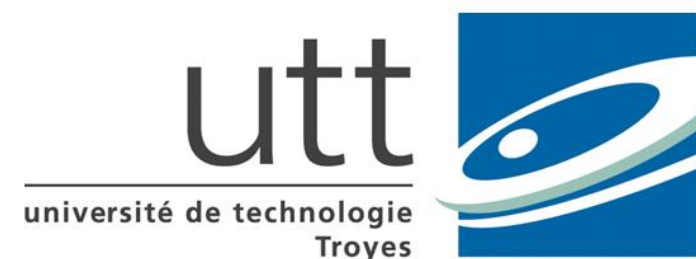




Modélisation stochastique de dégradation : estimation robuste par bootstrap

Quentin Chatenet – Big Ideas for Small Data

Mai 2019



Sommaire

- | | | | |
|---|-------------------------------------|---|-----------------------------|
| 1 | Problématique industrielle | 4 | Principe du bootstrap |
| 2 | Essais de cavitation en laboratoire | 5 | Mise en oeuvre |
| 3 | Modélisation de la dégradation | 6 | Conclusions et perspectives |

PROBLÉMATIQUE INDUSTRIELLE

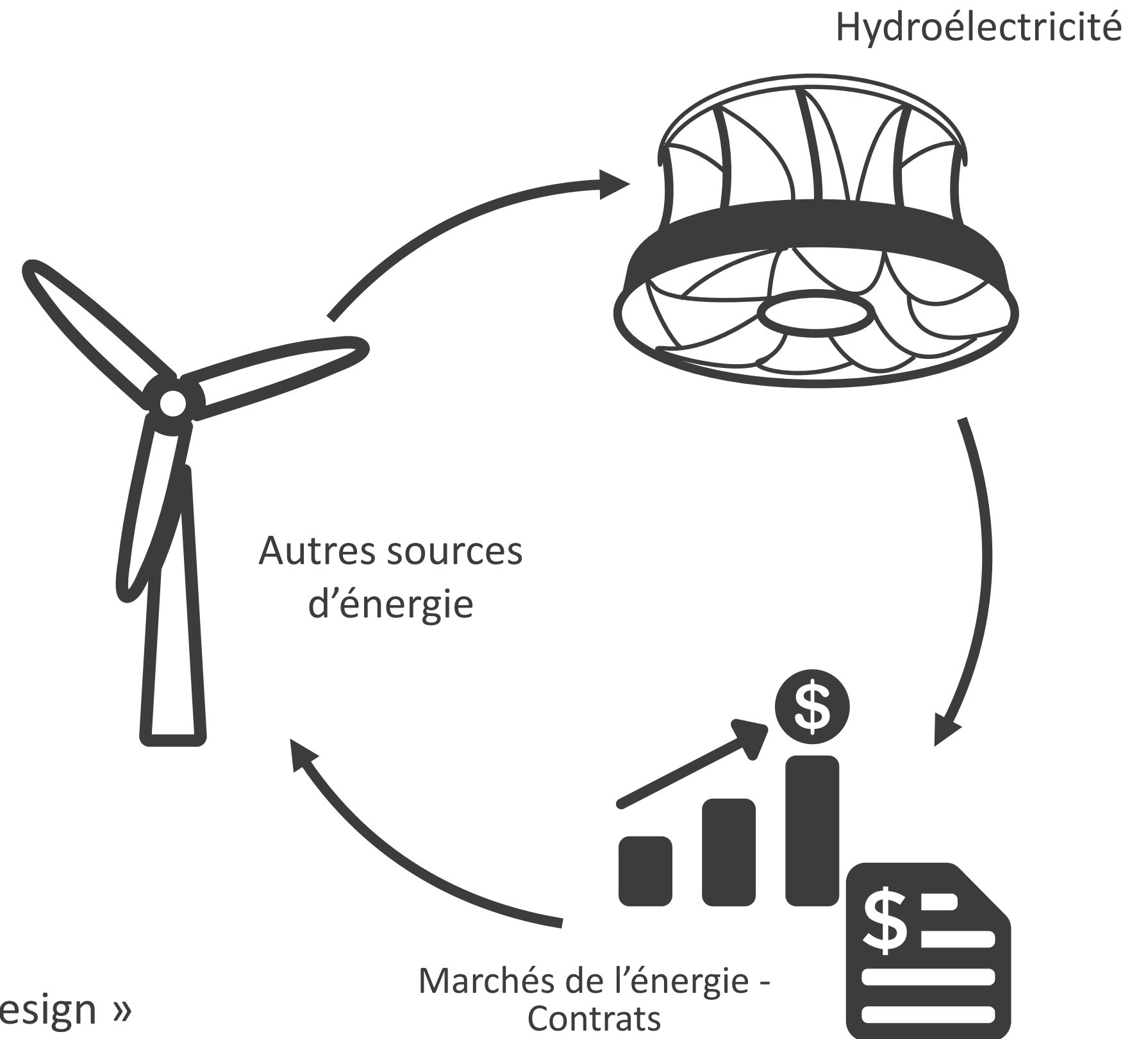
Problématique industrielle

Nouvelles contraintes de régulation du réseau avec les EnR



Marché de l'énergie de plus en plus concurrentiel

➔ Recours aux groupes Turbine/Alternateur pour lisser la production
Utilisation des groupes T/A dans des zones « hors design »



