



Approche décisionnelle pour l'estimation
de critères fiabilistes à l'aide de la
simulation numérique : point de vue
industriel et exemples d'application

Alberto Pasanisi

EDF R&D

Dépt. Management des Risques Industriels

JSTAR 2012, Rennes 26 octobre 2012



Remerciements

- Travaux et réflexions sur l'approche bayésienne menés avec :
 - Merlin Keller (EDF R&D)
 - Eric Parent (AgroParisTech)
- Et, en particulier pour l'application sismique, avec
 - Guillaume Damblin (doctorant EDF / AgroParisTech)
 - Pierre Barbillon (AgroParisTech)
 - Irmela Zentner, Anne-Laure Popelin (EDF R&D)
- Projet CSDL (Complex System Design Lab, 2009-2012)
- Organisateurs JSTAR 2012
 - en particulier, Pierrette Chagneau et Jean-François Dupuy
- Equipe projet « Incertitudes_2015 » d'EDF R&D :
 - Anne Dufloy, Bertrand Iooss, Nicolas Bousquet, Mathieu Couplet, Marc Berveiller, Géraud Blatman, Renaud Barate, Michaël Baudin ...



Sommaire

1. Introduction : Incertitudes en simulation numérique, contexte, enjeux
2. Estimation décisionnelle de grandeurs d'intérêt
3. Application : estimation de courbes de fragilité sismique

INTRODUCTION

Contexte

- Aléas économiques, environnementaux, techniques
- Risques industriels, sanitaires, alimentaires
- Mesures, essais, données quantitatives et qualitatives issues du retour d'expérience
- Modèles prévisionnels complexes → maîtrise de la prévision, décision sous incertitudes, conséquences plus ou moins graves



Ariane 5, 1996



Crise financière, 2008 ...



Roissy, Noël 2010



Columbia, 2003



Xynthia, 2010

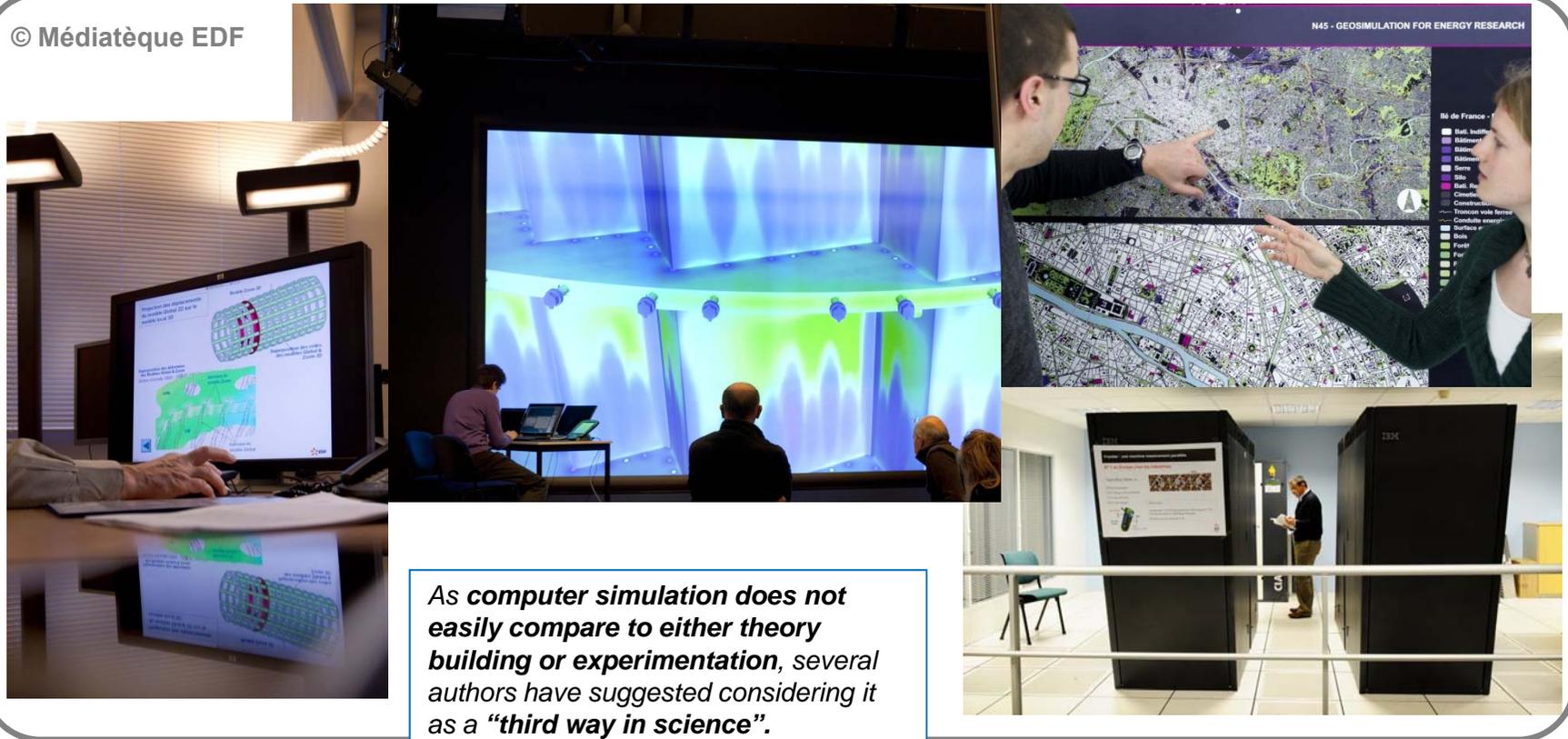


Sud-Est, fév. 2012

La simulation numérique : un atout

- Des codes de calcul de plus en plus complexes pour « modéliser » de manière de plus en plus fine la réalité → une chance pour la science et pour les ingénieurs

© Médiatèque EDF

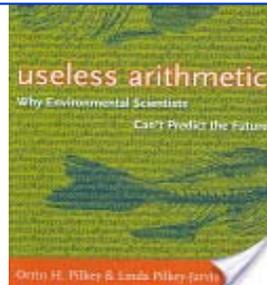


Nécessité d'un positionnement critique

- Nécessité d'un positionnement critique par rapport aux résultats fournis par un modèle/code de calcul. Attention : Garbage in – Garbage out !
- La prise en compte des incertitudes en simulation numérique est aujourd'hui une composante importante de la crédibilité des prévisions basées sur les modèles numériques ... et des décisions qui en découlent

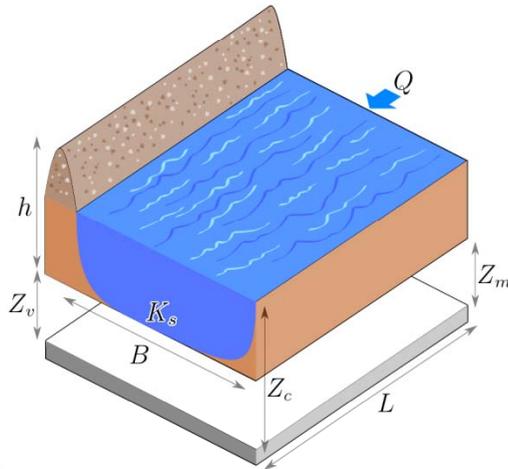
*The reliance on mathematical models has done tangible damage to our society in many ways. **Bureaucrats who don't understand the limitations of modeled predictions often use them.** [...] **Models act as convenient fig leaves for politicians, allowing them to put off needed action on controversial issues.** [...] Agencies that depend on project approvals for their very survival can and frequently do find ways to adjust models to come up with correct answers that will ensure project funding. (Pilkey & Pilkey-Jarvis, 2006)*

*Most simulation models will be complex, with many parameters, state-variables and non linear relations. Under the best circumstances, such models have many degrees of freedom and, **with judicious fiddling, can be made to produce virtually any desired behaviour, often with both plausible structure and parameter values.** (Hornberger & Spear, 1981) <http://sensitivity-analysis.jrc.ec.europa.eu/docs/quotes.html>*



