Séminaire IMPEC – CEA Saclay le 17 octobre 2006

Incertitudes et surfaces de réponses : premières expérimentations sur la robustesse quantiles

E. de Rocquigny (EDF R&D), Y. Barhoumi, L. Carraro (EMSE)

EDF R&D

Sommaire

- Incertitudes et SR : exemples EDF et approche globale
  - Quelques exemples de dossiers
  - D'où le schéma global et les grandes étapes
  - Plusieurs contextes d'utilisation des PE/SR/AS selon la finalité
  - Une exigence récente : robustesse de la prédiction à base de SR dans un calcul de risque

- Premières pistes pour réduire la variance de l'estimateur en combinant Wilks et une SR
  - Réduction par formulation asymptotique sous indépendance
  - Addition de quantile sous inégalité forte

- Approfondissements à venir

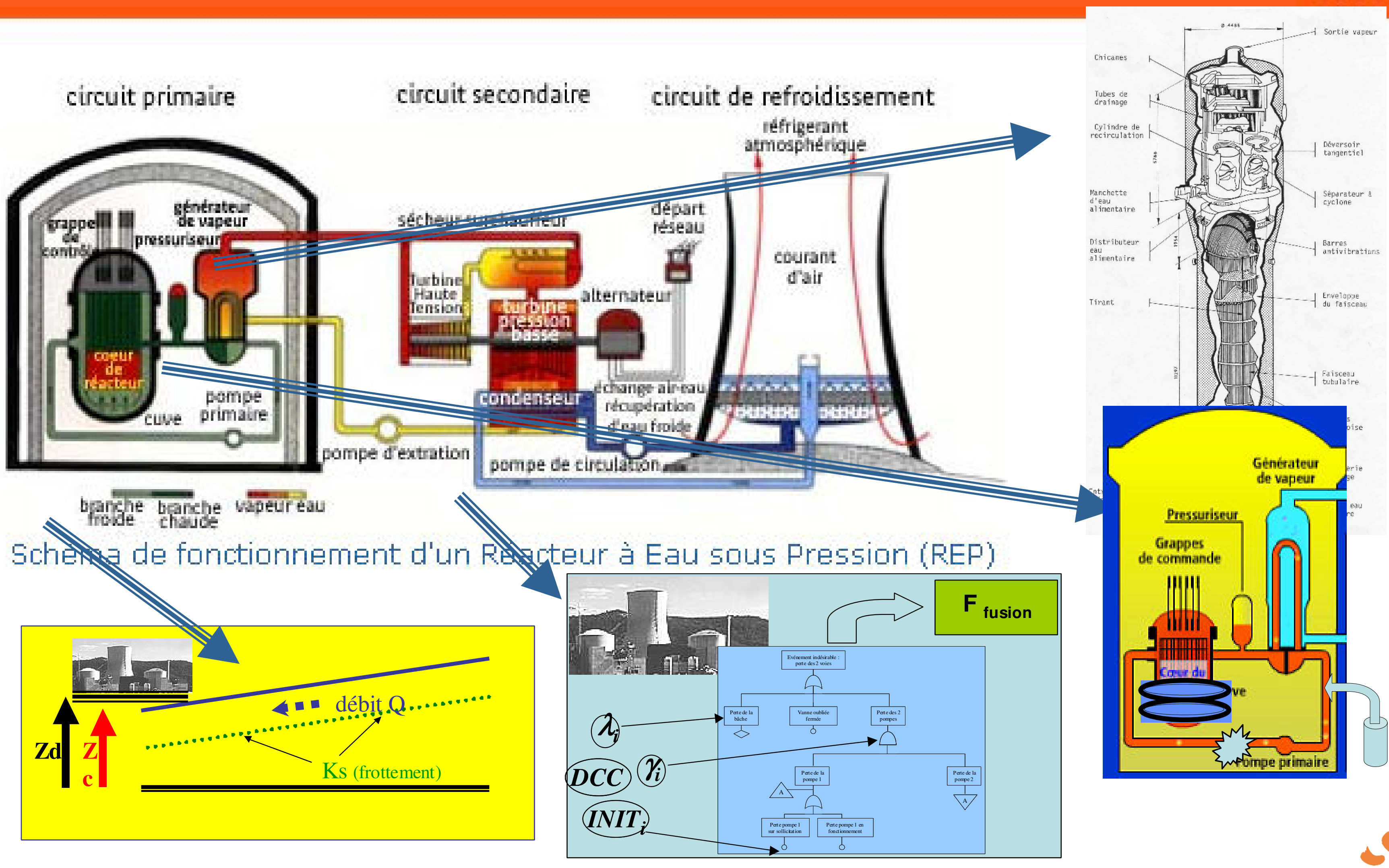
Exemples

Enjeu principal	Exemple	Domaine
Justification environnementale	**Maîtrise des émissions industrielles (e.g. émissions atmosphériques)**	Métrologie
	Impact des rejets aqueux sur l'environnement et la santé	Physico-chimie et Ecotoxicologie
Justification sûreté	**Sûreté du réacteur nucléaire – thermohydrau./mécanique**	Sûreté d'un procédé – physique probabiliste
	Protection contre les crues	Dimensionnement d'une protection
Hiérarchisation / priorités R&D	**Analyse de risque système – études proba. de sûreté (E.P.S.)**	Fiabilité des systèmes

Thermique	Nucléaire	Renouvelable / hydraulique

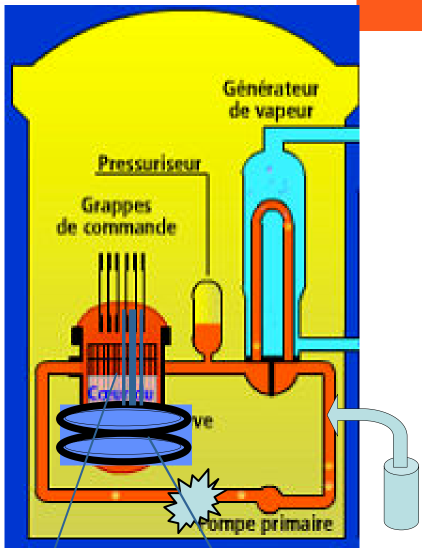
E. de Rocquigny

Exemples nucléaires



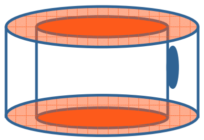
E. de Rocquigny

Sûreté du réacteur / thermo-hydraulique et mécanique
cuve



Gaine de crayon combustible

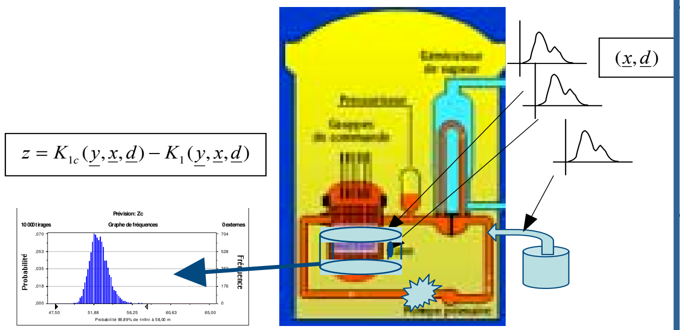
Épaisseur de la cuve



Finalité	**Justifier la sûreté du réacteur / accidents conventionnels**
Critère	Z= température gaine, rupture cuve … >> ***aujourd'hui*** marges **déterministes** $T_{/\,accident} < T_{\max gaine} \qquad (K_{1c} - K_{cp})_{\,pénalisées} > 0$ >> ***demain peut-être*** **proba.** / seuil $P(T_{/\,accident} < T_{\max gaine}) > 95\%$ $P[(K_{1c} - K_{cp}) > 0] < 10^{-x}$
Sources/ modèle	Variabilité des conditions de fonctionnement ; propriétés matériaux ; taille de défauts ; incertitudes sur coefficients thermohydrauliques etc. **Modèles** thermohydrauliques et mécaniques **complexes** (transitoires, éléments finis, …)

