



Calibration bayésienne et prédiction de réglages

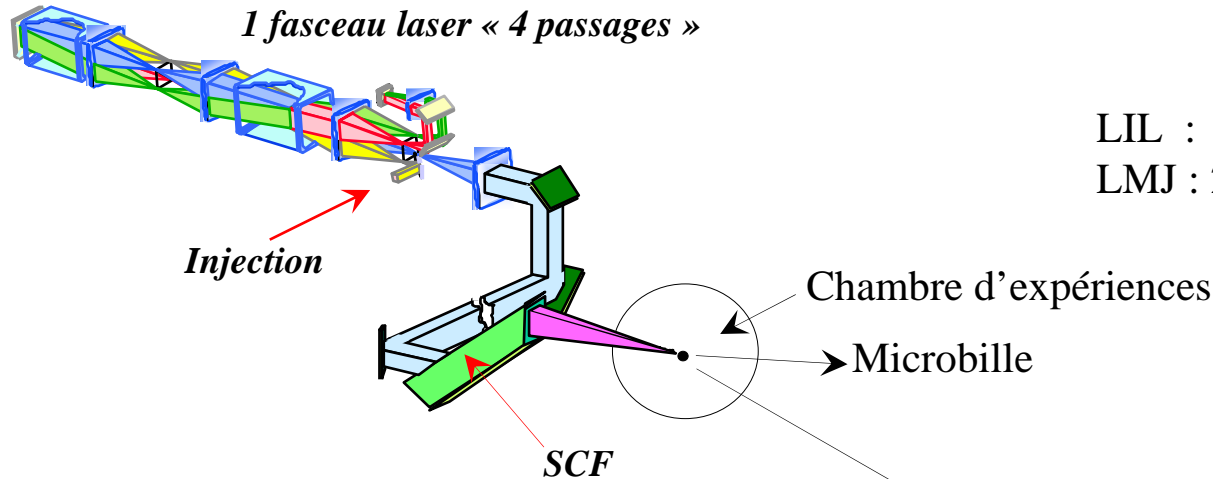
Marc Sancandi
CEA/CESTA

Séminaire « Incertitudes et Simulation », DIF 3-4 Octobre 2007

Qu'est-ce que la « *prédiction de réglage* » ?

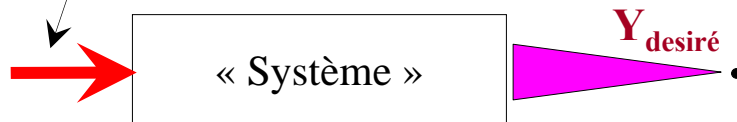


1 faisceau laser « 4 passages »



LIL : 4 faisceaux laser
LMJ : 240 faisceaux laser

Réglages :
Que mettre ici pour obtenir $Y_{\text{désiré}}$?



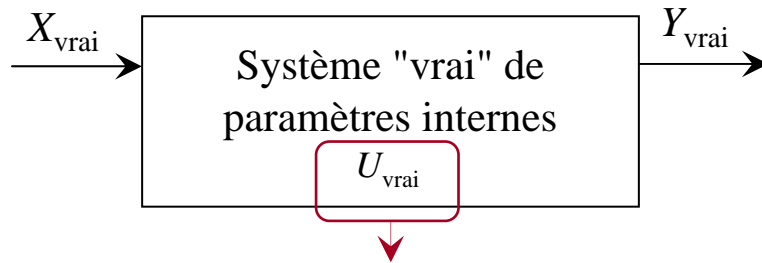
Spécifications du physicien :

- Puissance = $P(t)$
- Energie
- Isotropie d'éclairement
- ...

Pour prédire les réglages on doit connaître le système !!!

L'approche adoptée repose sur l'utilisation de codes de calculs et sur l'exploitation des résultats issus d'expériences passées (spécifiques ou « de physique »)

Système vrai et système apparent

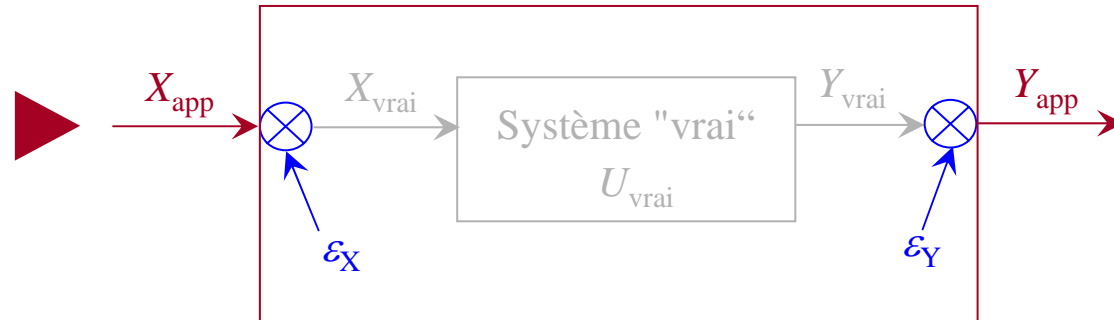


$$\text{Système "vrai" : } Y_{\text{vrai}} = \mathbf{F}_{\text{vrai}} (X_{\text{vrai}} \mid U_{\text{vrai}})$$

Paramètres « fixes » mais non mesurés

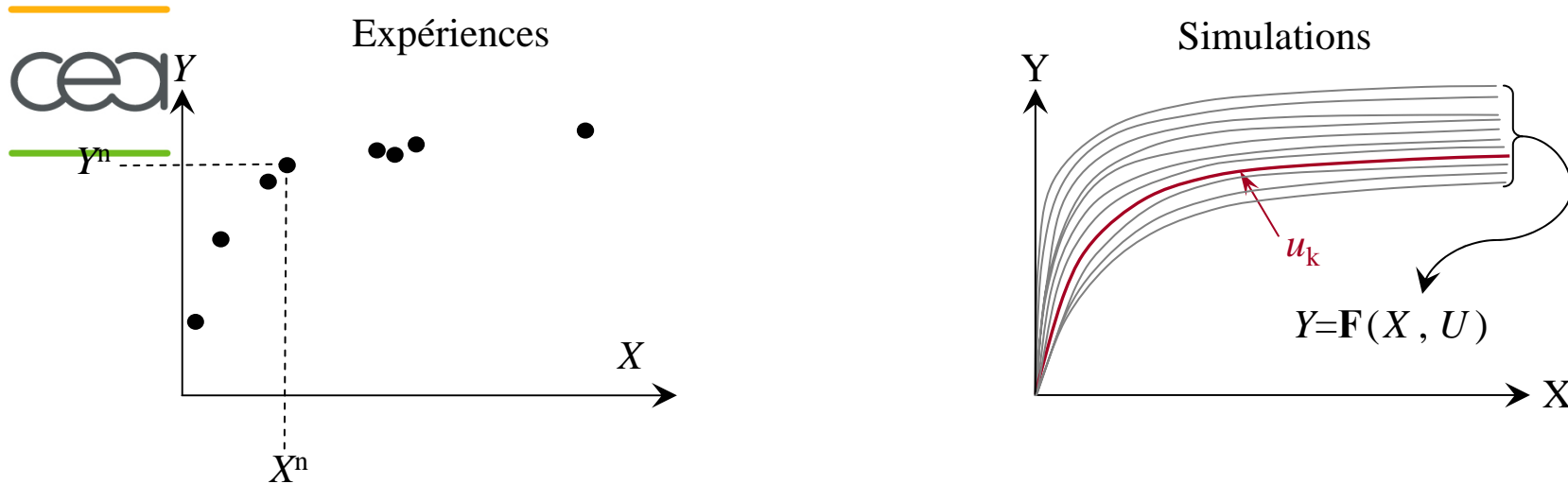
$$\text{Erreurs de mesure : } X_{\text{app}} = X_{\text{vrai}} + \varepsilon_X \quad ; \quad Y_{\text{app}} = Y_{\text{vrai}} + \varepsilon_Y$$

$$\text{Système "apparent"} \\ Y_{\text{app}} = \mathbf{F}_{\text{app}} (X_{\text{app}})$$

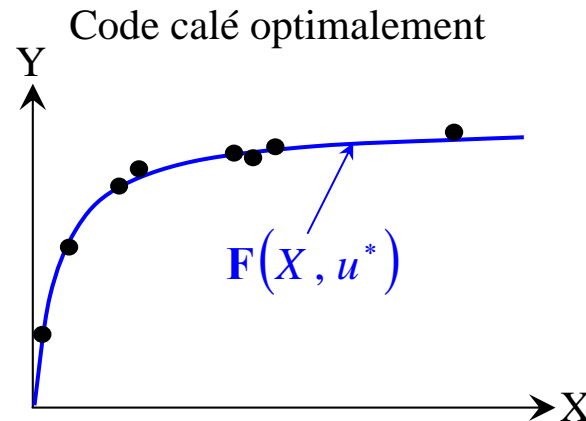


Connaître le système vrai c'est estimer U_{vrai} à partir de X_{app} et Y_{app} .