Préconisation et décision sous incertitude

- Les approches statistiques et probabilistes donnent un cadre logique (et mathématiquement fondé) pour quantifier les incertitudes
 - La formalisation des enjeux motivant l'étude par une fonction de coût (ou d'utilité) permet de résoudre le problème dans un cadre décisionnel formellement rigoureux
- Justification de la décision en environnement incertain
 - Deux étapes :
 - Quantification de l'incertitude qui affecte la prévision
 - Formalisation du problème de décision
- Exemple : dimensionner un ouvrage de protection hydraulique



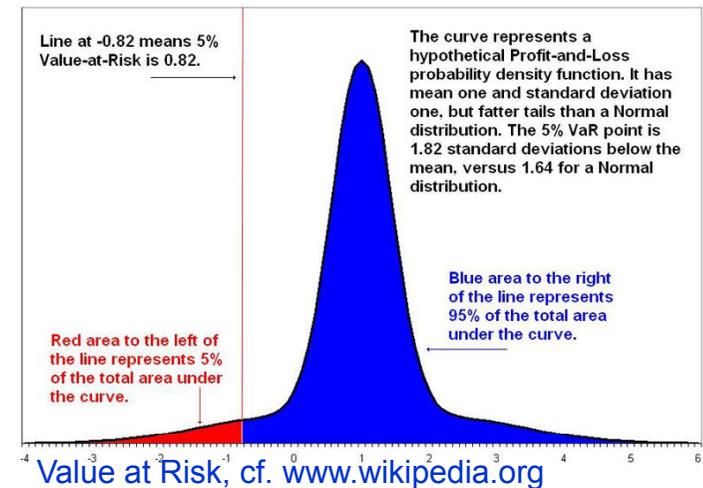
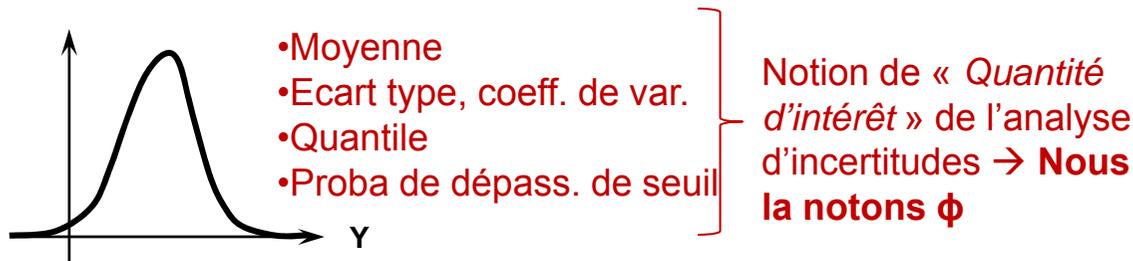
$$Y = Z_v + \left[\frac{Q}{K_s \cdot B \cdot \sqrt{(Z_m - Z_v)/L}} \right]^{3/5} \quad \leftarrow \text{Modèle physique}$$

$$C(h, y) = C_0(h) + C_S(h, y) \cdot \mathbb{1}_{\{y > h\}} \quad \leftarrow \text{Fonction de coût (partie fixe + dommages)}$$

$$\rho(h) = \int C(h, y) \cdot p(y|\mathcal{D}) dy \quad \longrightarrow \quad h_{\text{opt}} = \arg \min_h \rho(h)$$

Une vision « moins large » du problème

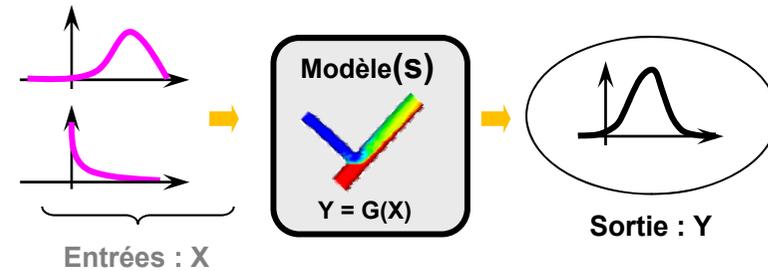
- **Mais l'analyste est souvent loin de la décision** → Plusieurs situations où l'étude n'est pas directement rattaché à un problème de décision
 - Approche « descriptive » : « *Exploration de codes* », analyse de sensibilité → mieux comprendre la modélisation physique
 - En métrologie : évaluer l'incertitude (e.g. écart type) liée à un procédé de mesure est le but de l'étude
 - Approche « normative » : une question très précise est posée à l'analyste sur la distribution de probabilité de la variable de sortie
 - Quelle est la probabilité que la variable de sortie soit plus grande (ou plus petite) qu'une valeur de seuil ?
 - Quelle valeur de Z a une probabilité de dépassement donnée ? → quantiles (e.g. hauteur de crue centennale, « value at risk »)
 - Ces différents problèmes se basent tous sur la distribution de probabilité de la sortie du code numérique en question



Cadre : Analyse d'incertitude, sensibilité, AFS

Analyse d'incertitude (problématique générale)

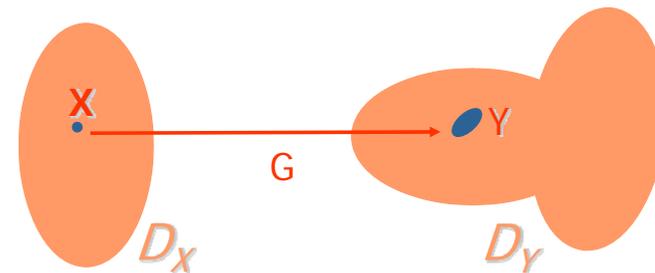
The study of the distribution of the code output that is induced by probability distributions on input
Kennedy & O'Hagan, 2001



Analyse de sensibilité

Questions centrales :

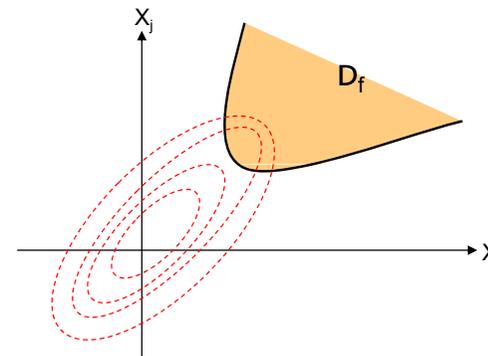
La sortie Y est-elle plus ou moins variable lorsque l'on fixe une des entrées X_i ?
 Quelles variables d'entrée sont les plus influentes ?



Analyse de fiabilité (des structures)

Quelle est la probabilité que la variable Z dépasse un seuil donné ?

$$p_f = \int \mathbb{1}_{\{G(x) \leq 0\}} p(x) dx = \int_{\mathcal{D}_f} p(x) dx$$



Enjeu typique : justifier un critère de sureté

Des problèmes différents, des questions différentes



Risque hydraulique :

Comment dimensionner un ouvrage de protection ?
Quelle est la probabilité que la plaine soit inondée ?
Deux questions différentes : phase de conception et de vérification d'un ouvrage.



Performance énergétique d'un bâtiment :

Comment définir la consommation énergétique d'un bâtiment (p.ex. dans le cadre d'un CPE) ?
Quels moyens mettre en œuvre pour atteindre ces objectifs ?