E. de Rocquigny

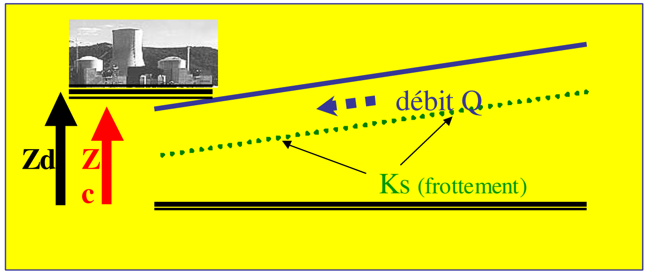
Sûreté du réacteur / thermo-hydraulique et mécanique
cuve



Critère	**Dépassement de seuil**
Quantification sources	Statistiques sur essais matériaux >> lois non gaussiennes *Méthodes inverses proba.*
Méthode calcul	*(si déterministe : un seul calcul)* Si proba : >> calculs Form-Sorm en mécanique (proba. très faibles – physique régulière) >> Monte-Carlo ou méthodes très robustes en thermo-hydraulique (proba. moins faibles – physique irrégulière)

- *Enjeu d'**évolution réglementaire** > proba./déterministe*
- *Travaux majeurs sur les **données/ sources** d'incertitudes*
- ***Méthodes** de propagation **sophistiquées** : modèles physiques lourds et proba. rares*

Protection des sites / crues



Finalité	**Justifier niveau de protection réglementé sur risque inondation**
Critère	>> ***aujourd'hui*** critère physique déterministe sur aléa probabiliste $$Z_c(\hat{Q}^{1000}{}_{pén.}) < Z_d$$ >> ***demain ?***
Sources/modèle	Aléa du débit de crue – incertitudes d'estimation de cet aléa – conditions d'écoulement variables / mal connues **Modèles** hydrodynamiques 1D ou 2D
Quantification / propagation	**Stat. extrêmes sur débit** Calculs déterministes avec sensibilités

- ***La réglementation*** *reste encore* ***semi-déterministe***
- ***Des marges importantes se cumulent***

Présentation du contexte général « Incertitudes »

Variabilité

Sources d'incertitudes

Objet « physico-industriel »

Enjeux:

Sûreté

Performance

Environnement …

(s… que)

Agressions météo aléatoires

Variabilité matériaux

Corrélations imprécises et erreurs de modèles etc…

Site de stockage déchets

Cœur réacteur

Cuve réacteur

Parc à optimiser etc…

Critère sûreté sur … Temp. gaine

Marge à la rupture mécanique

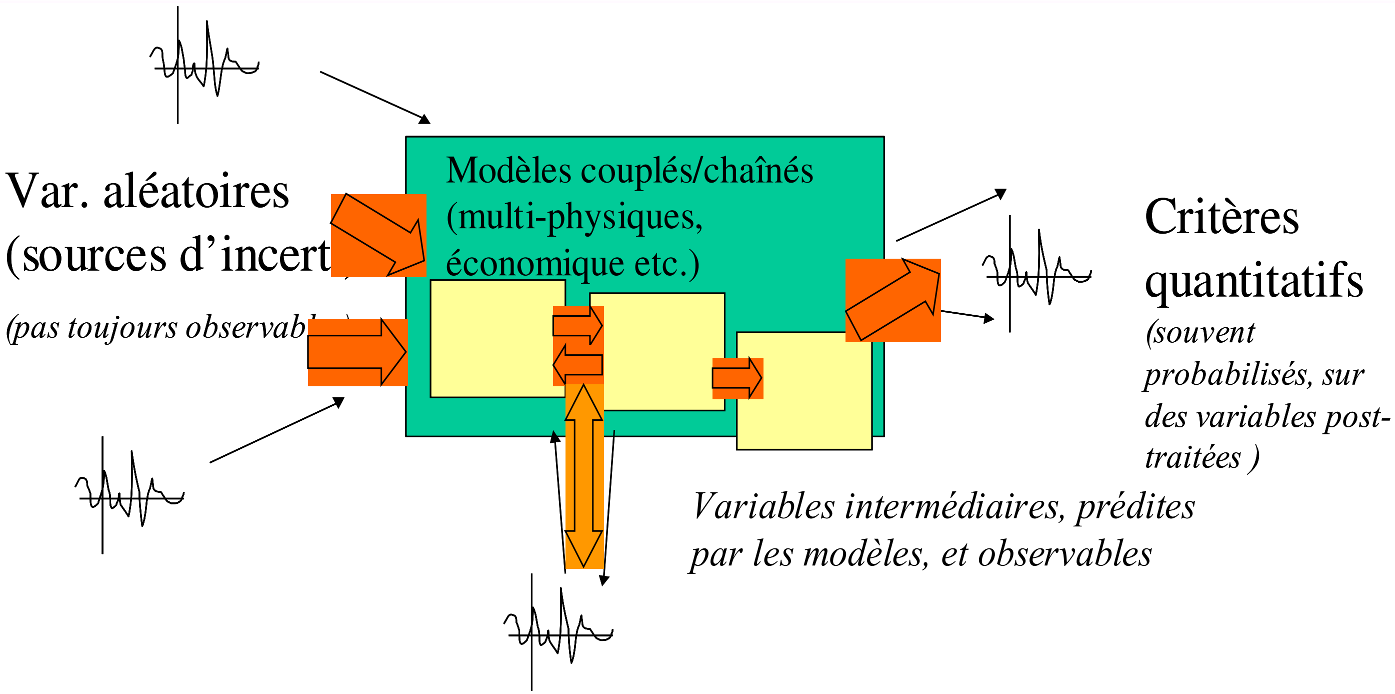
Autorisation rejet polluant etc …

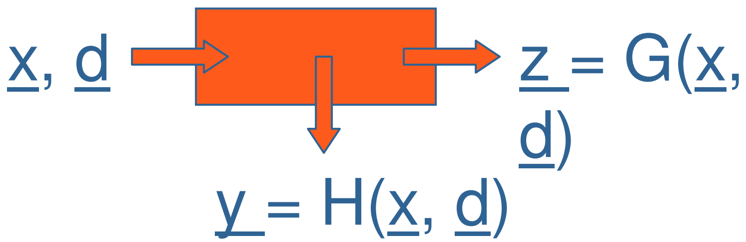
› Maîtrise des marges >> calculs d'incertitudes sur les grands codes physiques (et plus rarement sur de l'expérimental)

- *traditionnellement des marges déterministes « au jugement de l'ingénieur »*

EDF R&D

Cadre de modélisation (physico-probabiliste)





