilité sismique

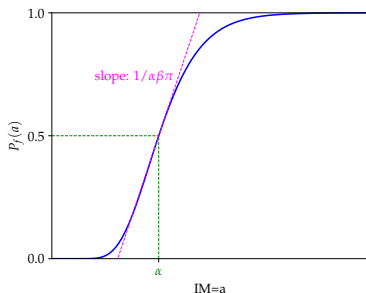


Figure: Exemple de courbe de fragilité sismique

Modèle Log-Normal:

$$\begin{aligned} \log DM = c \log IM + d + \sigma \mathcal{N}(0, 1) &\longrightarrow P_f(a) = \mathbb{P}(DM > c | IM = a) \\ &= \Phi\left(\frac{\ln a - \ln \alpha}{\beta}\right) \end{aligned}$$

- L'étude revient à la recherche du paramètre  $\theta = (\alpha, \beta)$ .

Base de données  $((a_1, z_1), \dots, (a_k, z_k))$   $a = IM$  et  $z \in \{0, 1\}$

- Méthode classique : Max de vraisemblance

$$\hat{\theta}_{MLE} = \arg \max_{\theta} l_k(z|a, \theta) = \arg \max_{\theta} \prod_{i=1}^k p(z_i|a_i, \theta)$$

→ Obtention de différents échantillons de  $\hat{\theta}_{MLE}$  par bootstrap.

→ Pb : Irrégulier avec peu de données

Base de données  $((a_1, z_1), \dots, (a_k, z_k))$   $a = IM$  et  $z \in \{0, 1\}$

- Méthode classique : Max de vraisemblance

$$\hat{\theta}_{MLE} = \arg \max_{\theta} l_k(z|a, \theta) = \arg \max_{\theta} \prod_{i=1}^k p(z_i|a_i, \theta)$$

→ Obtention de différents échantillons de  $\hat{\theta}_{MLE}$  par bootstrap.

→ Pb : Irrégulier avec peu de données

- Point de vue Bayésien:  $\theta$  est une variable aléatoire de loi  $\pi$

La loi a posteriori :  $p(\theta|z, a) \propto \pi(\theta) l_k(z|a, \theta)$   
 $\propto \pi(\theta) \prod_{i=1}^k p(z_i|a_i, \theta)$

→ Obtention de différents échantillons de  $\theta$  par simulation de  $p$

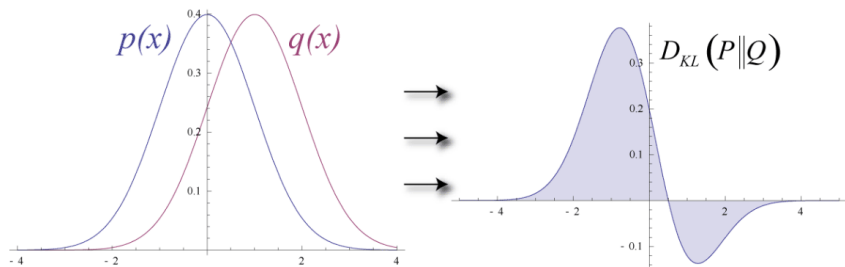
→ Pb : Comment choisir le prior  $\pi$  ?



# Information mutuelle et prior objectif

## Introduction au prior de référence

Idée: Maximiser l'information apportée par les données par rapport au prior. i.e. maximiser  $KL(p(\cdot|a, z)||\pi)$ .



# Information mutuelle et prior objectif

## Introduction au prior de référence

Idée: Maximiser l'information apportée par les données par rapport au prior. i.e. maximiser  $KL(p(\cdot|a, z)||\pi)$ .

### Définition (Information attendue moyenne)

$$\begin{aligned} I(\pi|\mathcal{M}^k) &= \int KL(p(\cdot|a, z)||\pi)d\mathbb{P}_{A,Z}(a, z) \\ &= \int_{\mathcal{A} \times \mathcal{Z}} \int_{\Theta} l_k(z|a, \theta) \log \frac{l_k(z|a, \theta)}{\int_{\Theta} l_k(z|a, \theta)d\pi(\theta)} d\pi(\theta) dc(z) d\mathbb{P}_A(a). \end{aligned}$$

$c =$  mesure de comptage sur  $\mathcal{Z} = \{0, 1\}^k$

# Information mutuelle et prior objectif

## Introduction au prior de référence

Idée: Maximiser l'information apportée par les données par rapport au prior. i.e. maximiser  $KL(p(\cdot|a, z)||\pi)$ .