ESTIMATION DÉCISIONNELLE DE GRANDEURS D'INTÉRÊT

Ingrédients et notations

- Un code de calcul (déterministe à entrées aléatoires) : $Y = G(X)$
- La loi de X appartient à une famille paramétrique : $p(x|\theta)$
- A disposition : de l'expertise et des données pour estimer θ : $\pi_0(\theta) \quad \mathcal{D}$
- Problème : estimer une « quantité d'intérêt » de la distrib. de Y : $\phi = \phi(\theta)$
- A θ fixé, la valeur de ϕ est fixée, même si son calcul peut être un défi numérique
- **On se focalise sur des grandeurs d'intérêt fiables** (quantiles, probabilités de défaillance)

- Mais θ n'est pas connu !
- Approche *classique* par plug-in : 

$$\hat{\theta} = \underset{\theta}{\text{ArgMax}} [L(\mathcal{D}|\theta)]$$
$$\hat{\phi} = \phi(\hat{\theta})$$

 - NB : On néglige l'incertitude sur θ

- En pratique θ est estimé à partir de l'info disponible (forcément limitée)
 - **Comment prendre en compte l'incertitude sur θ , i.e. l'incertitude sur le modèle proba $p(x|\theta)$?**

Pratiques courantes de l'ingénieur : plug-in

- Approche qui suppose une connaissance parfaite du modèle probabiliste des entrées (i.e. la valeur du paramètre θ)
 - Typiquement : estimation de θ par max de vraisemblance \longrightarrow
 - Méthode critiquable car le paramètre θ est en pratique connu avec incertitude
 - Notons au passage que l'incertitude affectant θ est de nature différente de celle de $X|\theta$
 - « aléatoire » vs. « épistémique »
- $$\hat{\theta} = \underset{\theta}{\text{ArgMax}} [L(\mathcal{D}|\theta)]$$
$$\hat{\phi} = \phi(\hat{\theta})$$
- Comment tenir compte de l'incertitude sur le modèle proba $p(x|\theta)$?
 - La tentation de l'extra-probabiliste \rightarrow un sujet d'actualité, mais plein de problèmes pas résolus !
 - Le cadre bayésien donne un cadre mathématique rigoureux ... qui permet de « ré-interpréter » certaines pratiques
 - Il y a plein de bayésiens qui s'ignorent (bayésiens « opportunistes » ?)

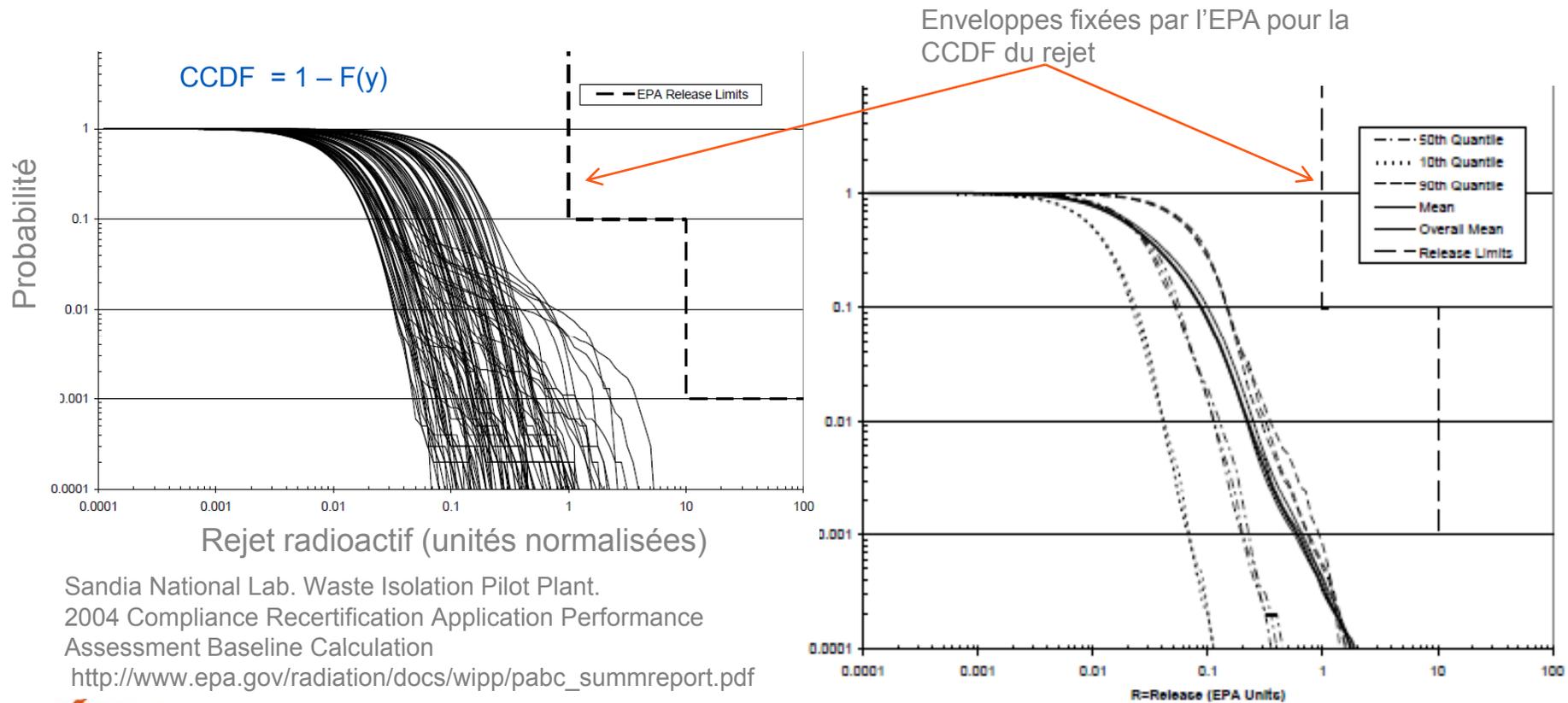
Pratiques courantes de l'ingénieur : évaluer $\pi(\theta)$

- Par l'avis d'expert uniquement (si pas de données) : $\rightarrow \pi_0(\theta)$
- Par la « loi asymptotique de l'estimateur de max de vraisemblance » : $\longrightarrow \mathcal{N}\left(\hat{\theta}, \frac{1}{n}\mathcal{I}^{-1}(\hat{\theta})\right)$
 - dans laquelle θ a été remplacé par son estimateur MLE
- Par la loi a posteriori de θ : $\longrightarrow \pi(\theta|\mathcal{D}) \propto \pi_0(\theta) \cdot L(\mathcal{D}|\theta)$
- Faisant ainsi, l'ingénieur se place automatiquement dans un cadre bayésien \rightarrow modéliser la connaissance de l'*Etat de la Nature* (ici θ) par une loi de probabilité
- $\mathcal{N}\left(\hat{\theta}, \frac{1}{n}\mathcal{I}^{-1}(\hat{\theta})\right)$ est la limite asymptotique de la loi a posteriori $\pi(\theta|\mathcal{D})$
- Tous des bayésiens ? Pas si simple.
 - Il ne s'agit juste d'établir $\pi(\theta)$, mais de l'utiliser !

Des approches « purement descriptives »

- Evaluation d'un certain nombre de valeurs de ϕ (ou clairement de toute la loi de Y), chacune desquelles conditionnelles à une valeur de θ

Très utilisées aux US en sûreté nucléaire. Exemple de résultats fournis par SANDIA à EPA



Estimation ponctuelle de ϕ en tenant compte de l'incertitude sur θ . L'estimateur « prédictif »

- Heuristique :

- Remplacer la loi de Y, connue seulement conditionnellement à θ , par sa loi prédictive $\longrightarrow p(y|\mathcal{D}) = \int p(y|\theta) \cdot \pi(\theta|\mathcal{D}) d\theta$
- Cette loi prédictive de Y est obtenu en propageant dans le modèle G, la prédictive de X $\longrightarrow p(x|\mathcal{D}) = \int p(x|\theta) \cdot \pi(\theta|\mathcal{D}) d\theta$
- C'est un plug-in : au lieu de remplacer θ par son estimateur, on remplace $p(y)$ par sa prédictive :

- Mise en œuvre : « Double Monte Carlo » ou « Nested Monte Carlo »

- Générer un échantillon de $\pi(\theta|\mathcal{D}) : \theta_1, \theta_2, \dots, \theta_n$
- Tirer aléatoirement : $x_i \sim p(x|\theta_i) \quad i = 1, \dots, n$
- Calculer : $y_i = G(x_i) \quad i = 1, \dots, n$
- Estimer ϕ à partir de l'échantillon de Y ainsi construits

Evidemment, si besoin, des techniques d'estimation plus astucieuses que le Monte Carlo brutal seront utilisées, mais le principe reste le même.