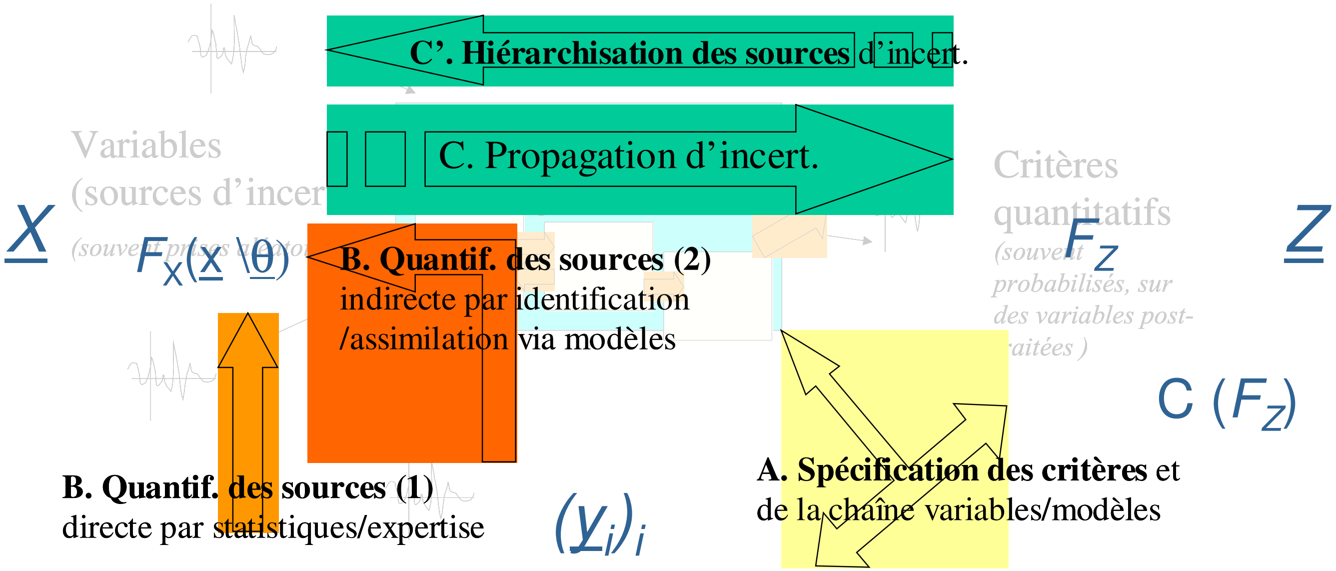
Partie déterministe : G et H sont calculés par le(s) modèle(s) physique(s) déterministe (pour x donné et d conditions connues)

Partie probabiliste :

x >> $X \sim F_X(x \backslash \theta)$ et un critère porte sur $F_Z(z)$

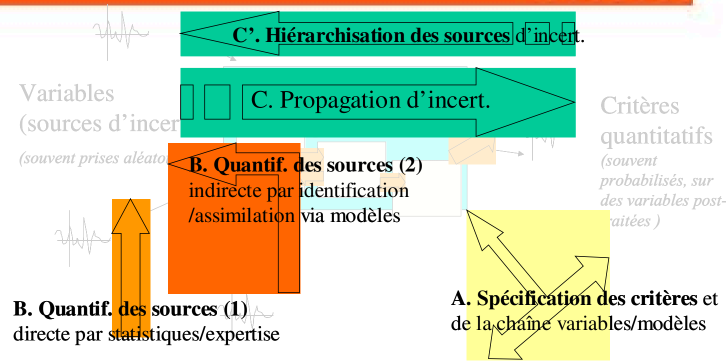
E. de Rocquigny

Les grandes étapes d'une étude d'incertitudes (2)



- *Deux finalités les plus fréquentes :*
  - ***Justification** d'un critère (l'étape C est l'étape finale)*
  - ***Hiérarchisation** des sources i.e. analyse de sensibilité (l'étape C' est l'étape finale)*
- *Egalement la **validation** (l'étape B' est finale) et l'**optimisation** (plusieurs étapes C itératives)*

Où interviennent les PE / SR ?



En simulation

- Pour la justification, en phase C
  - Accélérer la propagation … *surtout si $C_Z = P(Z > z_s)$*
  - Typiquement PE/SR puis Monte-Carlo ou PE/SR dans Form
- Pour la « hiérarchisation » (i.e. analyse de sensibilité) en phases C/C'
  - PE déterministe ou aléatoire
  - Mesures d'importance de type déterministe ou probabiliste plus ou moins complexes (Coef. de Corrélation, PRCC, Sobol …)
- Plus récemment: la validation / (méthodes inverses) et optimisation stochastique

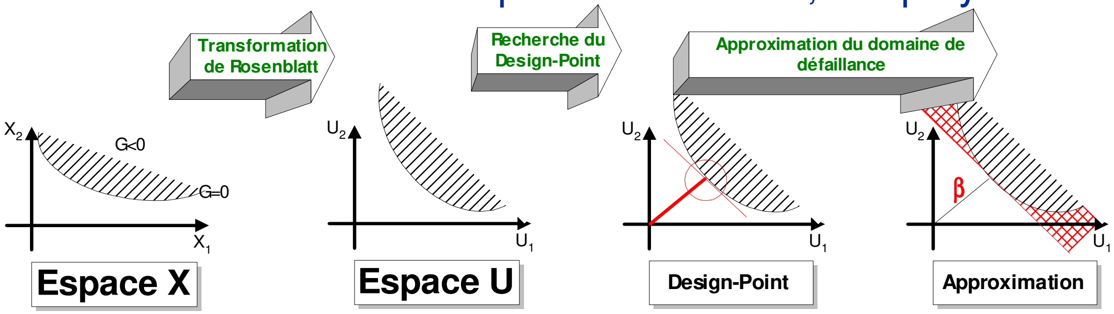
Et également dans l'expérimental: (micro-biologique, matériaux …)

Principaux PE/SR utilisés en propagation

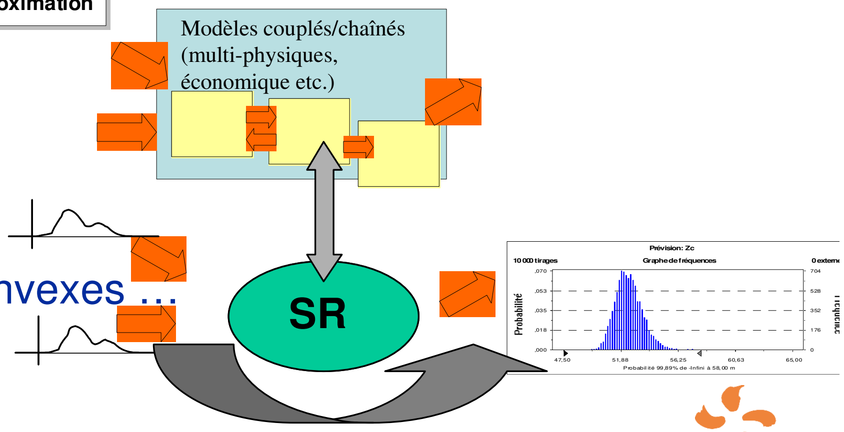
- Elémentairement dans le cumul quadratique

$$Inc_Z \approx \sqrt{\sum_{i=1}^{N} \left(\frac{\partial Z}{\partial X^i}\right)^2 Inc_{X^i}{}^{2}}$$

- Traditionnellement : accélérer l'estimation de $C_Z = P(Z > z_s)$
    - PE/SR adaptatif « naturel » dans Form/Sorm
        - dans l'espace transformé, SR polynômiale locale



- PE/SR global avant Monte-Carlo
    - dans l'espace physique
    - Globale
    - polynômial, spline, interpolants convexes …