### Définition (Information attendue moyenne)

$$\begin{aligned} I(\pi|\mathcal{M}^k) &= \int KL(p(\cdot|a, z)||\pi) d\mathbb{P}_{A,Z}(a, z) \\ &= \int_{\mathcal{A} \times \mathcal{Z}} \int_{\Theta} l_k(z|a, \theta) \log \frac{l_k(z|a, \theta)}{\int_{\Theta} l_k(z|a, \theta) d\pi(\theta)} d\pi(\theta) dc(z) d\mathbb{P}_A(a). \end{aligned}$$

$c$  = mesure de comptage sur  $\mathcal{Z} = \{0, 1\}^k$

### Définition (Prior de référence sur $\mathcal{P}$ )

Le prior de référence sur une classe de priors  $\mathcal{P}$  est

$$\pi_k^* = \arg \max_{\pi \in \mathcal{P}} I(\pi|\mathcal{M}^k)$$

[Bernardo 1979] "Reference Posterior Distributions for Bayesian Inference".

## Théorème (Expression implicite de $\pi^*$ )

*Sous certaines hypothèses de régularité et*

$\mathcal{P} = \mathcal{P}_1\{\pi \ll \nu, d\pi/d\nu \in \mathcal{C}^0\}$ :

$$p_k^* = \frac{d\pi_k^*}{d\nu} \propto f = \exp \left( \int l_k(z|a, \theta) \log \text{post}_k^*(\theta|a, z) dc(z) d\mathbb{P}_A(a) \right).$$

*Et si  $\mathcal{P} = \mathcal{P}_1 \cap \{\pi, \int g_i d\pi = c_i\}$ :*

$$p_k^* \propto f \times \exp \left( \sum_{i=1}^n \lambda_i g_i \right)$$

## Théorème (Expression implicite de $\pi^*$ )

Sous certaines hypothèses de régularité et

$\mathcal{P} = \mathcal{P}_1\{\pi \ll \nu, d\pi/d\nu \in \mathcal{C}^0\}$ :

$$p_k^* = \frac{d\pi_k^*}{d\nu} \propto f = \exp\left(\int l_k(z|a, \theta) \log \text{post}_k^*(\theta|a, z) dc(z) d\mathbb{P}_A(a)\right).$$

Et si  $\mathcal{P} = \mathcal{P}_1 \cap \{\pi, \int g_i d\pi = c_i\}$ :

$$p_k^* \propto f \times \exp\left(\sum_{i=1}^n \lambda_i g_i\right)$$

Approximation de  $\text{post}_k$ :

$$\text{post}_k(\theta|a, z) \simeq h_k(\theta|z, a) = \frac{l_k(z|a, \theta)h(\theta)}{\int_{\Theta} l_k(z|a, \theta)h(\theta)d\theta}$$

### Définition (Prior de référence asymptotique)

$\pi^*$  est le prior de référence asymptotique sur  $\mathcal{P}$  si  $\forall \pi \in \mathcal{P}, \forall \bigcup_i \Theta_i = \Theta$  union de compacts:

$$\liminf_{k \rightarrow \infty} I(\pi_i^* | \mathcal{M}^k) - I(\pi_i | \mathcal{M}^k) \geq 0$$

### Définition (Prior de référence asymptotique)

$\pi^*$  est le prior de référence asymptotique sur  $\mathcal{P}$  si  $\forall \pi \in \mathcal{P}, \forall \bigcup_i \Theta_i = \Theta$  union de compacts:

$$\liminf_{k \rightarrow \infty} I(\pi_i^* | \mathcal{M}^k) - I(\pi_i | \mathcal{M}^k) \geq 0$$

### Proposition

- *Si  $\mathcal{P}$  est convexe ce prior est unique.*
- *Si  $\forall \theta \in \Theta$  il existe un estimateur faiblement consistant de  $\theta$ , alors le prior de référence existe.*

---

[Muré 2018] “Objective Bayesian analysis of Kriging models with anisotropic correlation kernel”. Université Sorbonne Paris Cité.