Ajustement d'un code de calcul : l'approche conventionnelle



$$\text{Minimiser } : R(u) = \sum_{n=1}^N \left(Y_{\text{app}}^n - F(X_{\text{app}}^n, u) \right)^2 \rightarrow u^*$$





Question : Quelle confiance accorder à la valeur optimale u^* ?

- Quel aurait été le résultat si on avait eu des données expérimentales différentes, issues par exemple de répétitions des N expériences initiales ?
- Est-on, au moins, certain que le code \mathbf{F} représente fidèlement la réalité : $\mathbf{F} \stackrel{?}{=} \mathbf{F}_{\text{vrai}}$

On peut apporter des éléments de réponse ...

1. En faisant ces répétitions → cher → donc exclu !!!
2. En procédant par rééchantillonnage :
On estime u^ sur une partie des essais et on estime sa « qualité » sur la partie restante*
3. En simulant numériquement ce que qu'auraient pu donner ces répétitions onéreuses



Algorithme C

Hypothèse : $\mathbf{F} = \mathbf{F}_{\text{vrai}}$

Choisir M "grand"

Pour $m = 1$ à M faire :

Choix aléatoire de $\varepsilon_X^1, \dots, \varepsilon_X^N, \varepsilon_Y^1, \dots, \varepsilon_Y^N$

$$u_{\text{app}}^{*,m} = \text{argmin} \sum_{n=1}^N \left(Y_{\text{app}}^n + \varepsilon_Y^n - \mathbf{F} \left(X_{\text{app}}^n + \varepsilon_X^n, u \right) \right)^2$$

Fin Pour

SIMPLEX (Nelder & Mead 1965)

$$u_{\text{app}}^{\text{initial}} = u_{\text{vrai}}$$

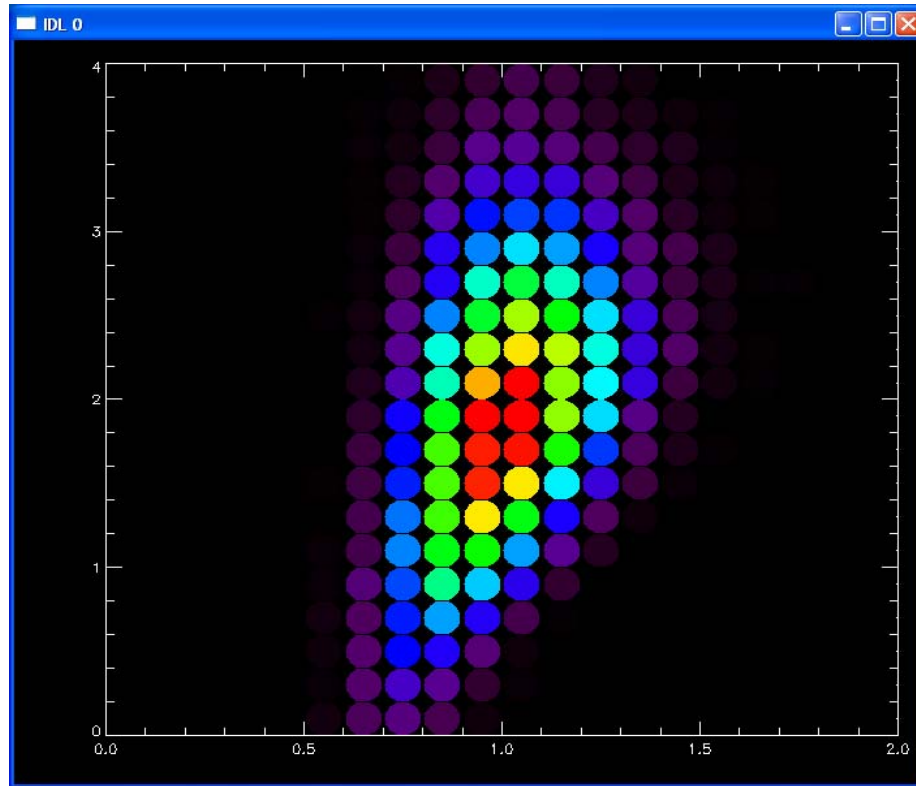
solution : u_{app}^*

$$Y = a \exp(-a \cdot b \cdot X)$$

$$\varepsilon_X \sim N(0, 0.2) \quad ; \quad \varepsilon_Y \sim N(0, 0.2)$$

Algorithme C

Distribution de $u_{\text{app}}^{*,m}$



Pour tous les exemples qui suivent :

$$N = 5$$

$$X_{\text{vrai}} = 0, 1, 2, 3, 4$$

$$u_{\text{vrai}} = (a_{\text{vrai}}, b_{\text{vrai}}) = (1, 2)$$

MA, MB 1.006 1.961
SA, SB, 0.179 0.870

Avantages et inconvénients de l'Algorithme C



Avantages :

- Simple à mettre en œuvre
- Permet de construire la distribution des solutions optimales

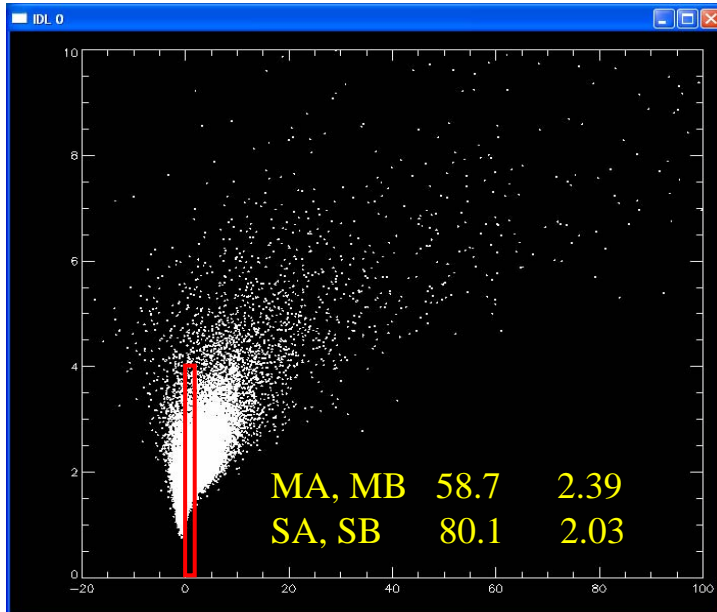
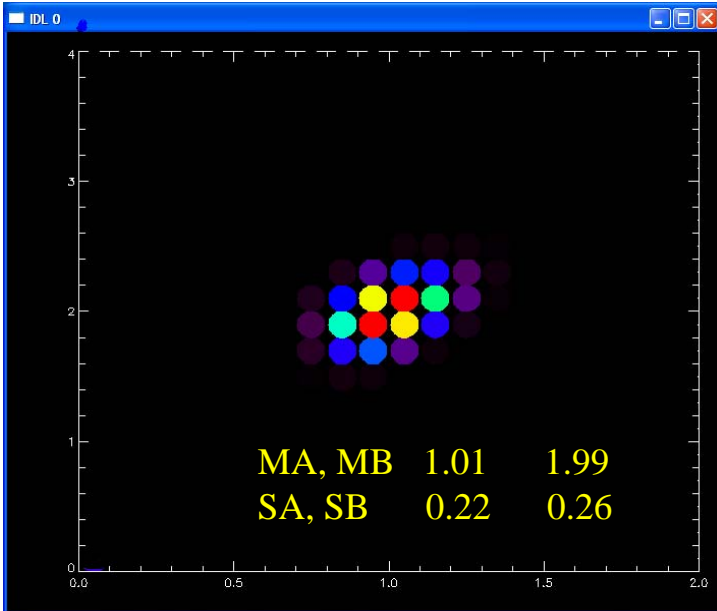
Inconvénients :

- Peut être « cher » (M problèmes de minimisation à résoudre)
- Et si le problème à plusieurs minima ?