Par exemple $\phi = \mathbb{E}(Y) \implies \hat{\phi} \approx \frac{1}{n} \sum_i G(x_i)$

What's wrong ?

- Le cadre bayésien donne (déjà) la solution du problème de l'estimation de $\phi(\theta)$

- C'est un problème de décision : *The supposed dichotomy between inference and decision is illusory, since any report or communication of beliefs following the receipt of information inevitably itself constitutes a form of action.* (Bernardo & Smith, 1984)

$$\hat{\phi}_{\text{BAY}} = \underset{d}{\text{ArgMin}} \left[\int C(\phi(\theta), d) \cdot \pi(\theta|\mathcal{D}) d\theta \right]$$

- $C(\phi(\theta), d)$ est une fonction qui quantifie les conséquences de l'estimation de la vraie grandeur inconnue ϕ par la valeur $d \rightarrow$ on choisit la valeur de d qui minimise ce coût

- Quid des estimateurs prédictifs ?

- Christensen et Huffman (1985) : **Si ϕ peut s'écrire comme une espérance** d'une quelconque fonction de Y , alors l'estimateur prédictif est un estimateur bayésien sous l'hypothèse de coût quadratique

$$\hat{\phi}_{\text{pred}} = \underset{d}{\text{ArgMin}} \left[\int (\phi(\theta) - d)^2 \cdot \pi(\theta|\mathcal{D}) d\theta \right]$$

- Exemples : $\phi = \mathbb{E}[Y]$, $\phi = \mathbb{V}[Y]$
 $\phi = \mathbb{P}[Y \geq y^*] = \mathbb{E}[\mathbb{1}_{\{Y \geq y^*\}}]$

Cas de l'AFS : probabilité d'occurrence d'un événement redouté. Peut on accepter une fonction de coût quadratique et donc symétrique ?
Autrement dit, surestimer ou sous-estimer, c'est vraiment pareil ?

What's wrong ?

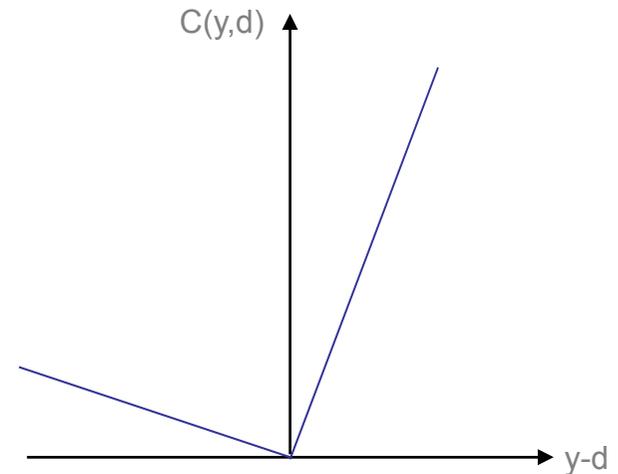
- Quid des estimateurs prédictifs ? (suite)

- **Cas où $\phi(\theta)$ est un quantile de probabilité α**
- L'estimateur prédictif est la solution du problème suivant (Parent et Bernier, 2007) :

$$\text{ArgMin}_d \left[\int C(y, d) \cdot p(y|\mathcal{D}) dy \right]$$

- où $C(y, d)$ est une fonction de coût linéaire par morceaux
 $C(y, d) = c_1 \cdot (y - d) \cdot \mathbb{1}_{\{y \geq d\}} + c_2 \cdot (d - y) \cdot \mathbb{1}_{\{d \geq y\}}$

avec : $\frac{c_1}{c_1 + c_2} = \alpha$



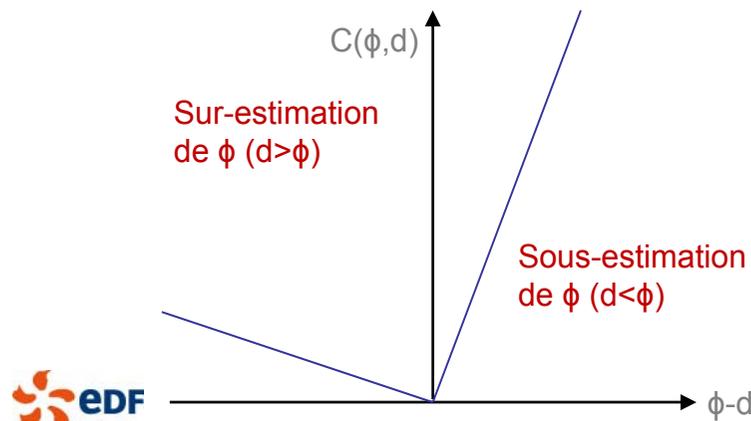
NB Cette heuristique n'est pas un estimateur bayésien de ϕ , mais de Y .
La fonction de coût dépend de l'ordre du quantile !

- Dans les deux cas, cette heuristique est un estimateur bayésien obtenu à partir de fonctions de coût « imposées » !

- **Problème de cohérence du double Monte Carlo :**
- **C'est à l'analyste de choisir, pas à la procédure utilisée pour les calculs !**
- Pourquoi utiliser des fonctions de coût différentes pour proba de dépassement et quantiles ?
- Symétrie dans le cas des proba de défaillance !

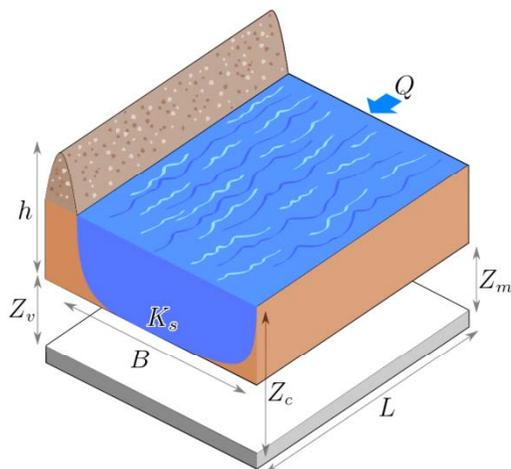
Intérêt en analyse d'incertitude

- Revenons au problème de départ : estimer une quantité d'intérêt (e.g. une « probabilité de défaillance »)
 - Cette grandeur, qui dépend de θ est incertaine
 - Comment choisir (et justifier) une valeur parmi toutes les valeurs possibles
 - Le cadre décisionnel bayésien fournit une solution : modéliser les conséquences de la sur-estimation ou sous-estimation par une fonction de coût $C(\phi, d)$
$$\hat{\phi}_{\text{BAY}} = \underset{d}{\text{ArgMin}} \left[\int_{\theta} C(\phi(\theta), d) \cdot \pi(\theta | \mathcal{D}) d\theta \right]$$
 - Parmi d'autres, deux fonctions de coût sont particulièrement simples à appréhender (et à manipuler) :
 - Quadratique (cas symétrique) \rightarrow l'estimateur de ϕ est sa moyenne
 - Linéaire par morceaux (cas dyssymétrique : sous-estimer est x fois plus grave que sur-estimer) \rightarrow l'estimateur de ϕ est un quantile donné, de proba $c_1/(c_1 + c_2)$



La théorie bayésienne permet de justifier des estimateurs ponctuels de la quantité d'intérêt ... ou de « réinterpréter » les heuristiques parfois utilisées par les ingénieurs

Exemple (1/2). Hypothèses et inférence



$$Y = Z_v + \left[\frac{Q}{K_s \cdot B \cdot \sqrt{(Z_m - Z_v)/L}} \right]^{3/5}$$