Plus récemment, les surfaces de réponse stochastiques

Des exigences spécifiques récentes et questions de R&D

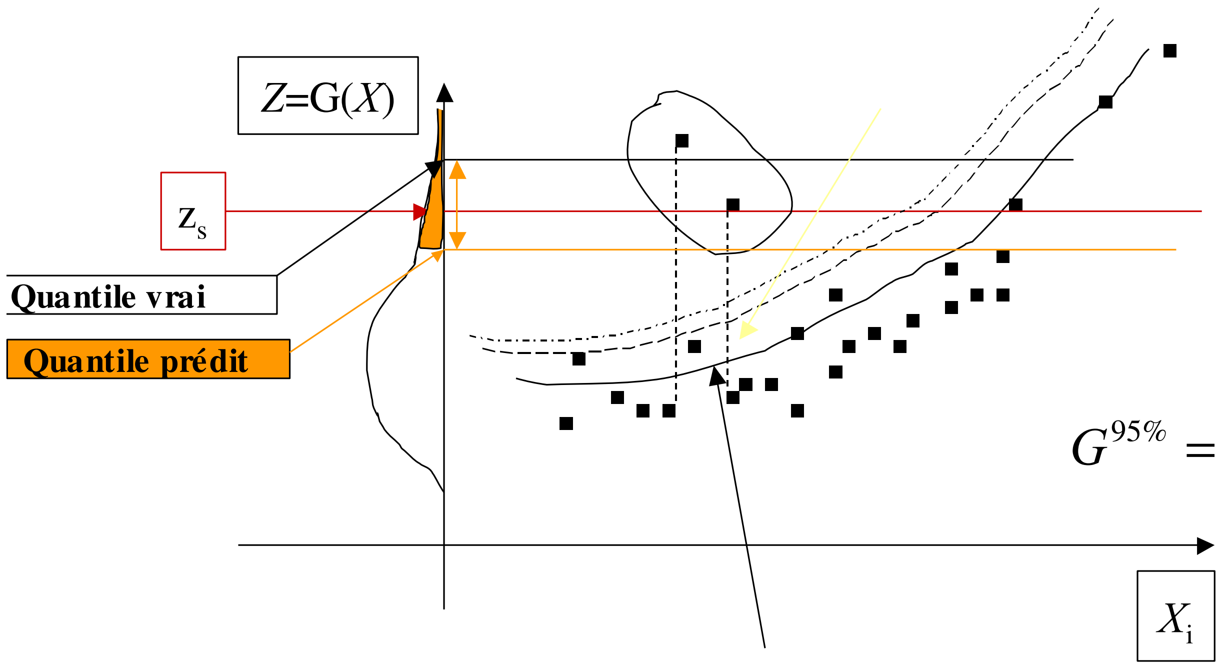
- De nombreux travaux de recherche appliquée continuent à EDF en sensibilités/incertitudes
- On exposera seulement deux points spécifiques

Des exigences spécifiques récentes
1. Robustesse des PE/SR en justification de sûreté

Justification de sûreté : estimer $C_Z = P(Z > z_s)$ de façon très robuste … par SR ?

$$G(x) = \tilde{G}(x) + u(x) = \sum_l \tilde{G}^l(x).\beta^l + u$$

Estimation $\hat{\beta}^l$ aux OMS / GLM



Prédit par la SR $\tilde{G}(x)$

$$G^{95\%} = F_Z^{-1}(.95) <?> [\tilde{G}(x) + u(x)]^{95\%}$$

E. de Rocquigny

1. Robustesse des PE/SR en justification de sûreté (suite)

Justification de sûreté : estimer $C_Z = P(Z > z_s)$ de façon très robuste … la méthode de référence : Monte-Carlo Wilks

- C'est-à-dire Monte-Carlo classique, mais à nb minimal de tirages
- En estimant l'intervalle de confiance β (non-paramétrique) sur l'estimateur du quantile α)

$$si \quad N \geq N_i(\alpha, \beta), \quad P(z^{\alpha} < Z^{n-i+1}) > \beta$$

$$P(Z^{n-i+1} < z) = \sum_{k=n-i+1}^{n} C_n^k F(z)^k (1-F(z))^{n-k}$$

		Minimum number of trials using the i-th max. of a n-sample			
α (quantile)	β (level of confidence)	Maximum $N_1(\alpha,\beta)$	2nd max. $N_2(\alpha,\beta)$	3rd max. $N_3(\alpha,\beta)$	4th max. $N_4(\alpha,\beta)$
95%	90%	45	77	105	132
95%	**95%**	**59**	**93**	**124**	**153**
95%	99%	90	130	165	198

- Wilks conduit à contrôler de façon probabiliste sûre l'erreur d'estimation du quantile dans l'étape C (i.e. erreur de propagation)

1. Robustesse des PE/SR en justification de sûreté (suite)

Contrôler l'erreur d'estimation de $C_Z = P(Z > z_s)$ avec la SR ?

- On peut imaginer calculer « un I.C. du quantile du prédit »

$$\mathrm{var}[\tilde{G}(x)+u] \approx \hat{\sigma}^2(1+\tilde{G}(x)'(\nabla\underline{\underline{\tilde{G}}}'\nabla\underline{\underline{\tilde{G}}})^{-1}\tilde{G}(x))$$

$$\tilde{G}(x)_\alpha^\beta \approx [\tilde{G}(x)+\phi^{-1}(\beta)\sqrt{\mathrm{var}\ \mathit{prédit}(x)}]_\alpha$$

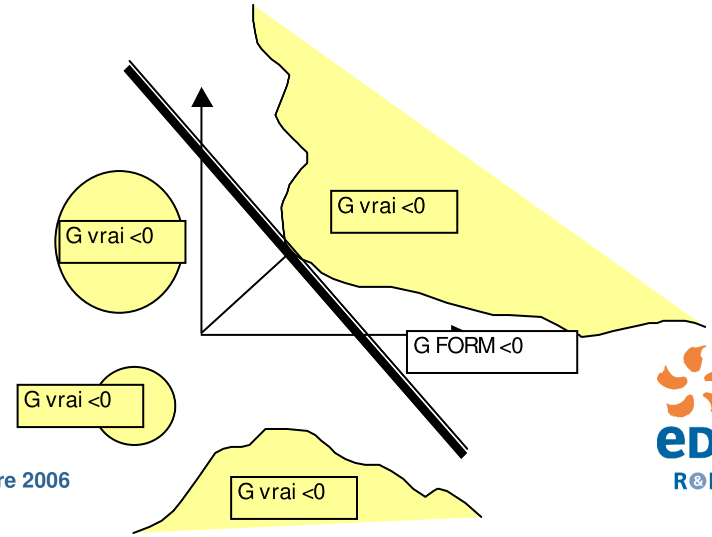
- Mais, en SR traditionnelle

  -… hypothèse de résidu gaussien non argumentée

  - … meilleure stat. de u >> beaucoup de simulations nécessaires

N.B. Form/Sorm (avec PE/SR adaptatif)

est une approx. peu contrôlable …



Premières pistes pour réduire la variance

Premières pistes pour réduire la variance de l'estimateur en combinant Wilks et une SR

- Réduction par formulation asymptotique sous indépendance (« réduction indépendante »)
- Addition de quantile sous inégalité forte

E. de Rocquigny

Utilisation d'une SR

On utilise typiquement une surface de réponse …
d'abord polynomiale d'ordre faible par régression
(mais on pourrait changer) :

$$Y=G(X) = Y_{SR}*\beta + \varepsilon_{SR}$$

$$=G_{SR}(X)*\beta + \varepsilon_{SR}$$

pour prédire un quantile de Y …en manipulant
l'« addition de quantiles » de $G_{SR}(X)$ et ε_{SR}