Problématique industrielle



Problématique industrielle

PERFORMANCES



RENTABILITÉ



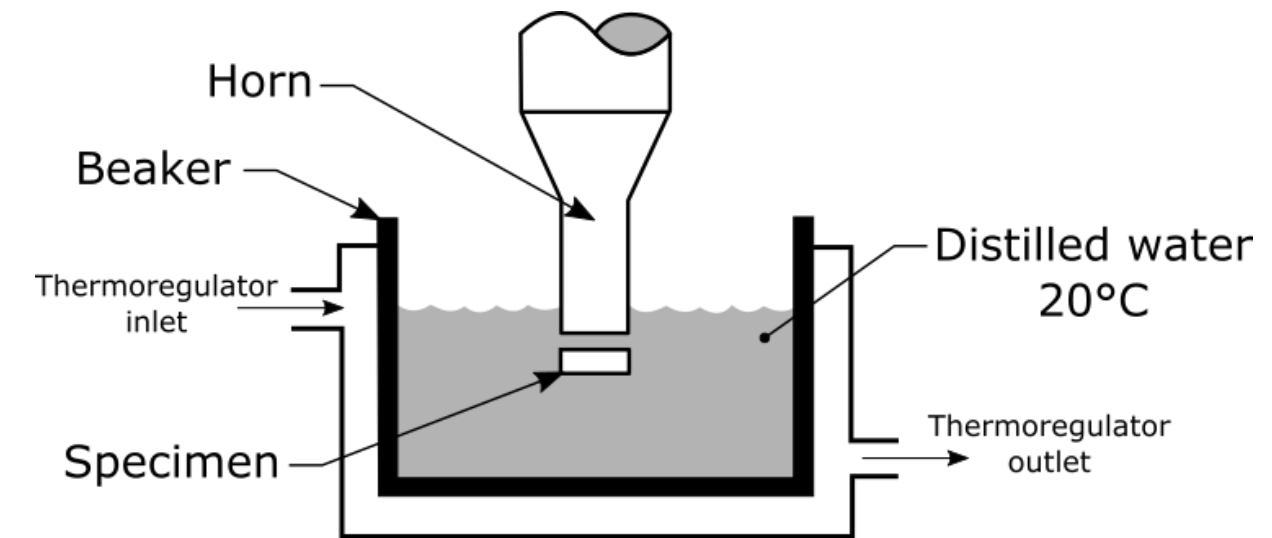
VIBRATIONS



ESSAIS DE CAVITATION EN LABORATOIRE

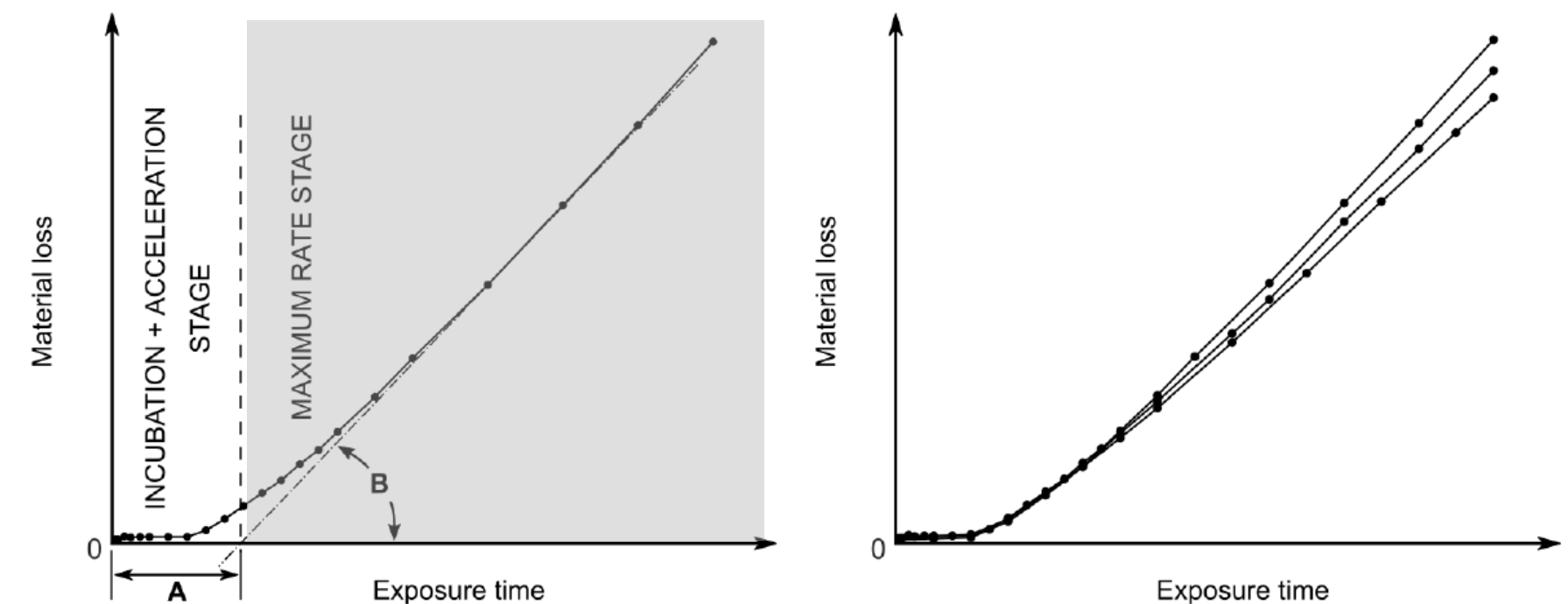
Essais en laboratoire

Création de cavitation à l'aide d'un cône vibrant



La masse de l'échantillon est régulièrement mesurée

D_{obs} = 30 observations disponibles dans la zone d'intérêt réparties sur 3 trajectoires de 10 incréments