← Modèle physique

$$\begin{aligned} Z_v &\sim \mathcal{T}(49, 50, 51) \\ Z_m &\sim \mathcal{T}(54, 55, 56) \\ K_s &\sim \mathcal{N}(30, 7.5) \end{aligned}$$

← Modèle proba. sur Z_m, Z_v, K_s

$$Q \sim \text{Gu}(\eta, \beta)$$

← Modèle proba. de Q

$\mathcal{D} =$	1306,	1235,	1830,	2442,	1128,	3330,
	1530,	3192,	2647,	238,	706,	1903,
	1594,	935,	1100,	2204,	1366,	1629,
	522,	642,	1173,	424,	1837,	1391,
	789,	383,	1858,	917,	1084,	1026)

← Données de débit

Prior

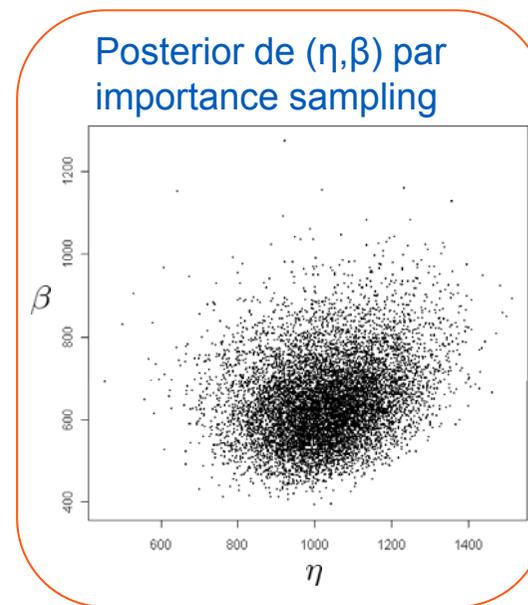
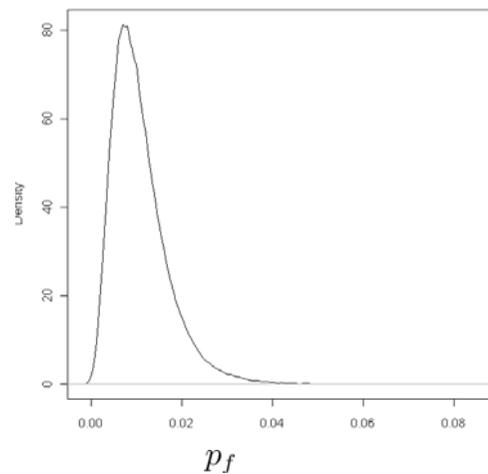
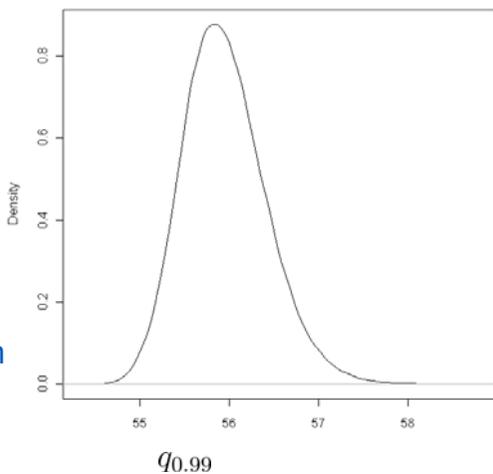
$$\pi_0(\eta) = \mathcal{G}(1, 500)$$

$$\pi_0(\beta) = \mathcal{G}(1, 200)$$

Posterior de (η, β) par importance sampling

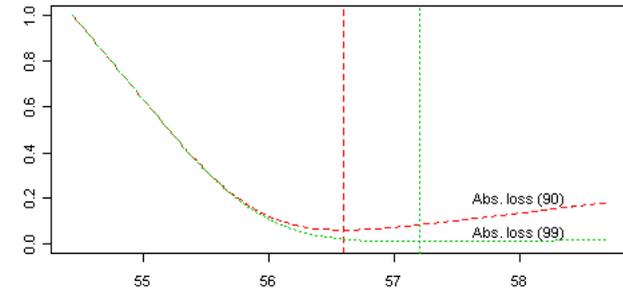
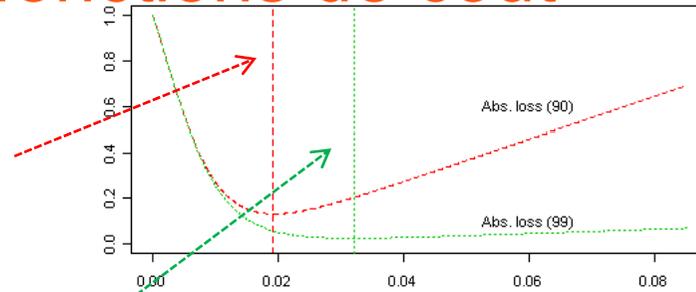
Posterior de deux quantités d'intérêt fiabiliste

- quantile à 99%
- proba de dépass. d'une digue de 56 m

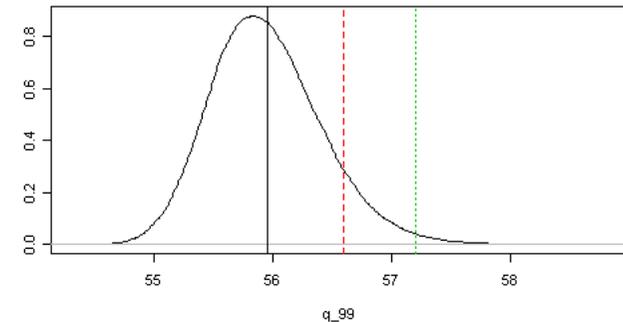
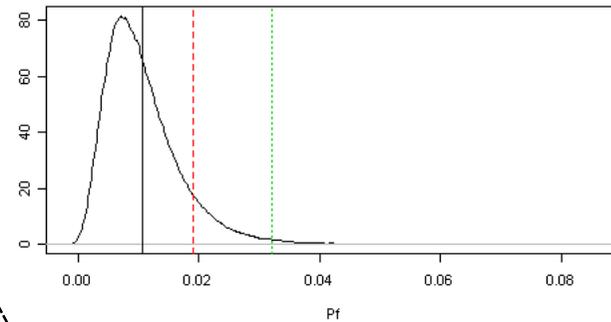


Exemple (2/2) Estimateurs bayésiens à partir de différentes fonctions de coût

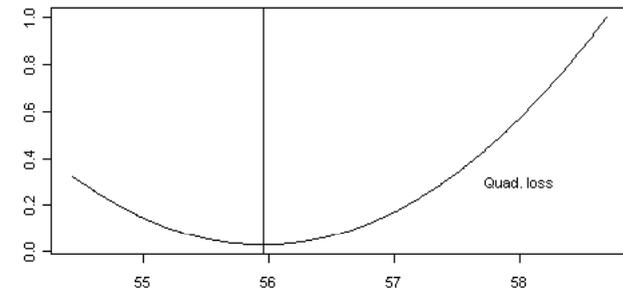
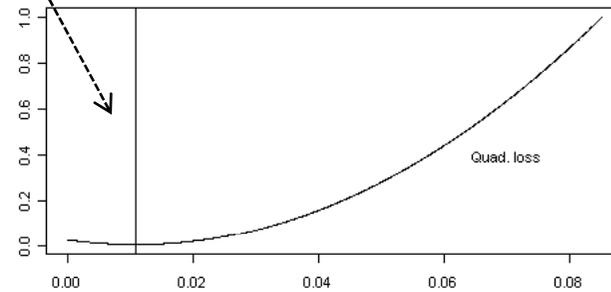
Espérance coût lin. avec sous-estim. pénalisée 9 fois plus que la sur-estim