Remarques :

- Cet algorithme ne garantit pas que la valeur moyenne de $u_{\text{app}}^{*,m}$ converge vers U_{vrai} lorsque M tend vers l'infini !
- En particulier car il ne tient pas compte d'informations a priori sur les valeurs les plus probables de U_{vrai}

Algorithmme C



$$Y = (a \cdot X + b) \exp(-k \cdot b \cdot X)$$

$$a \sim N(1, 1/3) \quad ; \quad b \sim N(2, 2/3)$$

Haut : $k = 0.15$

Bas : $k = 1$

Solution donnant le plus petit résidu

$\min \sqrt{\frac{R(a,b)}{N}}$	a^*	b^*
0.066	3.62	2.80

Solution la plus proche de la valeur vraie

a^*	b^*	$\sqrt{\frac{R(a,b)}{N}}$
1.0004	2.0075	0.555

Rappel : $\sigma_Y = 0.2$

Ajustement d'un code de calcul : l'approche bayésienne



... ou « Calibration Bayésienne »

Algorithme B

Hypothèse : $\mathbf{F} = \mathbf{F}_{\text{vrai}}$

Choisir M "grand"

Pour $m = 1$ à M faire :

Choix aléatoire de u^m

Pour $n = 1$ à N faire :

Calculer $\bar{Q}(\mathcal{H}_m^n \mid \mathcal{D}^n)$

Fin Pour

Calculer $\bar{Q}(\mathcal{H}_m \mid \mathbf{ID}^N)$

Fin Pour

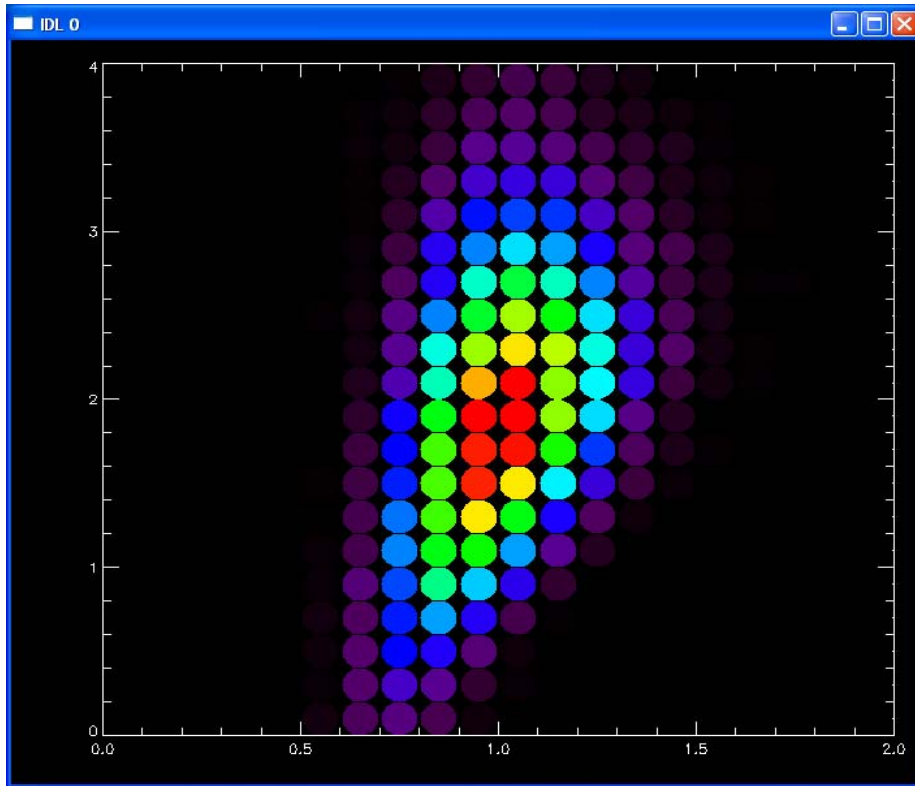
Exemple d'utilisation

$$\bar{u}_{\text{Bayes}} = \frac{\sum_{m=1}^M u^m \bar{Q}(\mathcal{H}_m \mid \mathbf{ID}^N)}{\sum_{m=1}^M \bar{Q}(\mathcal{H}_m \mid \mathbf{ID}^N)}$$

$$Y = a \exp(-a \cdot b \cdot X)$$

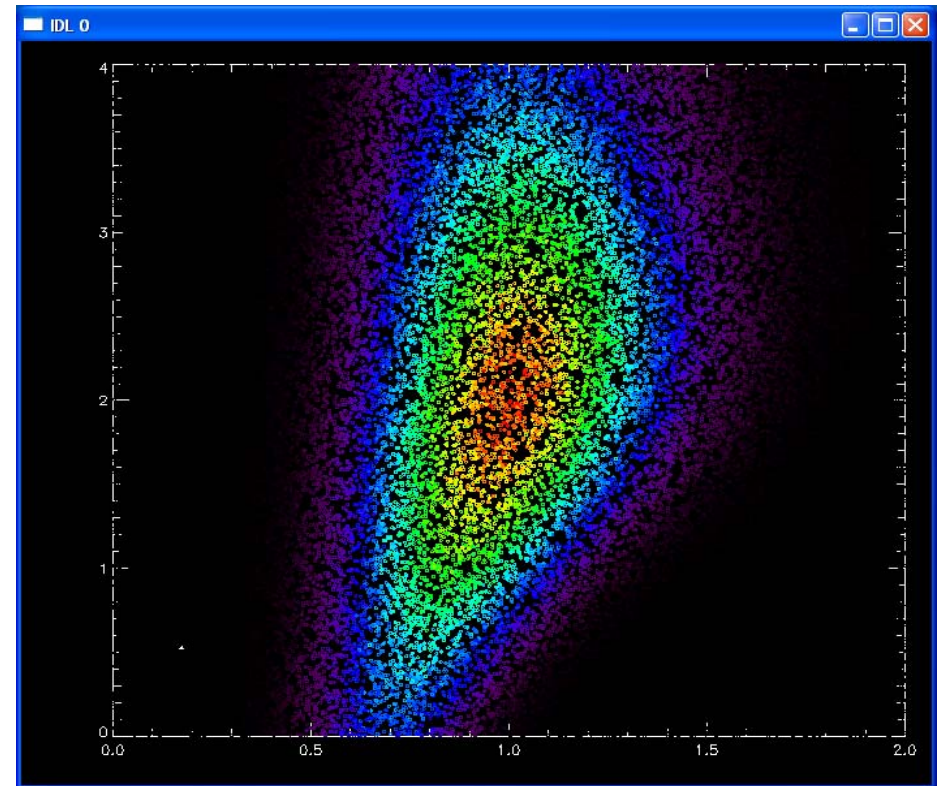
$$\varepsilon_X \sim N(0, 0.2) \quad ; \quad \varepsilon_Y \sim N(0, 0.2)$$

Algorithme C



MA, MB	1.006	1.961
SA, SB,	0.179	0.870

Algorithme B

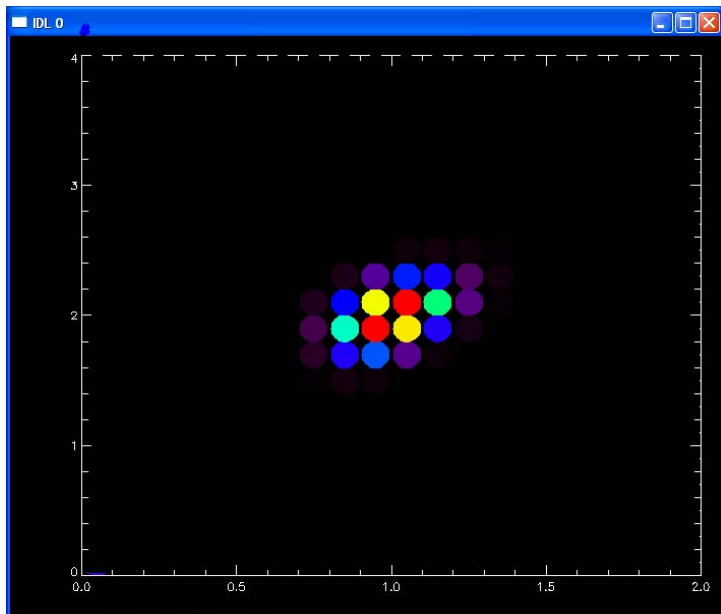


MA, MB	0.991	2.129
SA, SB,	0.235	0.935

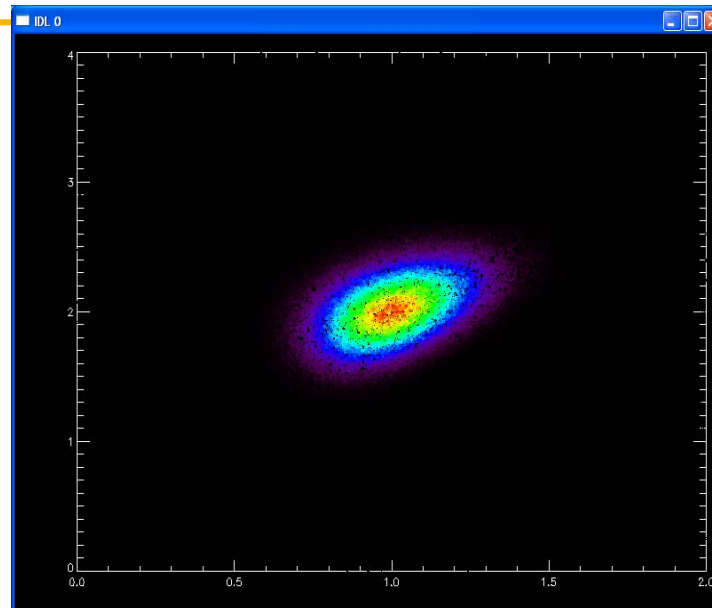
$$Y = (a \cdot X + b) \exp(-k \cdot b \cdot X) \quad a \sim N(0, 1/3) \quad ; \quad b \sim N(0, 2/3)$$