### Théorème

*Sous certaines hypothèses de régularité et sur le choix de  $h$ :*

$$\pi^* \propto \lim_{k \rightarrow \infty} f_k(\theta) = \exp \int l_k(z|a, \theta) \log h_k(\theta|z, a) dc(z) d\mathbb{P}_A(a)$$



# Information mutuelle et prior objectif

## Prior de référence asymptotique

### Théorème

*Sous certaines hypothèses de régularité et sur le choix de  $h$ :*

$$\pi^* \propto \lim_{k \rightarrow \infty} f_k(\theta) = \exp \int l_k(z|a, \theta) \log h_k(\theta|z, a) dc(z) d\mathbb{P}_A(a)$$

### Théorème

*Le prior de Jeffrey  $J(\theta) = \sqrt{|\det \mathcal{I}(\theta)|}$  est le prior de référence asymptotique.*

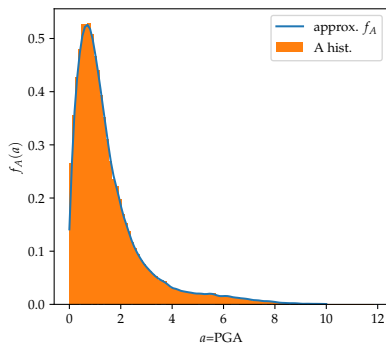
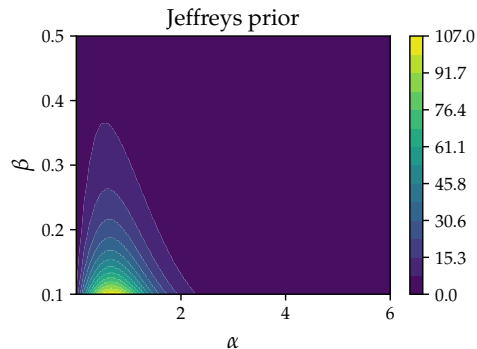
### Définition (Rappel : Information de Fisher)

Considérant la vraisemblance  $l(x|\theta)$ , l'information de Fisher est :

$$\mathcal{I}(\theta)_{ij} = \int_{\mathcal{X}} l(x|\theta) \frac{\partial^2}{\partial \theta_i \partial \theta_j} \log l(x|\theta) dx$$

# Applications numériques

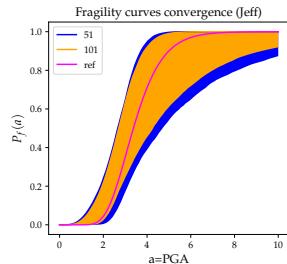
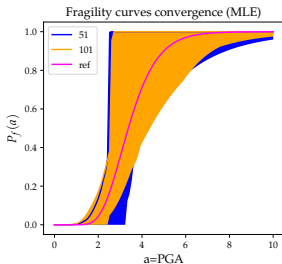
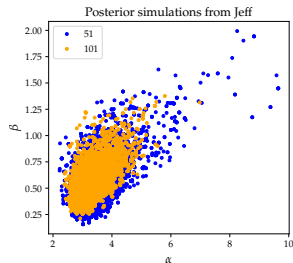
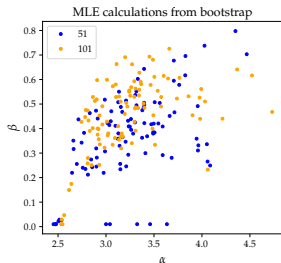
## Prior/posterior de Jeffrey



Prior de Jeffrey et densité de  $A$ .

# Applications numériques

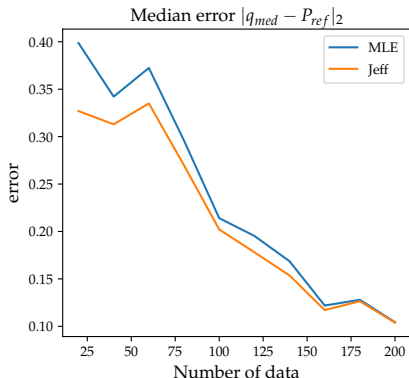
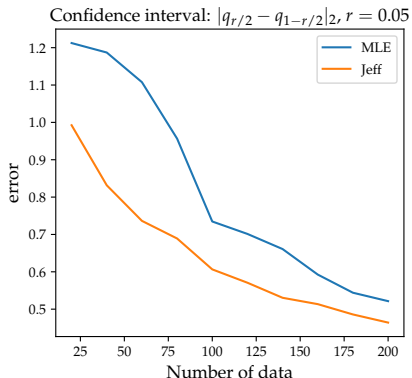
## Estimation de courbes de fragilité



Comparaison des méthodes MLE / Jeff.