Espérance coût lin. avec sous-estim. pénalisée 99 fois plus que la sur-estim



Espérance coût quadratique



Pour info, résultats plug-in MLE :

$$\hat{P}_f = 0.0085; \quad \hat{q}_{99\%} = 55.87.$$



Généralisation des résultats sur l'estimateur prédictif (Keller et al., 2011)

- Cadre méthodologique plus générale où l'on inclut le modèle numérique même dans l'Etat de la Nature : $\Theta = (F, G)$

Modèle probabiliste
des entrées du code

Modèle physique
(déterministe)

- On s'est intéressé à l'estimateur prédictif et à une classe de fonctions de coût particulières

- Nouvelles notations de l'estimateur prédictif : $\hat{\Phi}^{pred.} = \Phi(H(\cdot|\mathcal{D}))$

– Loi prédictive de Y

$$H(\cdot|\mathcal{D}) = \int H(\cdot|\Theta)\pi(\Theta|\mathcal{D}) d\Theta = \mathbb{E}[H(\cdot|\Theta)|\mathcal{D}]$$

Généralisation des résultats sur l'estimateur prédictif (Keller et al., 2011)

- Théorème qui généralise les résultats précédents

Soit $\Phi(H(\cdot|\Theta))$ une fonction d'intérêt de la distribution de Y . On suppose que :

1. Φ peut-être définie comme minimisant un coût :

$$\Phi(H(\cdot)) = \arg \min_d c(d; H(\cdot)).$$

2. $c(d; H(\cdot))$ est *linéaire* en $H(\cdot)$:

$$\sum_j p_j \times c(d; H_j(\cdot)) = c\left(d; \left\{ \sum_j p_j \times H_j(\cdot) \right\}\right),$$

Alors, $\hat{\Phi}^{pred}$ coïncide avec l'estimateur de Bayes relatif au coût $c(d; H(\cdot|\Theta))$.

Conclusions (partielles)

- La théorie bayésienne donne un cadre adapté à l'analyse d'incertitudes en simulation numérique dans un contexte industriel
 - Quantification de toutes les sources d'incertitude
 - Prise en compte des enjeux → conséquences de la sur-estimation / sous-estimation
 - Possibilité de prendre en compte l'expertise
- Justification de la valeur fournie pour un critère fiabiliste par des arguments simples et parlant aux ingénieurs
- La méthodologie peut être difficile à mettre en œuvre si le code est couteux

APPLICATION : COURBES DE FRAGILITÉ SISMIQUE

Qu'est ce qu'une courbe de fragilité F

- Définition : $F : s \in \mathbb{R}^+ \rightarrow F(s) \in [0, 1]$
- s : grandeur scalaire qui caractérise la sollicitation (e.g. accélération sismique)
- $F(s)$: probabilité de défaillance, conditionnelle à une valeur s de la sollicitation
- Hypothèses :
 - F continue et croissante
 - la structure défaille à partir d'une valeur limite de la sollicitation, notée c

- F est la fonction de répartition d'une v.a. (latente) C (capacité de la structure)

$$F(s) = \mathbb{P}[C \leq s]$$

- Données à disposition (expériences numériques ou physiques) : couples (s_i, y_i)

$$Y_i \sim \mathcal{B}(F(s_i))$$

Sollicitation Indicateur binaire de
défaillance

- Hypothèse : F appartient à la famille log-normale : $\ln C \sim \mathcal{N}(\mu, \tau)$
 - Estimer F revient à estimer la moyenne μ et la précision $\tau = \sigma^{-2}$ de loi de la log-capacité
 - Notation : F_θ avec $\theta = (\mu, \tau)$

Courbes de fragilité : plusieurs domaines d'application quelques exemples (parmi d'autres)

- **Ingénierie sismique**
- Enveloppes de bâtiments soumises à l'impact de débris emportés par le vent (Herbin et Barbato, 2012)
- Baies vitrées sollicitées par des explosions (Stewart et Netherton, 2008)
- Ponts sous l'effet de vagues de tsunami (Shoji et Oriyama, 2007)
- Ouvrages de protection hydraulique exposés à des niveaux d'eau exceptionnels, e.g. études post-Katrina (Ayyub et al., 2009)



New Orleans,
août 2005



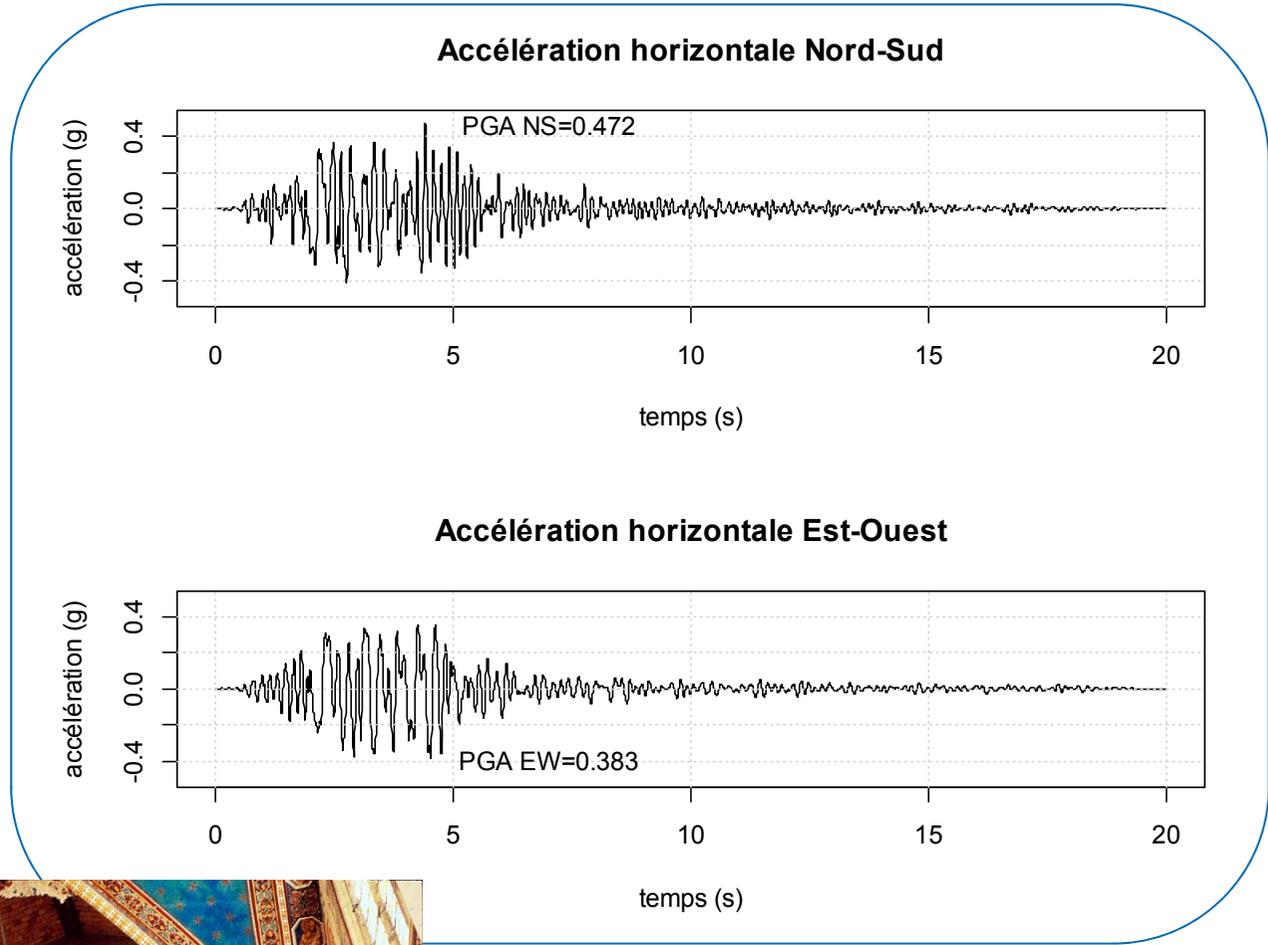
A propos de la sollicitation sismique

Un choix usuel :
PGA *moyenne* horizontale

Peak Ground accélération

Moyenne des PGA
mesurées selon deux
directions orthogonales

Exemple de signal
sismique et de PGA 



Prise en compte dans la pratique technique de l'incertitude affectant l'estimation de $\theta=(\mu,\tau)$