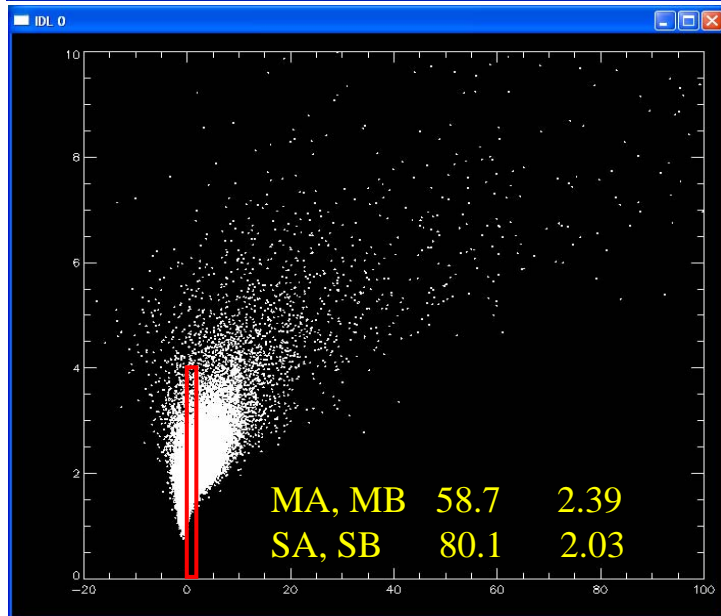
Algorithme C



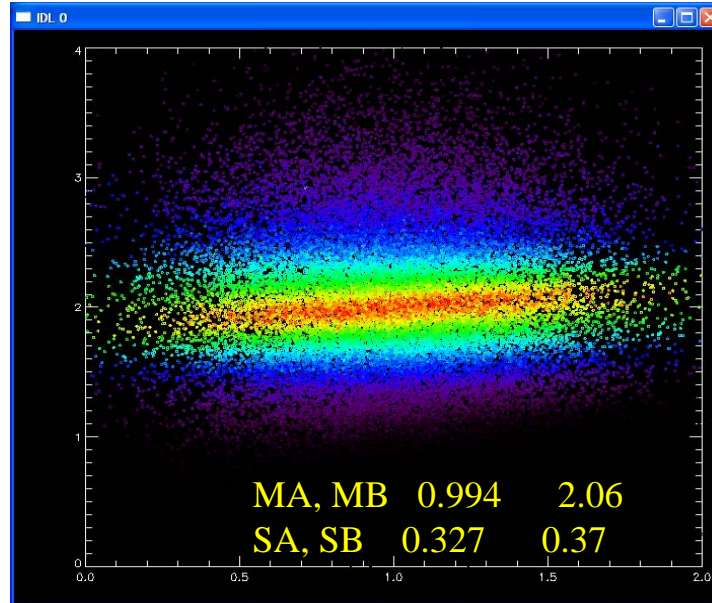
$k = 0.15$



Algorithme B



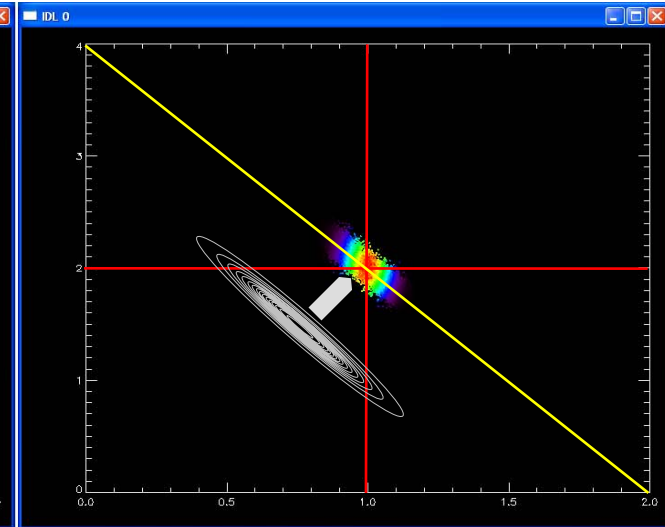
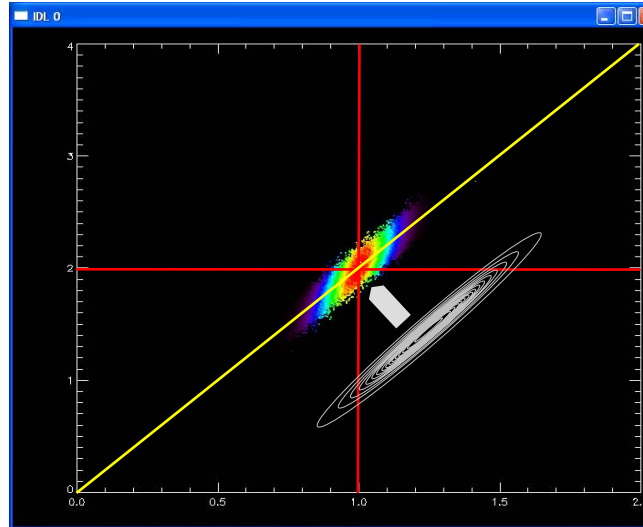
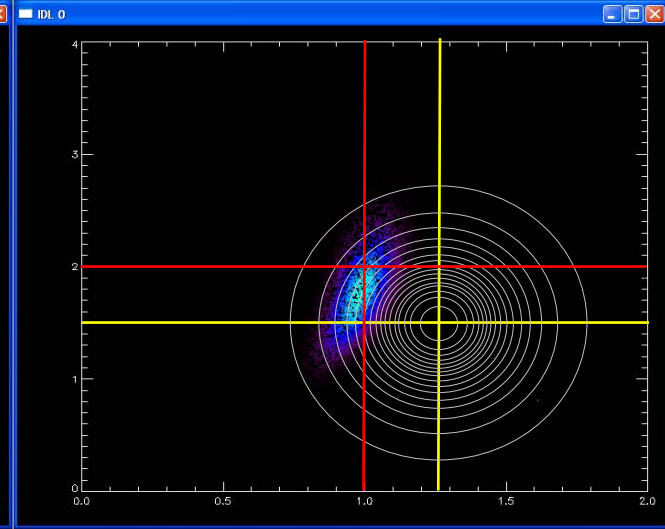
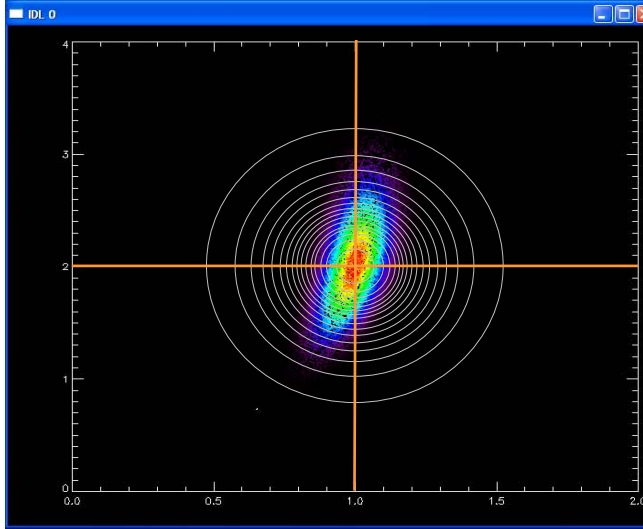
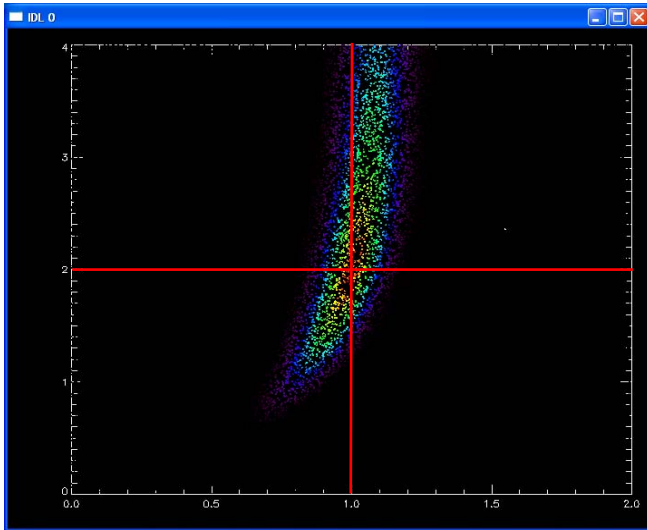
$k = 1$



$$Y = a \frac{1 - \exp(-b \cdot X)}{1 - \exp(-b)}$$



$a \sim U(0, 2)$; $b \sim N(0, 4)$
 $\rho = 0$



Deux remarques concernant l'Algorithme B



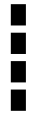
Algorithme complètement parallélisable

Algorithme « dynamique »