E. de Rocquigny

Formule asymptotique d'addition de quantiles sous hypothèses d'indépendance

On montre que :

$$\mathbb{P}\left(U+\sigma V \leq q\right) - \sum_{k=0}^{p} \frac{(-1)^k \mathbb{E}\left(V^k\right) F_U^{(k)}(q)}{k!} \sigma^k \underset{\sigma \to 0}{\sim} \frac{(-1)^{p+1} \mathbb{E}\left(V^{p+1}\right) F_U^{(p+1)}(q)}{(p+1)!} \sigma^{p+1}$$

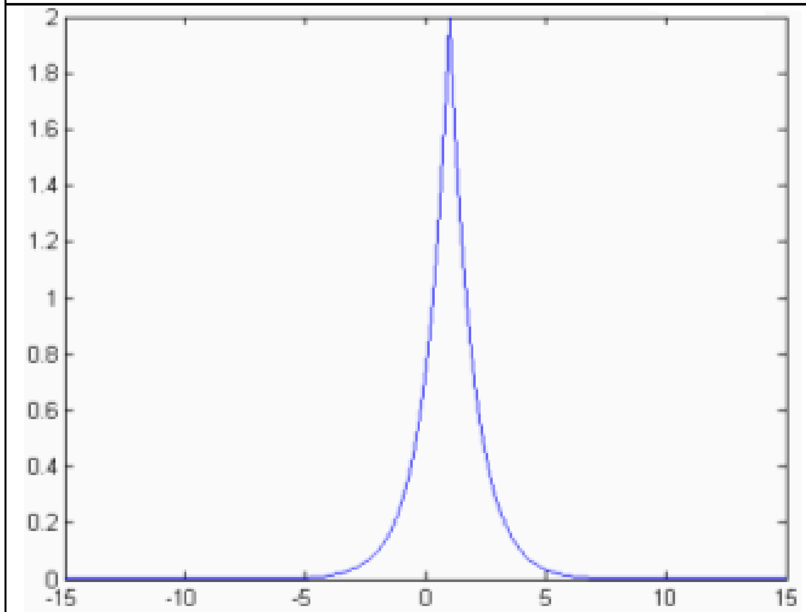
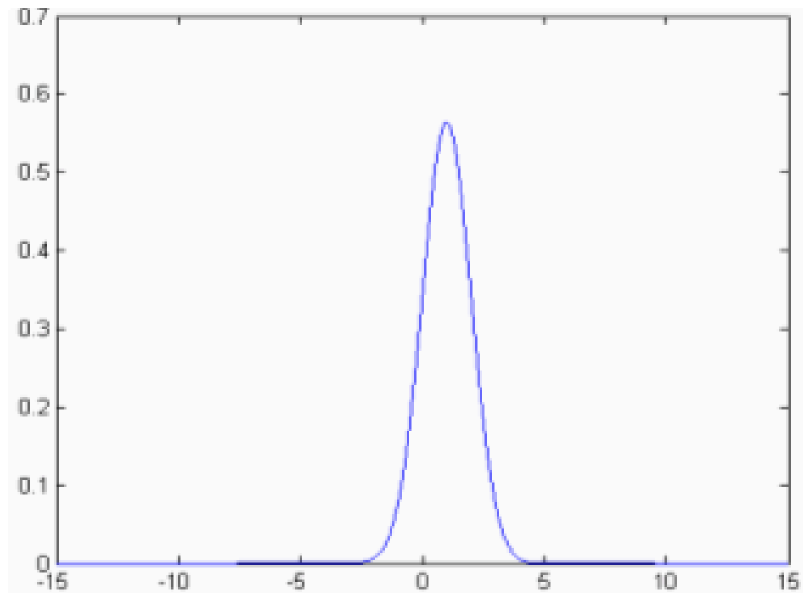
Pour avoir le quantile, on inverse la formule et on obtient :

$$F_{U+\sigma V}^{-1}(\alpha) \underset{\sigma \to 0}{=} F_U^{-1}(\alpha) - \mathbb{E}(V)\sigma + \left(\frac{1}{2}\mathbb{E}\left(V^2\right) - \left(\mathbb{E}(V)\right)^2\right) \frac{F_U'' \circ F_U^{-1}(\alpha)}{F_U' \circ F_U^{-1}(\alpha)} \sigma^2 + o\left(\sigma^2\right)$$

> U sera la SR et V le résidu (réduit) …

E. de Rocquigny

Tests numériques sur des modèles 1D physiquement « irréguliers » …

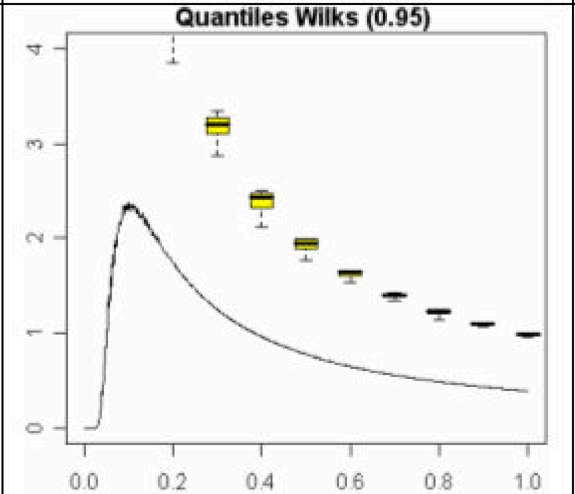
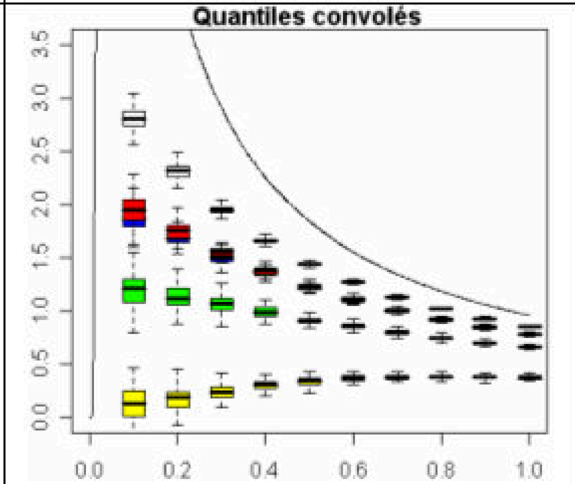
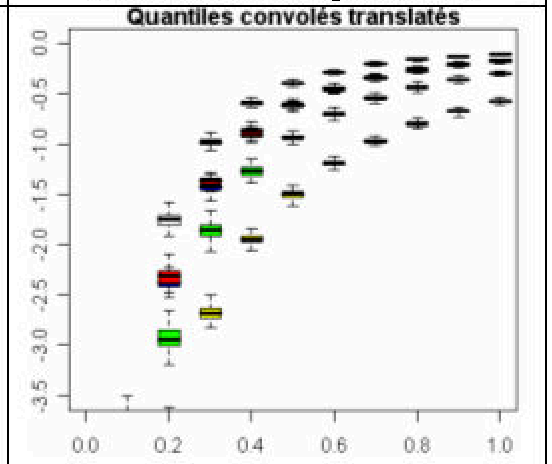
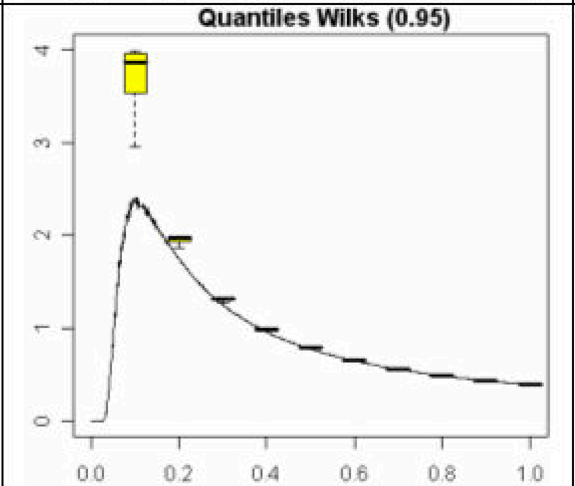
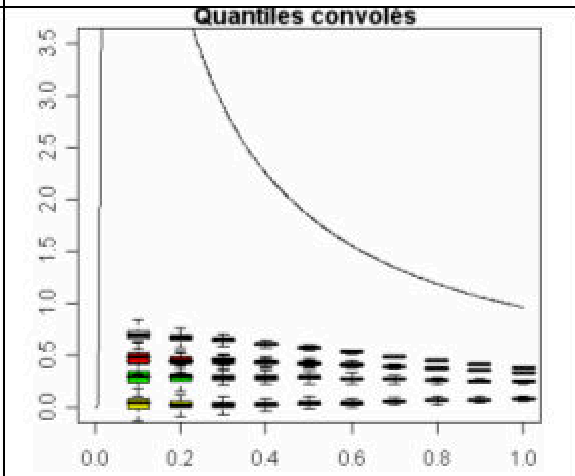
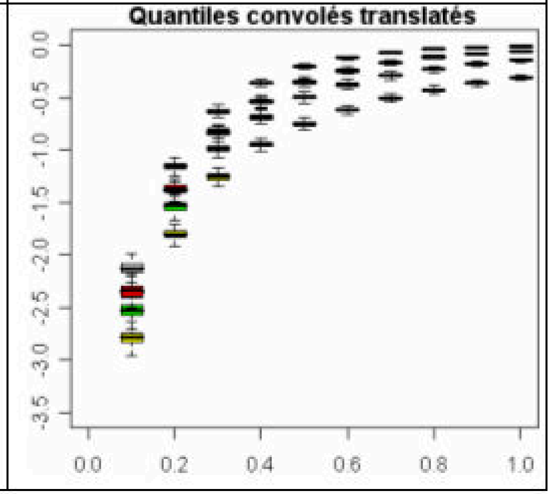
$x \mapsto \frac{1}{\lambda} e^{-\frac{1}{2}\left\vert\frac{x-1}{\lambda}\right\vert}$	$x \mapsto \frac{1}{\lambda\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-1}{\lambda}\right)^2}$
	