- Shinozuka et al. (2000) :

- Max de vraisemblance, puis par bootstrap construction d'un échantillon de $(\hat{\mu}, \hat{\tau})$

- Encadrement « à 90% » de F par : $F(\cdot|\hat{\mu}_{5\%}, \hat{\tau}_{mv})$ $F(\cdot|\hat{\mu}_{95\%}, \hat{\tau}_{mv})$

Quantiles
« bootstrap » de μ
MLE de τ

- Choisir comme courbe de fragilité : $\int F(\cdot|\hat{\mu}, \hat{\tau}_{mv}) h(\hat{\mu}) d\hat{\mu}$

Loi calée sur l'échantillon bootstrap de μ

- Modélisation statistique inspirée par la « méthode EPRI » (expertise)

- Straub et Der Kiureghian (2008)

- Cadre bayésien → Loi a posteriori de (μ, τ)

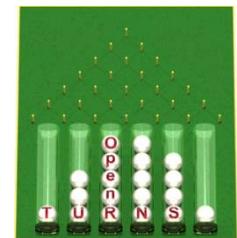
- Estimateur « prédictif » de la courbe de fragilité : $\int \int F(\cdot|\mu, \tau) \pi(\mu, \tau|\mathcal{D}) d\mu d\tau$.

Loi a posteriori de (μ, τ)

- Objectif du travail présenté : construction d'estimateurs bayésiens de F

Canevas bayésien pour l'estimation de F

- L'estimation de F repose sur les 4 étapes suivantes :
 - Quantifier l'incertitude a priori sur F par une loi de probabilité a priori $\Pi_0(F)$
 - Mise a jour de la loi a priori par l'observation des données $(x_i = \ln(s_i), y_i)$
Loi a posteriori : $\Pi(F|(x, y)) \propto L((x, y)|F) \times \Pi_0(F)$
 - Définition d'une fonction de coût : $C(\tilde{F}, F)$
 - Choisir la courbe de fragilité comme celle qui minimise le coût espéré : $\hat{F} = \underset{\tilde{F}}{\text{ArgMin}} \int_{F \in \Psi} C(\tilde{F}, F) \Pi(F|(x, y)) dF$
- Calculs menés sur un cas benchmark
 - Données (50) obtenues par simulation numérique (*Salomé_Méca*)
 - Sollicitation sismique : PGA moyenne horizontale
 - Signaux sismiques générés au hasard dans un catalogue significatif



Inférence bayésienne sur θ : loi a priori

- Loi a priori Gamma-Normale : $\theta \sim \mathcal{GN}(a, b, m, \lambda)$

$$\begin{cases} \pi_0(\mu|\tau) = \mathcal{N}(m, \lambda\tau) \\ \pi_0(\tau) = \mathcal{G}(a, b) \end{cases}$$

- Prior impropre non informatif (pour $a, b, \lambda \rightarrow 0$) :

$$\begin{cases} \pi_0(\mu|\tau) \propto 1 \\ \pi_0(\tau) \propto \tau^{-1} \end{cases}$$

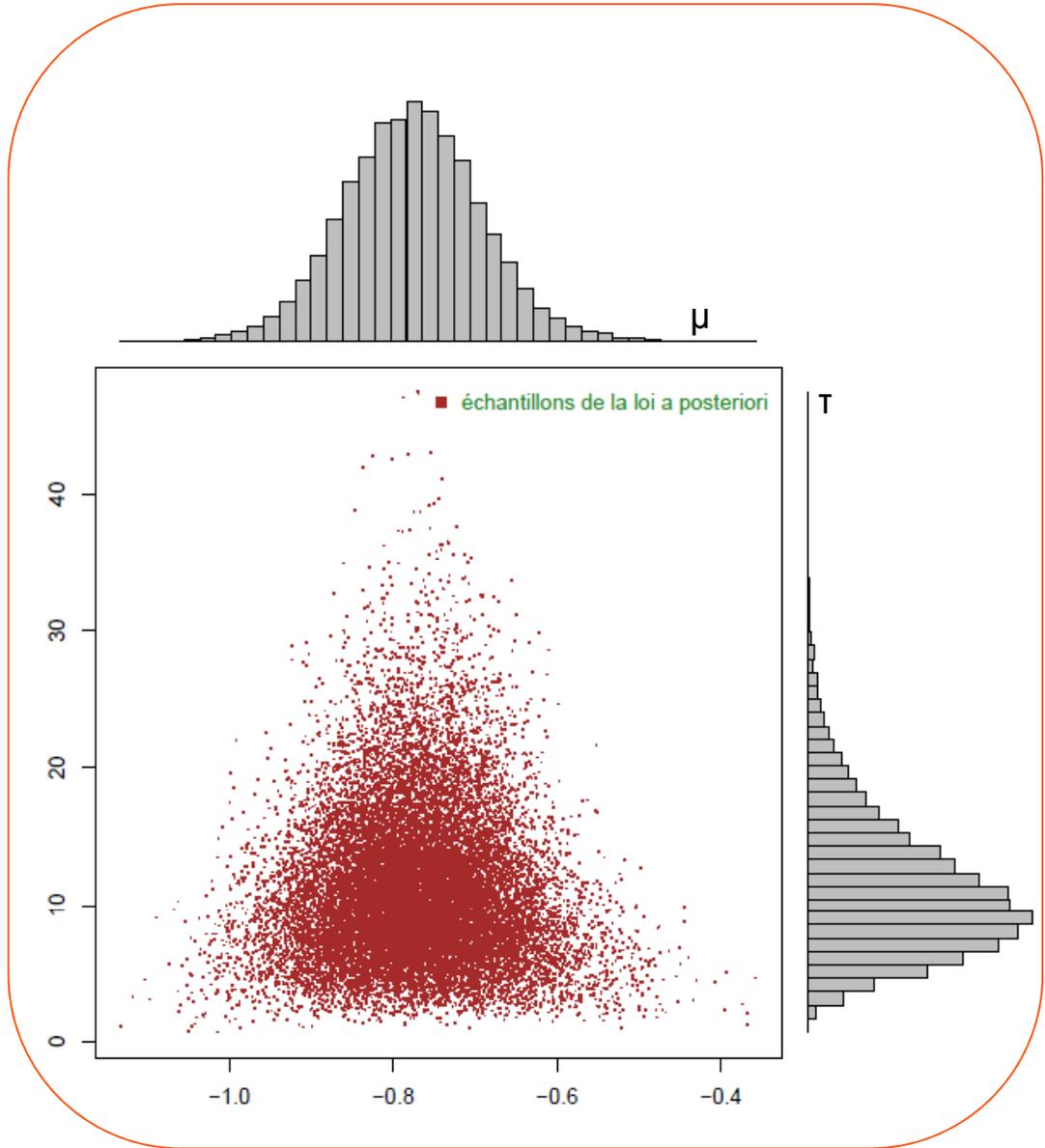
- Le modèle statistique est conjugué, conditionnellement aux données latentes (log-capacités)

Loi a posteriori de θ

Echantillonneur de Gibbs avec une étape d'*augmentation de données* : simulation des log-capacités (latentes) z , sachant les données et la valeur courante de θ :

$$\pi(\{z\}_{i=1..n} | x, y, \theta) \sim \prod_{i=1}^n [\mathcal{N}_{z_i \leq x_i}(z_i | \theta)^{y_i} \cdot \mathcal{N}_{z_i > x_i}(z_i | \theta)^{1-y_i}]$$