Remarque : $\text{prior}(U) \equiv \bar{Q}(\mathcal{H}_m | \emptyset)$

$$\bar{Q}(\mathcal{H}_m | \emptyset), \mathcal{D}^1 \rightarrow \bar{Q}(\mathcal{H}_m | \mathbf{ID}^1)$$

$$\bar{Q}(\mathcal{H}_m | \mathbf{ID}^1), \mathcal{D}^1 \rightarrow \bar{Q}(\mathcal{H}_m | \mathbf{ID}^2)$$



$$\bar{Q}(\mathcal{H}_m | \mathbf{ID}^N), \mathcal{D}^{N+1} \rightarrow \bar{Q}(\mathcal{H}_m | \mathbf{ID}^{N+1})$$

Rappel

$$\mathcal{D}^n = \{ X_{\text{app}}^n, Y_{\text{app}}^n, \tilde{\sigma}_X^n, \tilde{\sigma}_Y^n \}$$

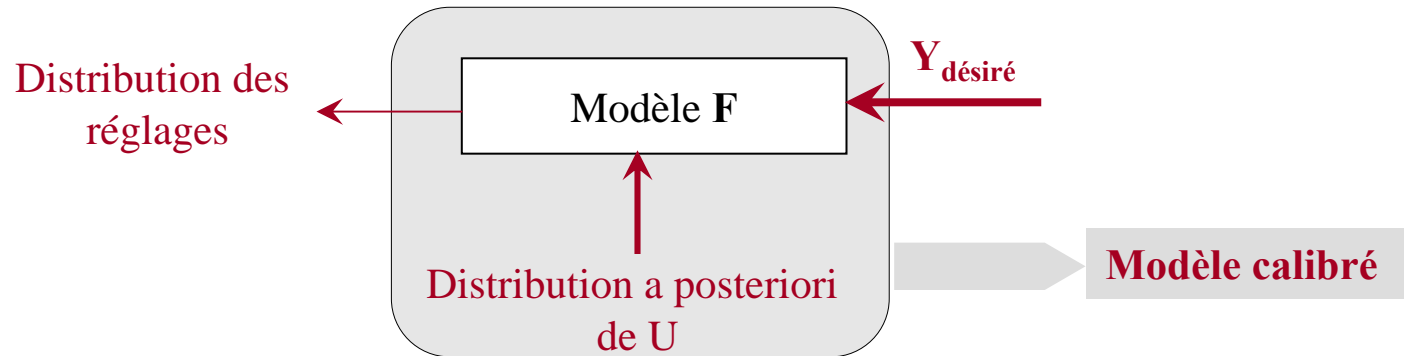
$$\mathbf{ID}^N = \mathcal{D}^1 \dots, \mathcal{D}^N$$

⇒ rapidité potentielle ... mieux « tirage selon un a posteriori » ???