# Applications numériques

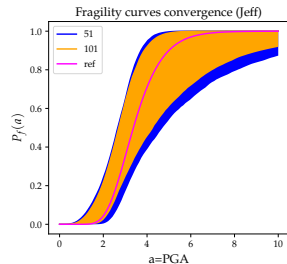
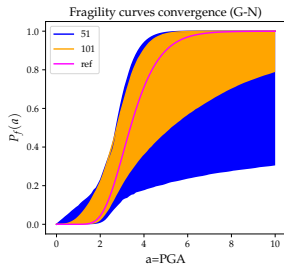
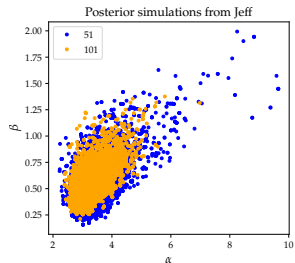
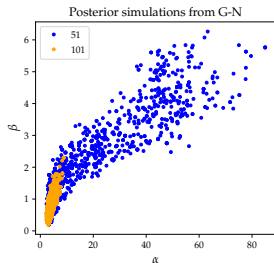
## Convergence des courbes de fragilité



Pour chaque simulation : 10000 jeux de données sont générés artificiellement par bootstrap et 10000  $\theta$  sont générés selon le posterior via MCMC.

# Applications numériques

## Comparaison avec un prior Gamma-Normal



Comparaison des simulations selon G-N / Jeff.

# Caractère propre de Jeffreys lorsque $\beta \rightarrow 0$

## Développement asymptotique de la vraisemblance

On rappelle la vraisemblance :

$$\begin{aligned}L(\theta) = p(z|a, \theta) &= \prod_{i=1}^k \Phi(\gamma_i)^{z_i} (1 - \Phi(\gamma_i))^{1-z_i} \\ &= \exp \left[ \sum_{i=1}^k (z_i \log \Phi(\gamma_i) + (1 - z_i) \log(1 - \Phi(\gamma_i))) \right]\end{aligned}$$

Avec  $\gamma_i = \beta^{-1} \log \frac{a_i}{\alpha} \xrightarrow{\beta \rightarrow 0} \pm\infty$ .

# Caractère propre de Jeffreys lorsque $\beta \rightarrow 0$

## Développement asymptotique de la vraisemblance

On rappelle la vraisemblance :

$$\begin{aligned}L(\theta) = p(z|a, \theta) &= \prod_{i=1}^k \Phi(\gamma_i)^{z_i} (1 - \Phi(\gamma_i))^{1-z_i} \\ &= \exp \left[ \sum_{i=1}^k (z_i \log \Phi(\gamma_i) + (1 - z_i) \log(1 - \Phi(\gamma_i))) \right]\end{aligned}$$

Avec  $\gamma_i = \beta^{-1} \log \frac{\alpha_i}{\alpha} \xrightarrow{\beta \rightarrow 0} \pm\infty$ .

**Proposition (Développement asymptotique de la vraisemblance)**

$$\begin{aligned}L(\theta) &\underset{\beta \rightarrow 0}{=} C \frac{e^{-\frac{K}{\beta^2} + o(1)}}{(2\beta^{-1}\sqrt{2\pi})^N} \quad C, K, N > 0 \\ &\underset{\beta \rightarrow 0}{=} O\left(\beta^N e^{-\frac{K}{\beta^2}}\right).\end{aligned}$$

## Théorème (Caractère impropre du prior de Jeffreys)

Supposons que  $A \sim \mathcal{LN}(\mu, \sigma^2)$  i.e.  $f_A(a) = \frac{1}{\sqrt{2\pi}\sigma a} e^{-\frac{(\log a - \mu)^2}{2\sigma^2}}$ .  
Alors le prior de Jeffreys est impropre.

$$J(\theta) \underset{\beta \rightarrow 0}{\sim} E' \beta^{-3/2} \quad (1)$$



## Théorème (Caractère impropre du prior de Jeffreys)

Supposons que  $A \sim \mathcal{LN}(\mu, \sigma^2)$  i.e.  $f_A(a) = \frac{1}{\sqrt{2\pi}\sigma a} e^{-\frac{(\log a - \mu)^2}{2\sigma^2}}$ .  
Alors le prior de Jeffreys est impropre.

$$J(\theta) \underset{\beta \rightarrow 0}{\sim} E' \beta^{-3/2} \quad (1)$$

## Corollaire (Caractère propre du posterior)

Sous les mêmes hypothèses la loi a posteriori est bien une loi de probabilité.