La variable aléatoire d'entrée X est une gaussienne centrée réduite.

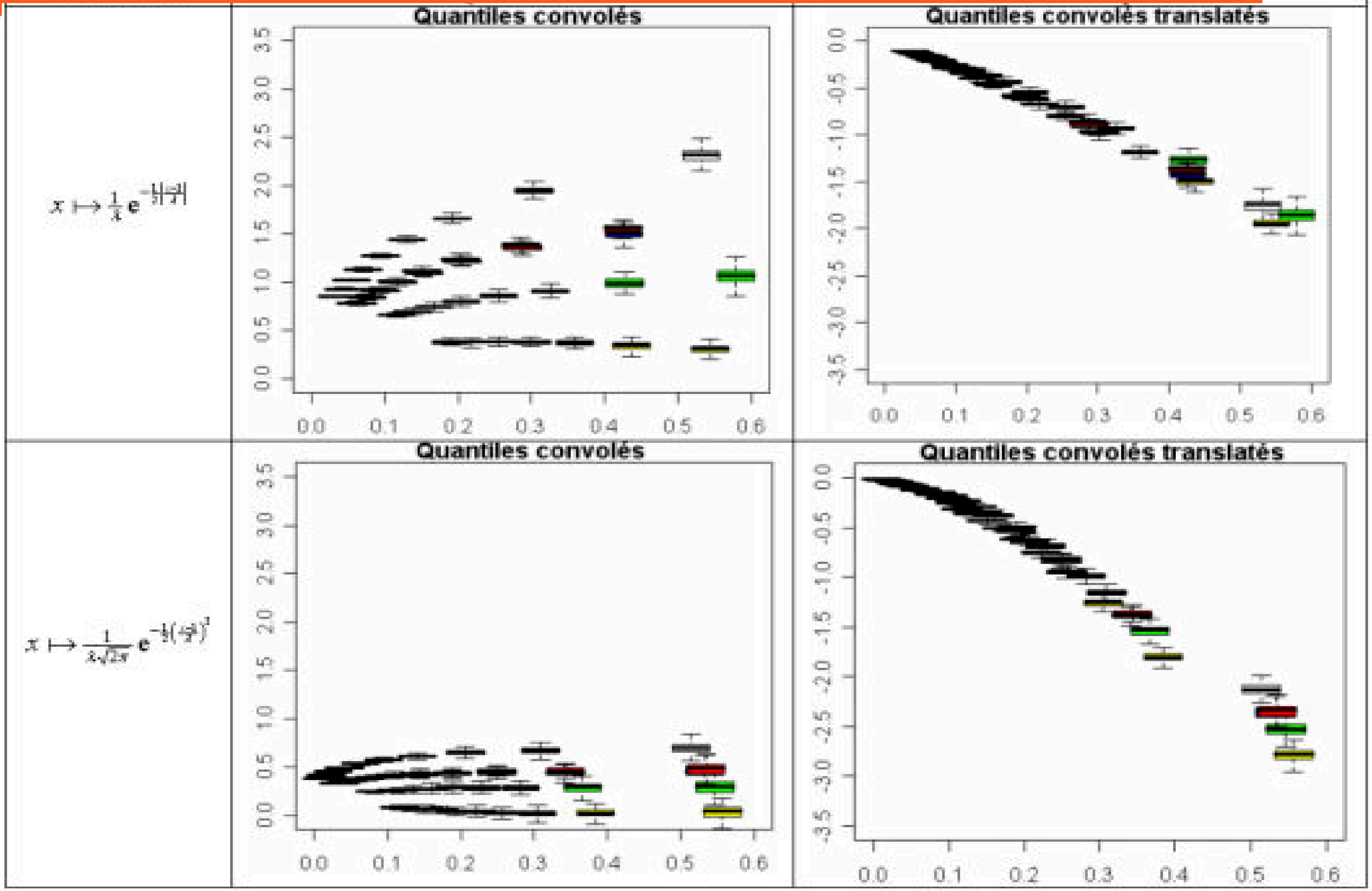
Le paramètre λ des fonctions caractérise leur aplatissement / leur pic.