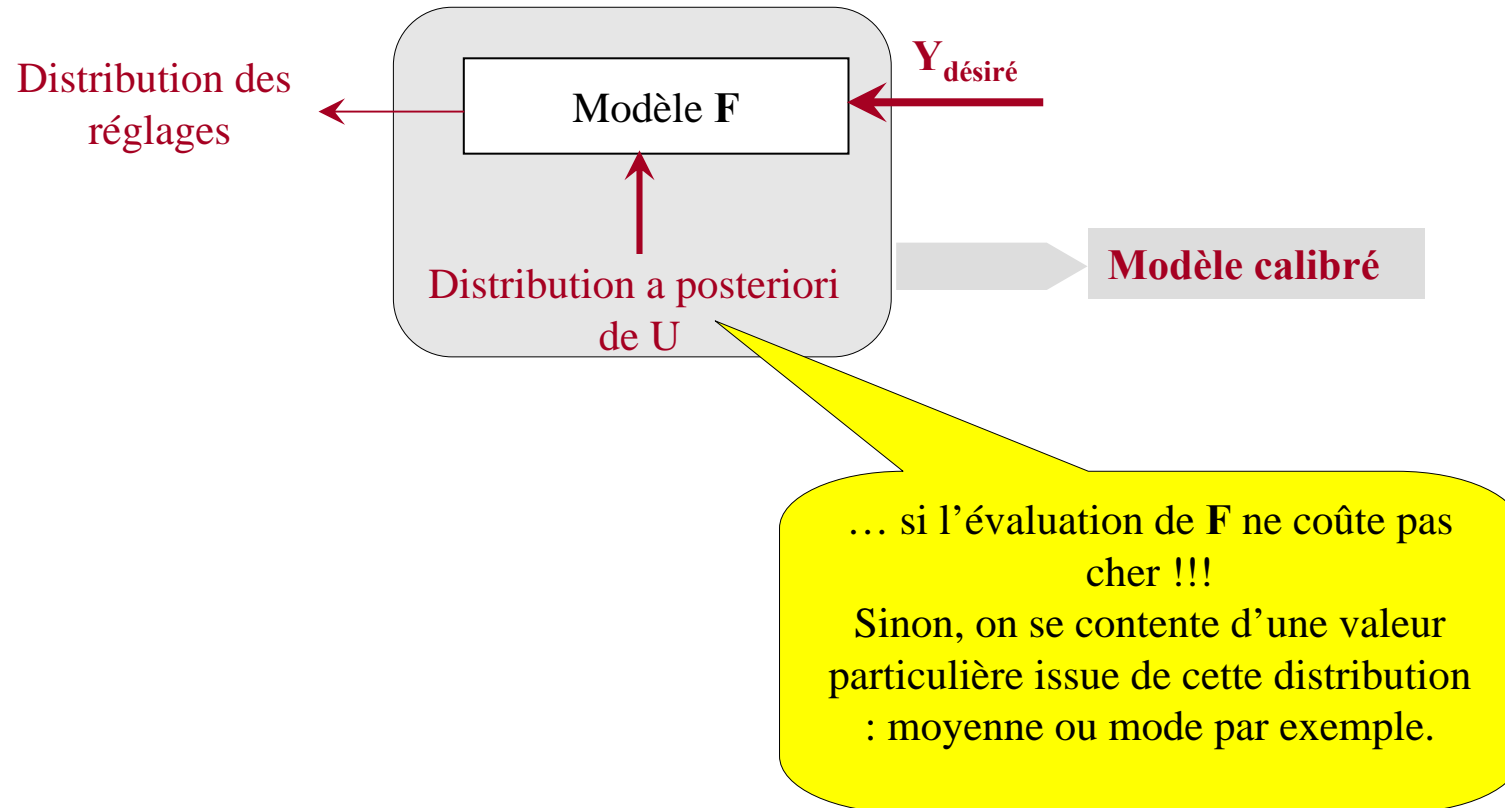
⇒ robustesse

⇒ performances

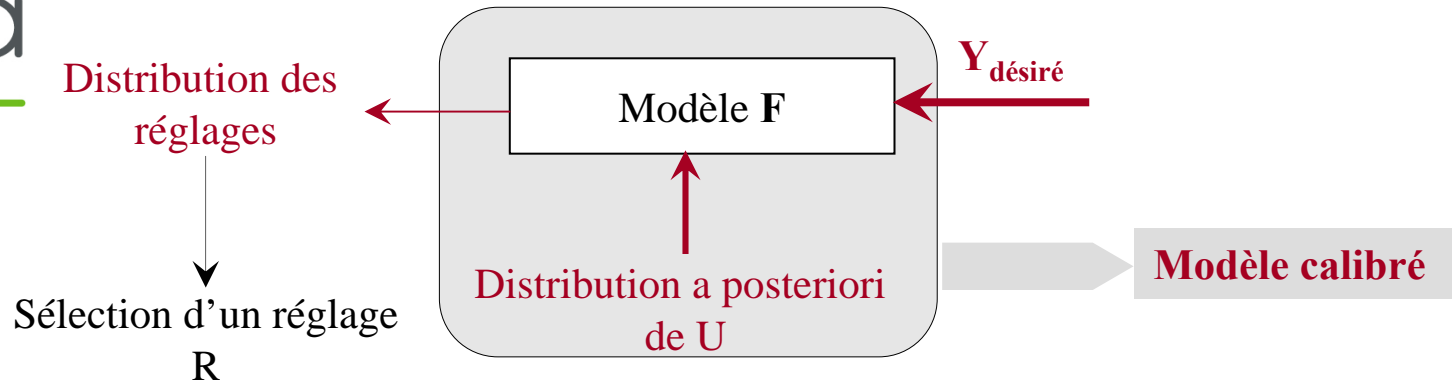
La prédiction des réglages



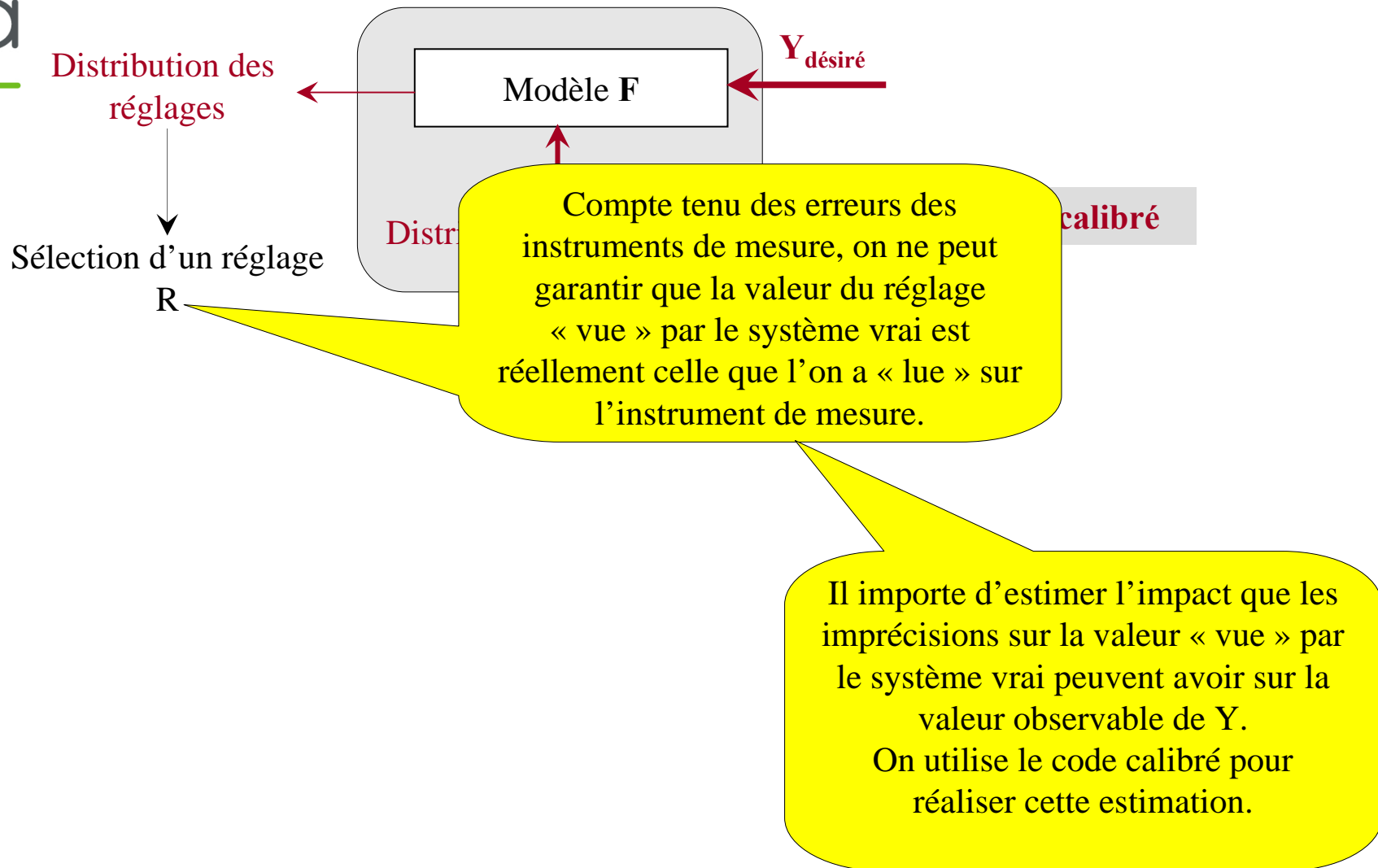
La prédiction des réglages



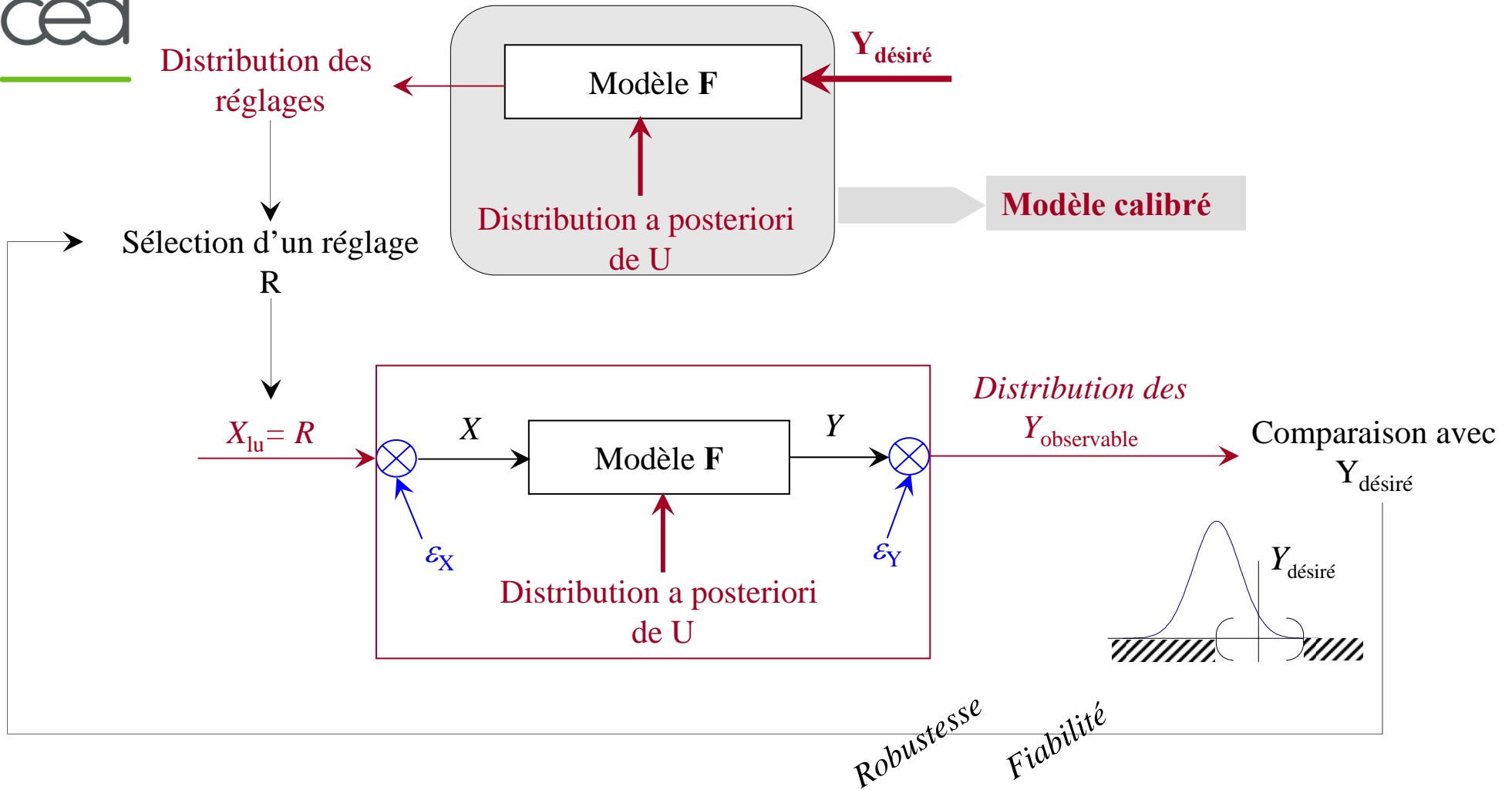
La prédiction des réglages



La prédiction des réglages



La prédiction des réglages



① Calibration : problème **inverse**



$$\bar{Q} \left(\mathcal{H}_m \mid \mathbf{ID}^N \right), \mathcal{D}^{N+1} \xrightarrow{\text{Algorithme B}} \underbrace{\bar{Q} \left(\mathcal{H}_m \mid \mathbf{ID}^{N+1} \right)}_{\text{Distribution a posteriori de } U}$$

② Prédiction des réglages : problème **inverse**

$$\bar{Q} \left(\text{réglages} \mid \emptyset \right), \bar{Q} \left(\mathcal{H}_m \mid \mathbf{ID}^{N+1} \right), Y_{\text{désiré}} \xrightarrow{\text{Algorithme B}} \underbrace{\bar{Q} \left(\text{réglages} \mid \mathbf{ID}^{N+1}, Y_{\text{désiré}} \right)}_{\text{Distribution a posteriori de } X}$$