Résultats sous un prior « peu informatif » Gamma-Normal →



Fonctions de coût utilisées

Coût quadratique

$$C_2(F_{\hat{\theta}}, F_{\theta}) = \int_0^{+\infty} (F_{\hat{\theta}}(s) - F_{\theta}(s))^2 dP(s)$$

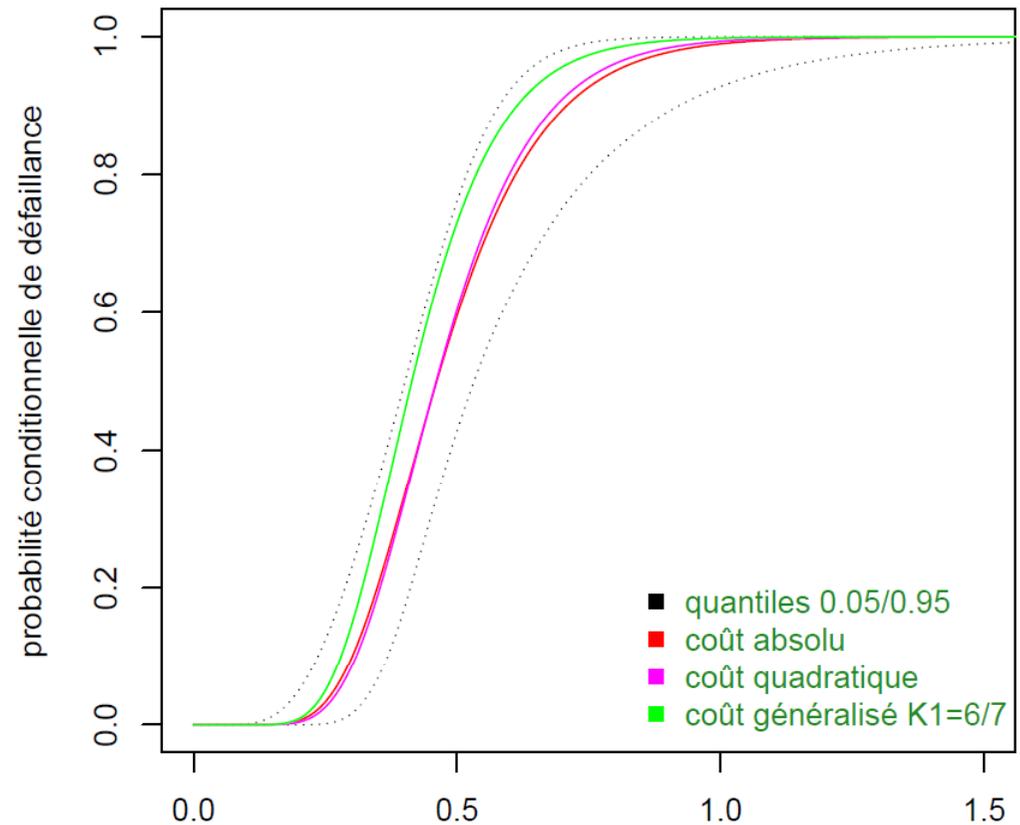
Coût absolu « généralisé » : pénalise différemment la sur-estimation et la sous-estimation de la probabilité de défaillance (pénalisations réglées par le rapport K_1/K_2)

Manière pragmatique de « formaliser » le conservatisme (cf. Parent et Bernier, 2007)

$$C_{K_1, K_2}(F_{\hat{\theta}}, F_{\theta}) = \int_0^{+\infty} K_1 |F_{\hat{\theta}}(s) - F_{\theta}(s)| \mathbb{1}_{F_{\hat{\theta}}(s) > F_{\theta}(s)} dP(s) + \int_0^{+\infty} K_2 |F_{\hat{\theta}}(s) - F_{\theta}(s)| \mathbb{1}_{F_{\hat{\theta}}(s) < F_{\theta}(s)} dP(s)$$

Exemple de résultats : courbes de fragilité

- Trois fonctions de coût
 - Quadratique
 - Absolu ($K_1=K_2$)
 - Absolu ($K_1/K_2=6/7$)
- En pointillé :
 - Encadrement à 90% de la proba de défaillance à niveau de sollicitation fixé



Commentaires

- Approche (relativement) simple mais rigoureuse pour justifier le choix d'une courbe de fragilité log-normale
- Interprétation (relativement) simple du conservatisme
- On prend bien en compte l'incertitude sur les deux paramètres de la loi
 - Sans se restreindre à un seul paramètre (typiquement μ)
 - Sans faire d'hypothèses asymptotiques
- La méthode EPRI, basée sur l'expertise, (non abordée ici) peut fournir (éventuellement) un prior informatif

Perspectives

- Coupler la modélisation de la courbe de fragilité à celle de l'aléa sismique (i.e. loi de la PGA)
 - Travailler directement sur la probabilité de défaillance $\int F(s)dP(s)$
 - En pratique, il est commode de séparer les deux problèmes (fragilité, aléa)
- Etendre la famille de lois paramétriques (e.g. transformation Box-Cox)
 - mais problèmes de transfert vers les études
- Extension non-paramétrique
- Papier à soumettre dans les prochains jours
 - Damblin G., Keller M., Pasanisi A., Barbillon P., Parent E. (2012) Approche décisionnelle bayésienne des incertitudes dans un contexte industriel. Application aux courbes de fragilité sismique (à soumettre au Journal de la Société Française de Statistique).

Références

- Ayyub B.M., Foster J., McGill W.L. (2009) Risk analysis of a protected hurricane-prone region. I : Model development. *Natural Hazards Review*, 10(2), 38-53.
- Bernardo, J. M., Smith, A. F. M. (1994). *Bayesian Theory*. Chichester: Wiley.
- Christensen R., Huffan M.D. (1985) Bayesian Point Estimation Using the Predictive Distribution, *The American Statistician*, 39(4), 319-321.
- Herbin A.H., Barbato M. (2012) Fragility curves for building envelope components subject to windborne debris impact. *Journal of Wind Engineering and Industrial Aerodynamics*, 107,108(0), 285–298.
- Heymann M. (2010) Understanding and misunderstanding computer simulation: The case of atmospheric and climate science - An introduction. *Stud. Hist. Phil. Mod. Phys.* 41, 193-200.
- Hornberger G.M., Spear R. C. (1981) An approach to the preliminary analysis of environmental systems. *Journal of Environmental management*, 12, 7-18.
- Keller M., Pasanisi A., Parent E. (2011) Réflexions sur l'analyse d'incertitudes dans un contexte industriel : Information disponible et enjeux décisionnels. *Journal de la SFdS*, 152(4), 60-77.
- Kennedy M. , O'Hagan, A. (2001). Bayesian calibration of computer models (with discussion). *Journal of the Royal Statistical Society, Series B.* 63, 425-464.
- Parent É., Bernier J. (2007) *Le raisonnement bayésien. Modélisation et inférence*. Paris : Springer.
- Pasanisi A., Keller M., Parent E. (2012) Estimation of a quantity of interest in uncertainty analysis: some help from Bayesian decision Theory. *Reliability Engineering & System Safety*, 100, 93-101.
- Pilkey O.H., Pilkey-Jarvis L. (2006) *Useless Arithmetic: Why Environmental Scientists Can't Predict the Future*. Columbia University Press
- Shinozuka M, Feng M., Lee J. et Naganuma T. (2000) Statistical analysis of fragility curves. *Journal of Engineering Mechanics*, 126(12),1224-1231
- Shoji G., Moriyama T. (2007) Evaluation of the structural fragility of a bridge structure subjected to a tsunami wave load. *Journal of Natural Disaster Science*, 29(2):73-81.
- Stewart G., Netherton D. (2008) Security risks and probabilistic risk assessment of glazing subject to explosive blast loading. *Reliability Engineering and System Safety*, 93(4):627-638,.
- Straub D. et Der Kiureghian A. (2008) Improved seismic fragility modeling from empirical data. *Structural Safety*, 30(4):320-336, 2008.

MERCI DE VOTRE ATTENTION