E. de Rocquigny

Résultats, selon λ *en comparaison à Wilks*

Fonction :	Quantile Wilks :	Quantiles :	Translation des quantiles :
$x \mapsto \frac{1}{\lambda} e^{-\frac{1}{2}\left\lvert\frac{x-1}{\lambda}\right\rvert}$			
$x \mapsto$ $\frac{1}{\lambda\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-1}{\lambda}\right)^2}$			

Résultats, selon σ



E. de Rocquigny

Conclusion provisoire sur 1ère formule asymptotique

- Cette méthode n'est plus conservative
- … la correction résiduelle – par formule asymptotique - ne « rattrape » pas l'erreur de SR en quantile
- … et l'hypothèse d'indépendance SR / résidu est peut-être également en cause

Deuxième piste : quantile par SR, mais avec une inégalité conservative

On utilise encore une variable de contrôle de type surface de réponse polynomiale calculée par régression linéaire :

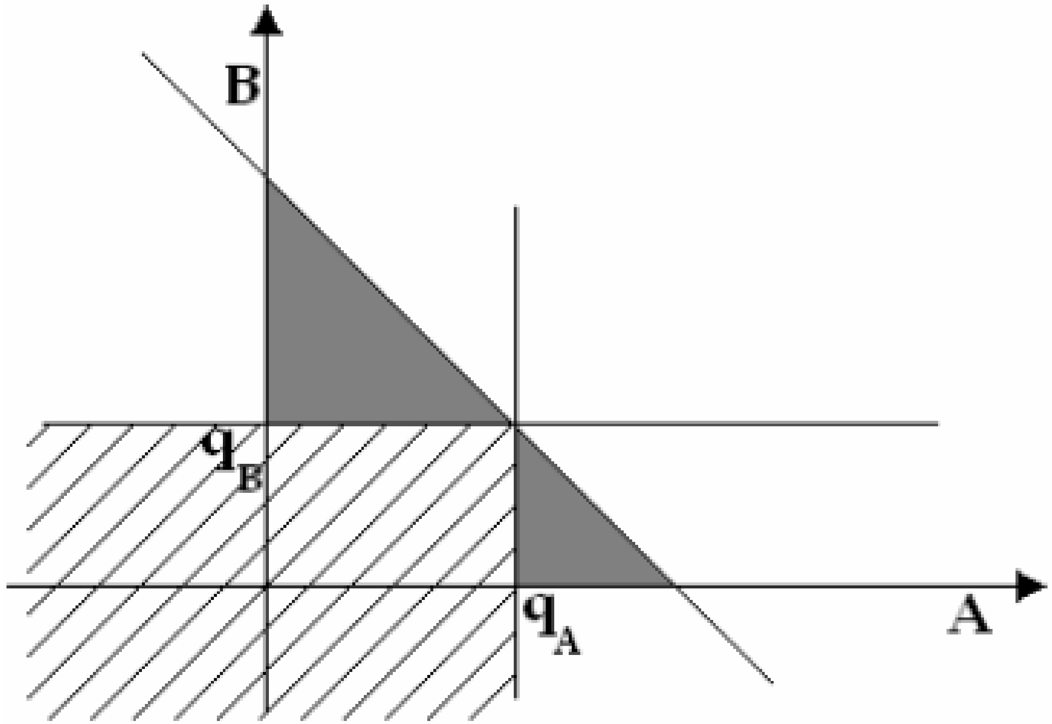
$$Y = G(X) = Y_{SR} * \beta + \varepsilon_{SR}$$

$$= G_{SR}(X) * \beta + \varepsilon_{SR}$$

Mais l'indépendance (entre $G_{SR}(X)$ et ε_{SR}) n'est pas ici le facteur primordial.

E. de Rocquigny

Une inégalité élémentaire … mais certaine, sur l'addition de quantiles …



$$\{A \le q_A\} \cap \{B \le q_B\} \subset \{A + B \le q_A + q_B\}$$

E. de Rocquigny

Inégalités probabilistes (1)

En prenant la probabilité de cette inclusion ensembliste, on a alors :

$$\mathbb{P}\left(A+B\le q_\alpha(A)+q_\beta(B)\right)=F_{A+B}\left(q_\alpha(A)+q_\beta(B)\right)$$
$$\ge \mathbb{P}\left(\{A\le q_\alpha(A)\}\cap\{B\le q_\beta(B)\}\right)$$

… que l'on peut interpréter en « copule »

$$\forall(u,v)\in[0,1]^2\ C_{(A,B)}(u,v)=\mathbb{P}\left(\{A\le q_u(A)\}\cap\{B\le q_v(B)\}\right)$$
$$=\mathbb{P}\left(A\le q_u(A),B\le q_v(B)\right)$$
$$=F_{(A,B)}\left(F_A^{-1}(u),F_B^{-1}(v)\right).$$

Inégalités probabilistes (2)

La fonction $C_{(A,\ B)}$ étant la fonction de copule du couple $(A,\ B)$

Elle vérifie les inégalités fondamentales :

$$C_{(A,B)}(\alpha,\beta) \in \left[\alpha+\beta-1, \min(\alpha,\beta)\right]$$

Inégalités probabilistes (3)

$$q_\alpha(A) + q_\beta(B) \geq q_{C(\alpha,\beta)}(A+B) \geq q_{\alpha+\beta-1}(A+B)$$

$$q_\alpha(Y) = q_\alpha(Y_{SR} + \sigma\varepsilon_{SR}) \leq q_\delta(Y_{SR}) + q_{1+\alpha-\delta}(\sigma\varepsilon_{SR})$$

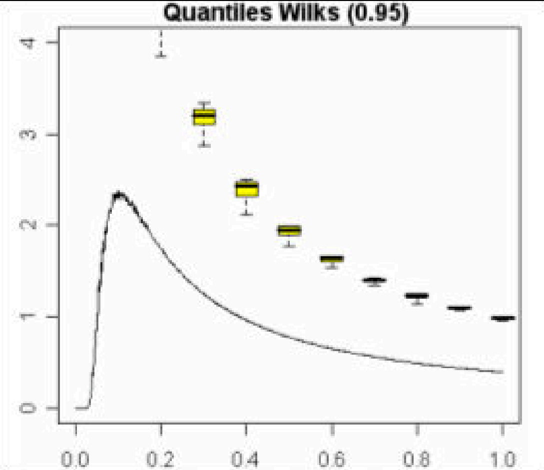
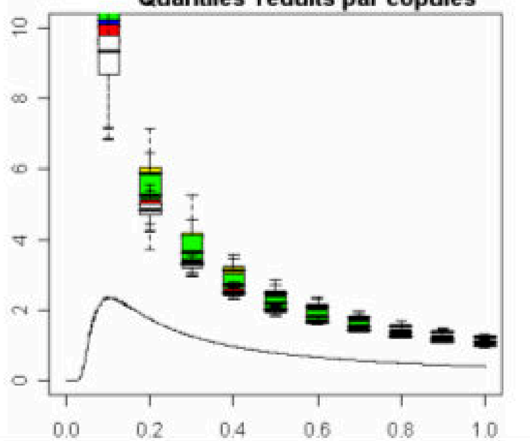
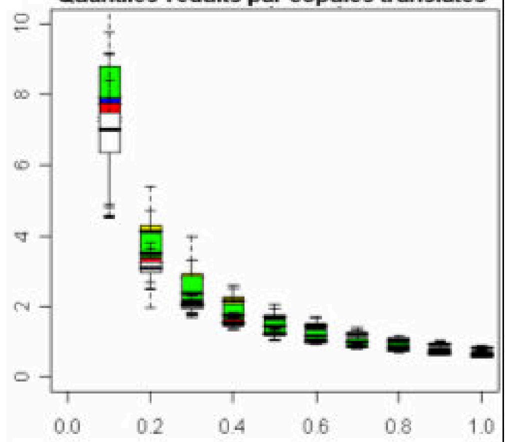
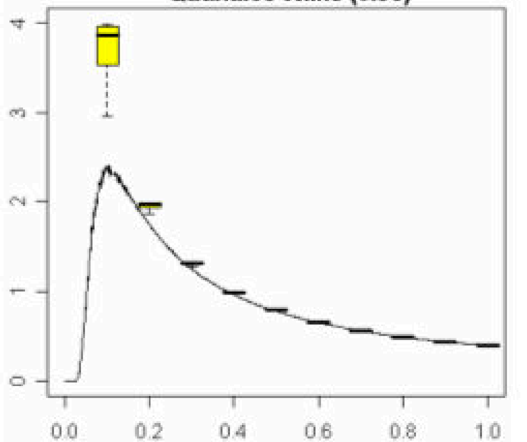
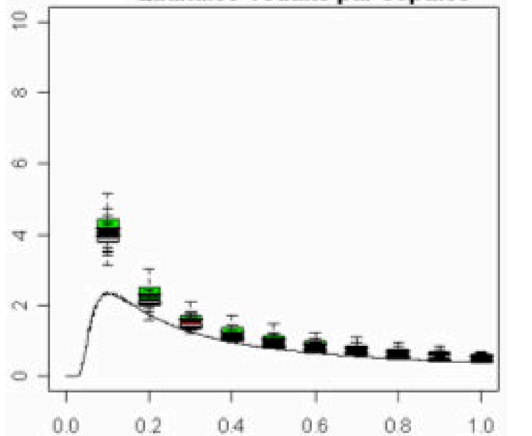
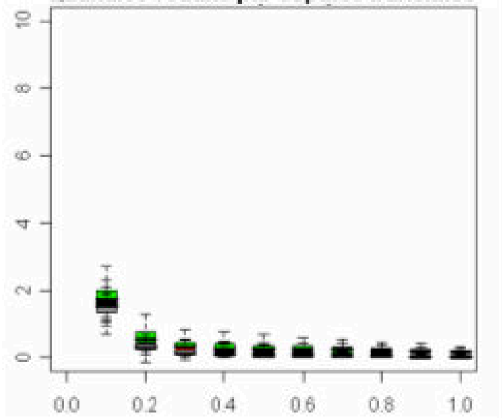
$$\leq q_\delta(Y_{SR}) + \sigma q_{1+\alpha-\delta}(\varepsilon_{SR})$$