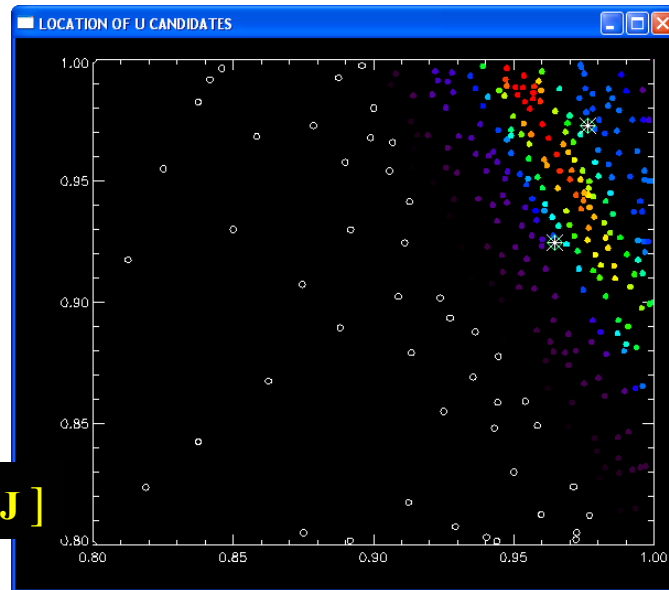
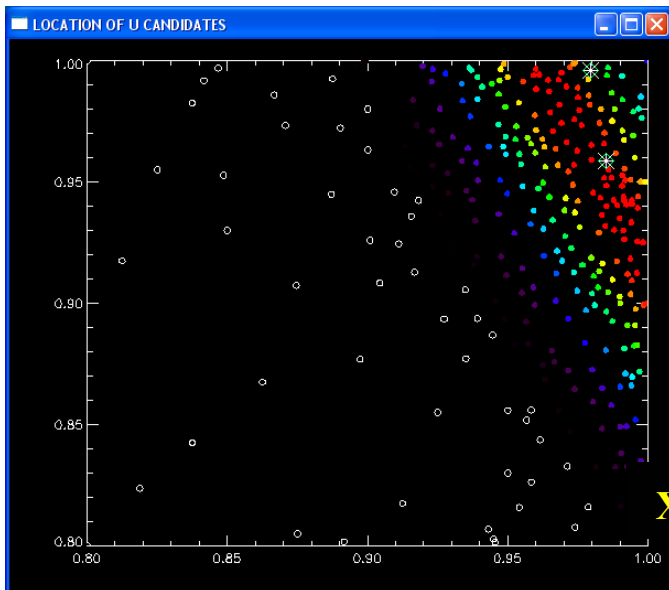
③ Qualification du réglage χ : problème **direct** de propagation d'incertitude

$$\chi \sim \bar{Q} \left(\text{réglages} \mid \mathbf{ID}^{N+1}, Y_{\text{désiré}} \right)$$
$$\bar{Q} \left(\mathcal{H}_m \mid \mathbf{ID}^{N+1} \right), Y_{\text{désiré}}, \chi \rightarrow \text{Distribution des } Y \text{ observables}$$

④ Choix d'un réglage « optimal » : en présence d'incertitudes c'est un problème **inverse** complexe

Calibration sur deux paramètres

(réflexions de miroirs)

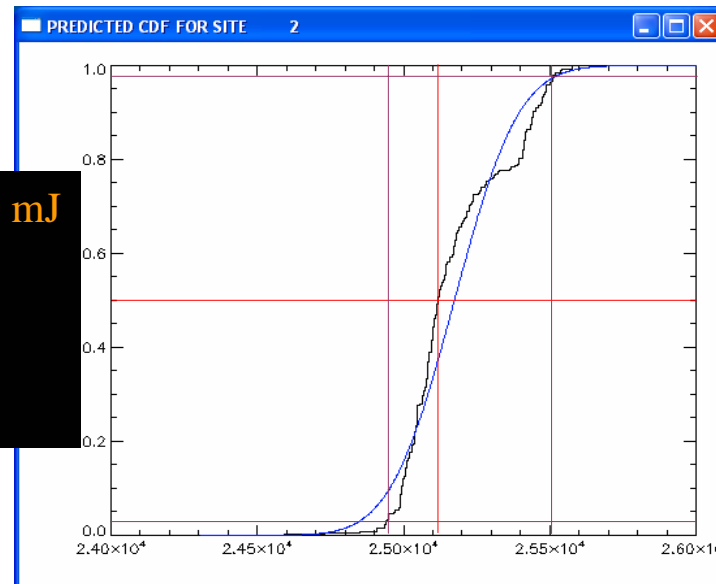
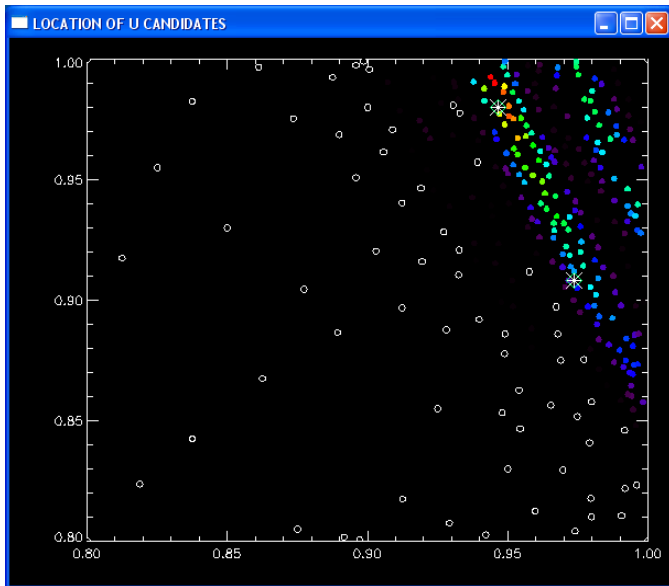


$X_{app} \in [5 \text{ mJ}, 90 \text{ mJ}]$

N=2

N=4

N=12



Prédiction à $x=200 \text{ mJ}$

Moyenne 25123

Médiane 25107

Mode (+) 25088

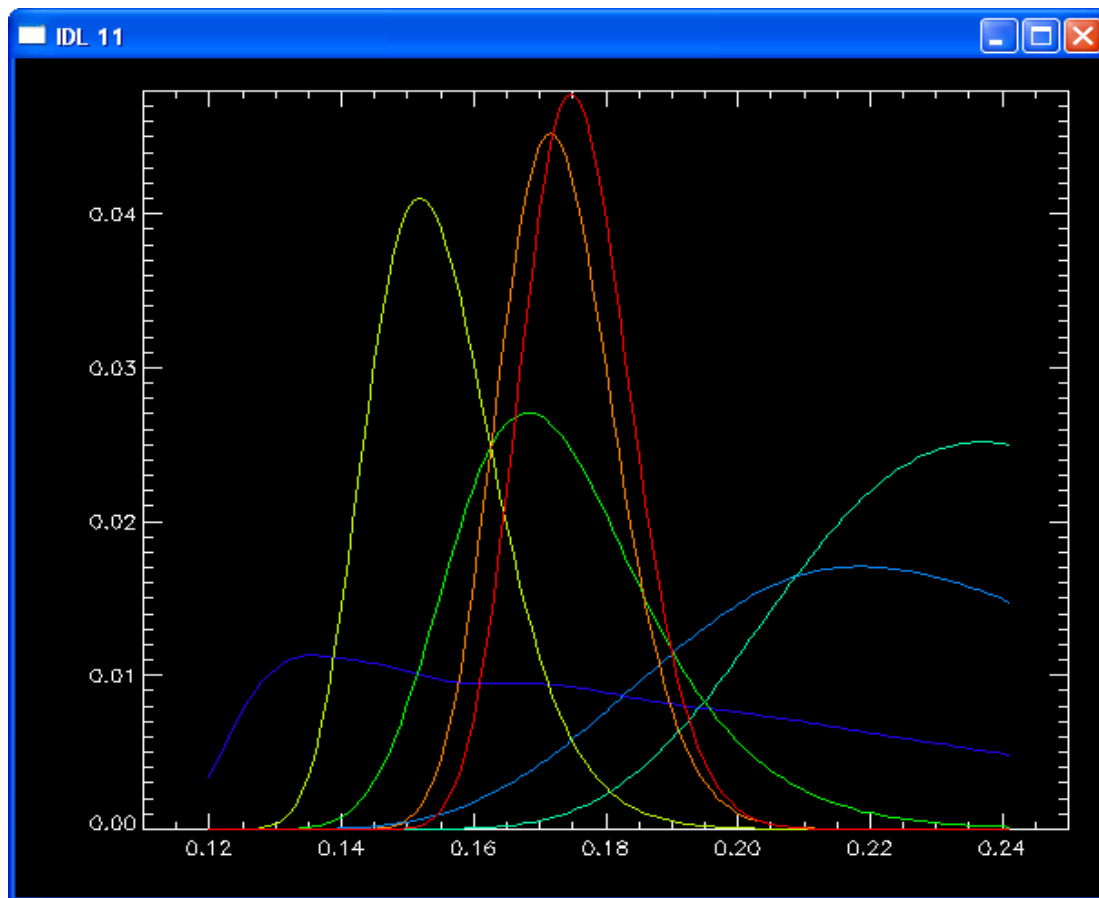
Théorique 25087

Prédiction de réglage



Le tir 8 a donné **13685 J** pour **172.1 mJ** injectées : partant de $Y_{\text{désirée}}=13685 \text{ J}$, retrouver le réglage à appliquer

Calibration bayésienne « dynamique » (sur 1 tir, puis 2, puis 3, ...)
Prédiction de la distribution des réglages après chaque calibration



Nb tirs	Moyenne	Ecart type
1	174.51	33.05
2	208.91	20.57
3	218.61	15.75
4	173.85	15.90
5	154.84	10.15
6	173.16	8.96
7	175.95	8.41

Conclusion



CALIBRATION :

- **L'approche « par minimisation » doit être maniée avec beaucoup de précautions**
- L'approche bayésienne est
 - Souple, robuste, « universelle », « maîtrisable » et facilement parallélisable
 - Mais attention aux temps de calculs ...
... l'algorithme B est inefficace : c'est une illustration et non l'algorithme réel utilisé.