- Etude théorique des résultats sur le prior de référence.
- Application au contexte des courbes de fragilité sismique et calcul du prior de Jeffrey.
- Résultats numériques sur un cas pratique du CEA qui démontrent:
  - Une suppression des irrégularités qu'impliquent la méthode classique du MLE.
  - Une convergence des estimations issue du prior de Jeffrey, qui est meilleure que celle obtenue avec les autres méthodes étudiées.
- Validation mathématiques des bonnes propriétés des lois simulés.

FIN

# Annexe

## Algorithme de Metropolis-Hasting adaptatif

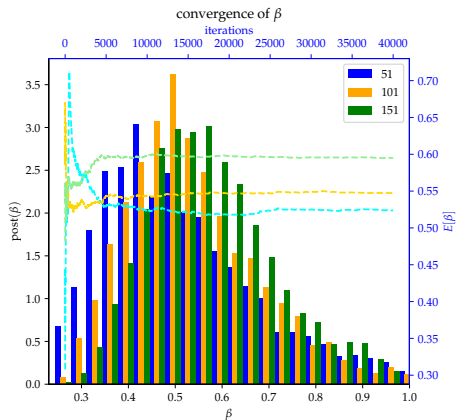
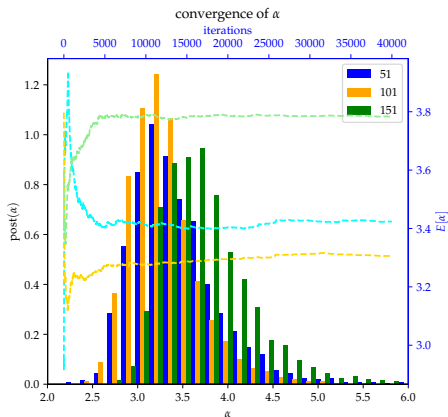
**Require:**  $\pi$ : target distribution  
 $t_0$ : Markov chain initialization  
 $N_{iter}, N_{step}, \sigma_p, b, s$ : calibration parameters

**for**  $k = 1 \dots N_{iter}$  **do**  
  **if**  $k < N_{step}$  **then**  
     $t^* = t_{k-1} + \mathcal{N}(0, \sigma_p I)$   
  **else**  
     $t^* = t_{k-1} + s(1 - b)\mathcal{N}(0, \hat{\Sigma}_{k-1}) + b\mathcal{N}(0, \sigma_p I)$   
  **end if**  
   $\alpha_k = \pi(t^*)/\pi(t_{k-1})$   
   $t_k = t^*$  with probability  $\alpha_k$ ,  $t_k = t_{k-1}$  otherwise  
   $m_k = \frac{1}{k} \sum_{i=1}^k t_i$   
   $\hat{\Sigma}_k = \sum_{i=1}^k (t_i - m_k)(t_i - m_{k+1})^T$   
**end for**

**Ensure:**  $t, \alpha$ : Markov chain and acceptance rates sequences

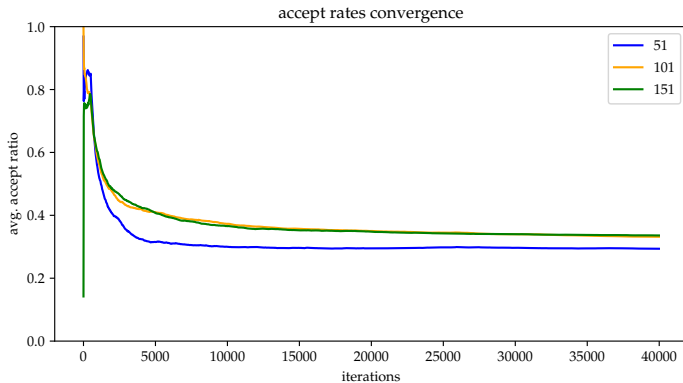
# Annexe

## Convergence du MH



# Annexe

## Convergence du MH



# Annexe

## Comparaison visuelle des médianes