$$\widehat{Y^\alpha} = \overbrace{(Y_{SR})^\delta}^{(MC)} + \overbrace{(\sigma\varepsilon_{SR})^{1+\alpha-\delta}}^{(Wilks,\beta)}$$

On teste alors l'estimateur sur les fonctions précédentes.

E. de Rocquigny

Résultats, selon λ *en comparaison à Wilks*

Fonction :	Quantile Wilks :	Quantile réduit par copules :	Translation du quantile :
$x \mapsto \frac{1}{\lambda} e^{-\frac{1}{2}\left\lvert\frac{x-1}{\lambda}\right\rvert}$			
$x \mapsto$ $\frac{1}{\lambda\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-1}{\lambda}\right)^2}$			

E. de Rocquigny

Conclusion provisoire sur 2ème piste / inégalité conservative

- Cette méthode est trop conservative …
- Mieux l'estimer suppose par exemple de bien estimer la copule entre la SR et son résidu … ce qui coûte cher pour être aussi robuste que Wilks … vu qu'aucune famille de copule n'a de raison *a priori* d'être « la bonne » …

Approfondissements …

- ➢Mieux gérer la variable de contrôle …(cf. exposé Cemracs)
- ➢Exploiter les considérations physiques …

Robustesse des PE/SR en justification de sûreté (suite)

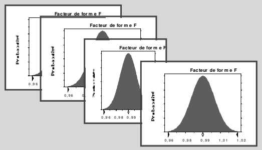
Des idées à creuser : valoriser connaissances physiques sur G (et donc $\tilde{G}(x)$ et u) …

- Un premier exemple : monotonie partielle et approximation de l'une des sous-briques du modèle physique couplé

- Au-delà : modèle physico-stat. sur le résidu (a priori sur zones non régulières etc.)& choix physique des fonctions de base de la SR … et optimisation des PE

NB : les SR n'ont plus à être régulières … un modèle physique simple peut être bien meilleur qu'une autre approx. Fonctionnelle (RN …) … les A, D, …-optimalités des PE encore pertinentes ?

Un cas physique où la monotonie partielle rend une SR constante par morceaux conservative



Vecteur aléatoire d'incertitudes ($\underline{u}$)

u_Φ , $u_{RT\,NDT}$ et u_{KIc} gaussiennes

T_{is} (plusieurs lois c[7°,40°])

Calcul (analytique) du transitoire
Tfluide(t, T_{is}), H(t, T_{is})

Code_Aster : calcul T(x,t), σ(x,t)

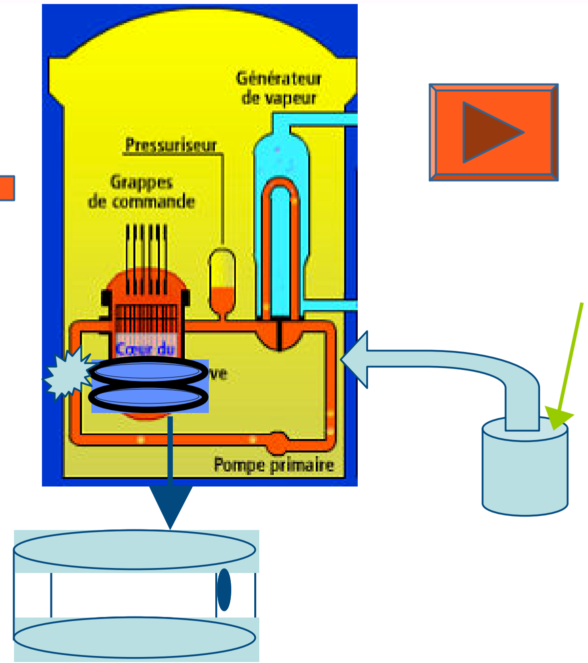
Calcul marge instantanée à rupture

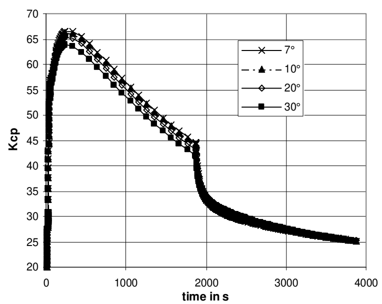
$G(t, \underline{u}) = K1c\,(T_B(t), \underline{u}) - Kcp(T(x,t), \sigma(x,t))$

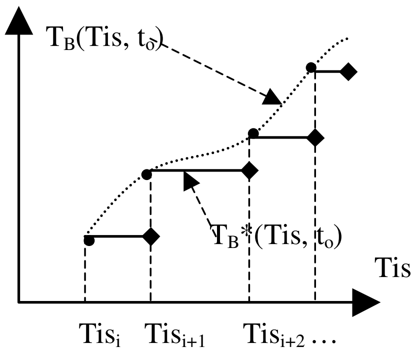
N.B. Temps calcul global $G_{min}(\underline{u})$ (pour une valeur de $\underline{u}$) ~1-10 mn CPU

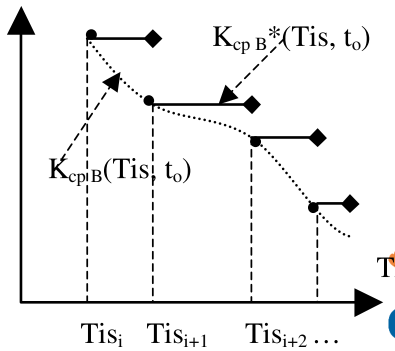
$G_{min}(\underline{u})=Min_t\,(G(t, \underline{u}))$

Défaillance = « risque global » = $\{G_{min}(\underline{u})<0\}$









E. de Rocquigny