PRÉDICTION des RÉGLAGES

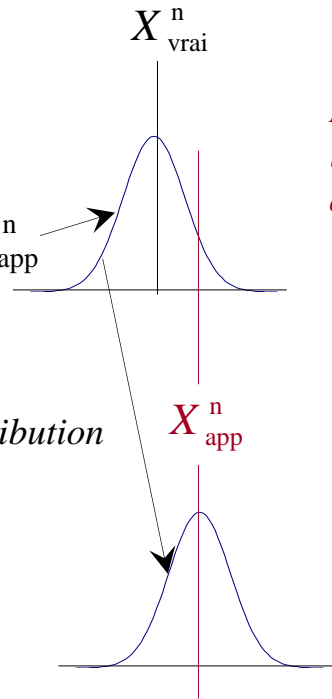
- **Mêmes remarques que précédemment**
- **Le choix du meilleur réglage est un problème ouvert**

POINTS DURS

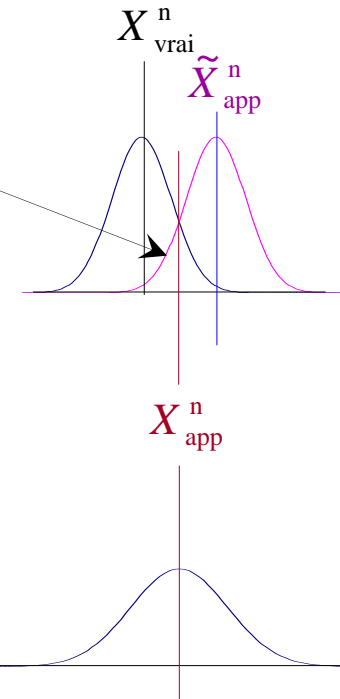
- Erreurs sur « les entrées » X :
Très rarement traitées (correctement)
- Données X et Y fonctionnelles :
Actuellement un vaste sujet de recherches
- L'invariance de U_{vrai} est fautive (vieillesse, dégradations) :
Idée : calibration « par lots » → classification non supervisée
- Enfin, on peut avoir plusieurs faisceaux partageant des éléments communs $30 \times 2 \times 4 \times 1$
??? Où va-t-on ???

Distribution a priori de X_{app}^n

Même distribution



Mais X_{app}^n pourrait être issu de cette distribution avec la même vraisemblance.



Distribution à posteriori de X_{vrai}^n sachant X_{app}^n

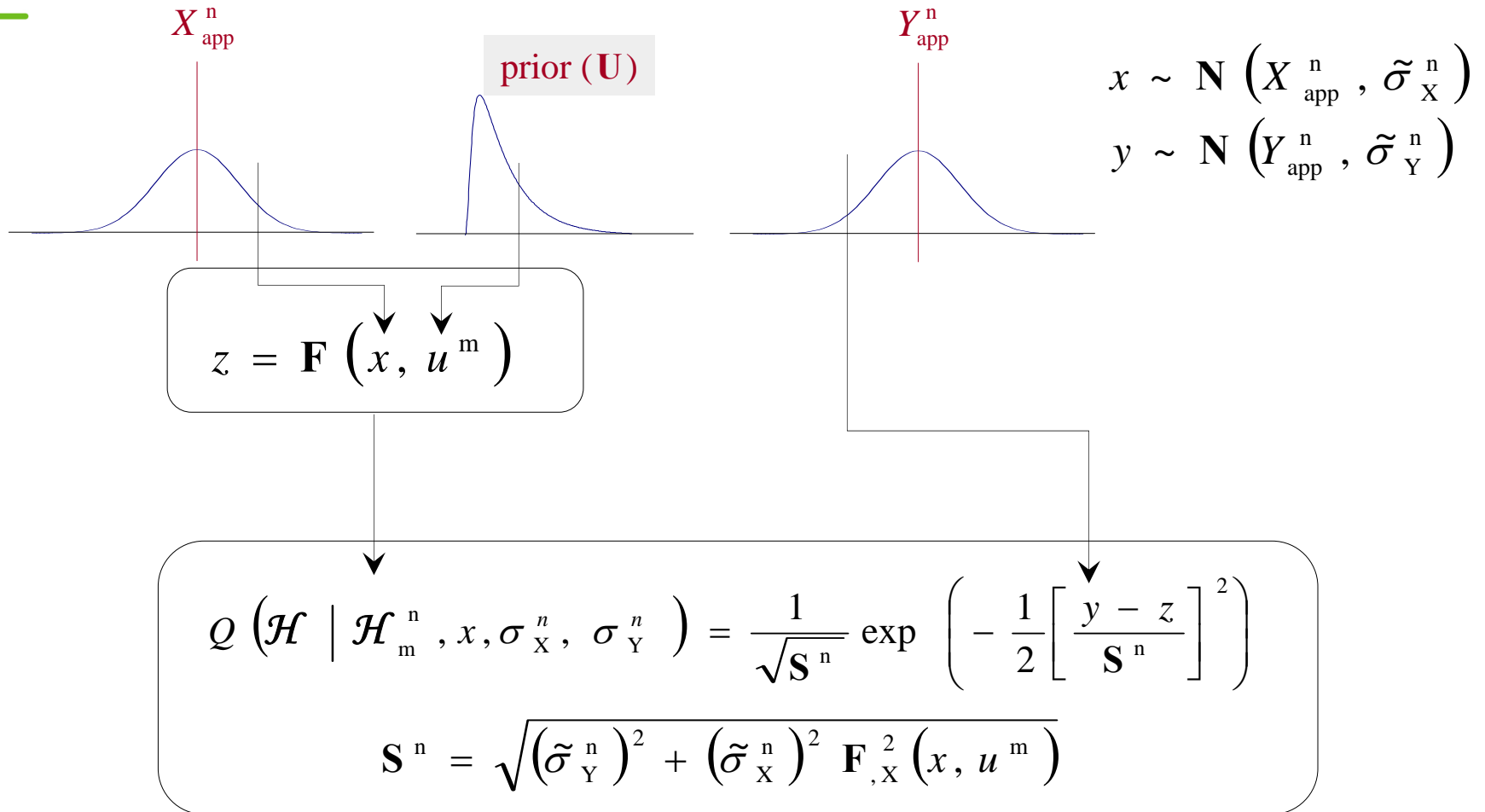
$$\underbrace{\text{postpdf}}_{\Phi_X^n} \left(X_{vrai}^n \mid X_{app}^n \right) = \text{pdf} \left(X_{app}^n \right) * \text{pdf} \left(X_{app}^n \right)$$

Exemple

$$\text{pdf} \left(X_{app}^n \right) \sim N \left(X_{vrai}^n, \sigma_X \left(X_{vrai}^n \right) \right) \rightarrow \Phi_X^n \left(X_{vrai}^n \mid X_{app}^n \right) \sim N \left(X_{app}^n, \sigma_X \left(X_{app}^n \right) \sqrt{2} \right)$$

Hypothèse \mathcal{H}_m^n : $u^m = U_{\text{vrai}}$ compte tenu des informations au site n

Hypothèse \mathcal{H} : $z = y$



$$\mathcal{D}^n = \left\{ X_{\text{app}}^n, Y_{\text{app}}^n, \tilde{\sigma}_X^n, \tilde{\sigma}_Y^n \right\}$$

Probabilités conditionnelles (formules de Bayes)

$$\bar{Q}(\mathcal{H}_m^n | \mathcal{D}^n) = \int Q(\mathcal{H}_m^n | \mathcal{H}, x, \sigma_X^n, \sigma_Y^n) \Phi_X^n(x) \Phi_Y^n(y) dx dy$$

$\bar{Q}(\mathcal{H}_m^n | \mathcal{D}^n)$ est la vraisemblance de l'hypothèse \mathcal{H}_m^n

Hypothèse \mathcal{H}_m : $u^m = U_{\text{vrai}}$ compte tenu des informations aux sites 1, ..., N

$$\mathbf{ID}^N = \mathcal{D}^1 \dots, \mathcal{D}^N$$

$$\bar{Q}(\mathcal{H}_m | \mathbf{ID}^N) = \prod_{n=1}^N \bar{Q}(\mathcal{H}_m^n | \mathcal{D}^n)$$

$\bar{Q}(\mathcal{H}_m | \mathbf{ID}^N)$ est la vraisemblance de l'hypothèse \mathcal{H}_m