Formalisation et notations

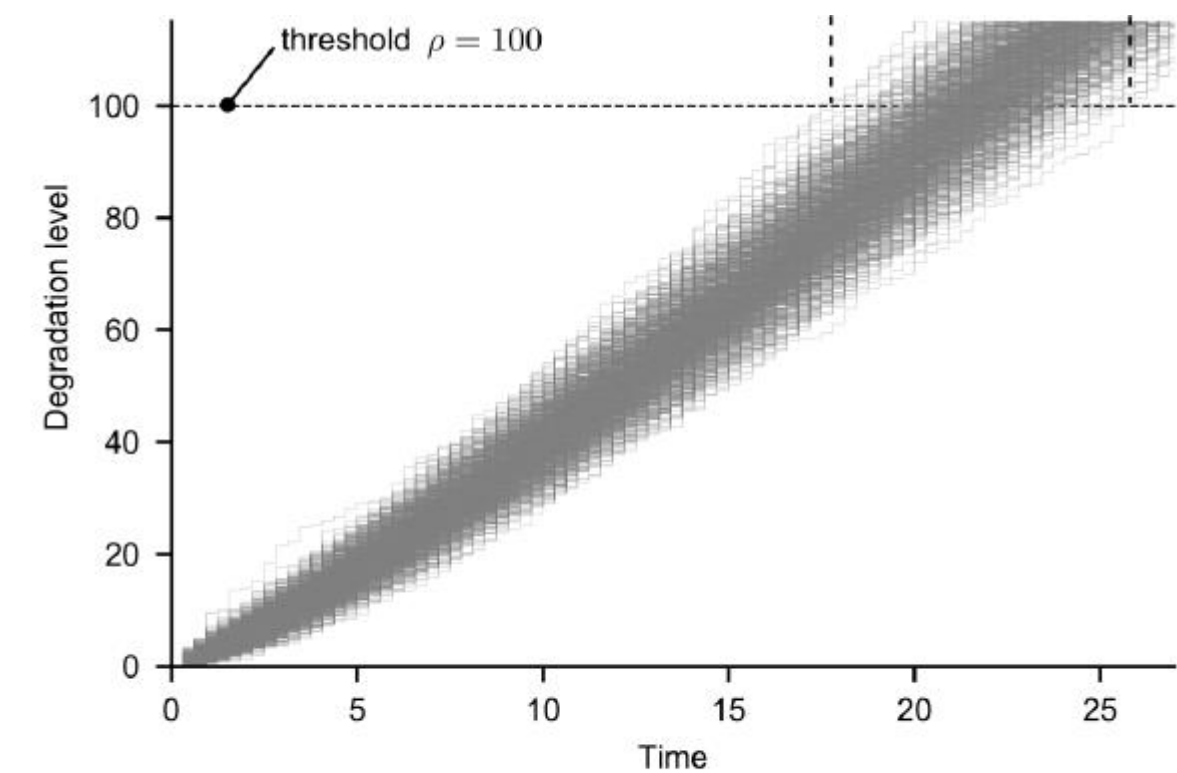
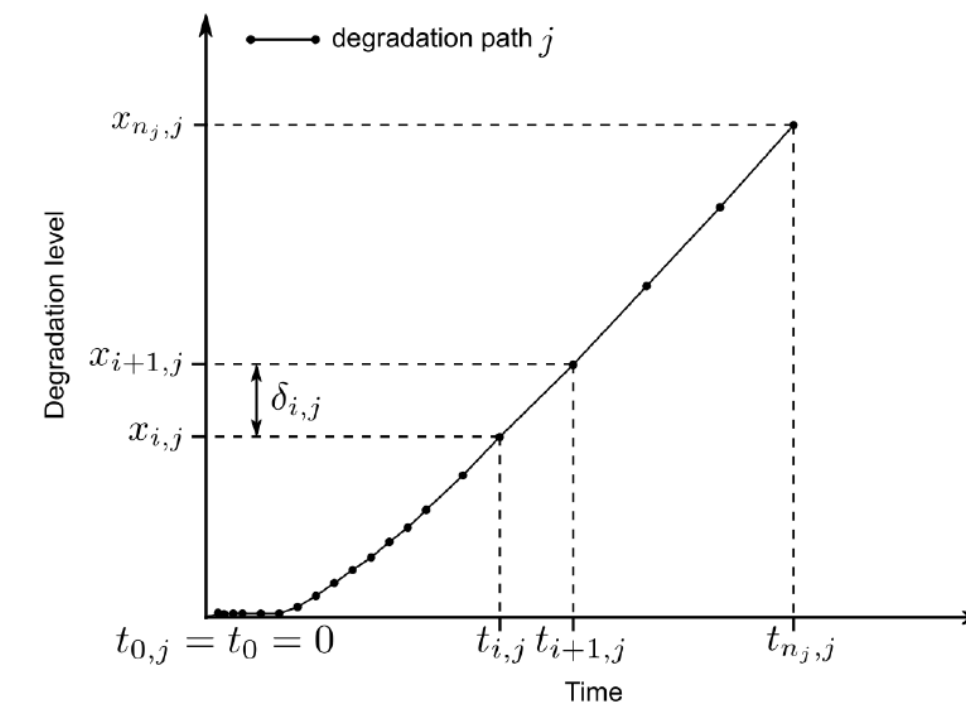
$$\mathbf{D}_{\text{obs}} = \bigcup_{j=1}^m \bigcup_{i=0}^{n_j-1} \{(t_{i,j}, t_{i+1,j}, \delta_{i,j})\}$$

m : nombre de trajectoires

n_j : nombre de points par trajectoire

Temps d'atteinte $\tau_\rho = \min\{t_{i,j} > 0 \mid x_{i,j} \geq \rho\}$

ρ : seuil de dégradation limite



Hitting time distribution for 10^3 simulated degradation paths.

MODÉLISATION DE LA DÉGRADATION

Modélisation de la dégradation

Processus stochastique utilisé : gamma non-homogène

- $X(0) = 0$,
- $X(\tau) - X(t) \sim Ga(u, v(\tau) - v(t)) \quad \forall \tau > t > 0$,
- $X(t)$ a des incréments indépendants et disjoints. La dégradation future est indépendante du passé.

Fonction de densité de probabilité de la loi gamma :

$$f_X(x|u, v(t)) = Ga(x|u, v(t)) = \frac{u^{v(t)}}{\Gamma(v(t))} x^{v(t)-1} \exp(-ux) I_{\mathbb{R}_+^*}(x)$$

$$\text{avec } v(t) = ct^b$$

Estimation des paramètres

Vraisemblance du processus gamma :

$$\mathcal{L}(\boldsymbol{\theta} | \mathbf{D}_{obs}) = \mathcal{L}(u, c, b | \mathbf{D}_{obs}) = \prod_{j=1}^m \prod_{i=1}^{n_j} Ga(\delta_{i,j} | u, v(t_{i,j}) - v(t_{i-1,j})) \quad \text{avec } v(t) = ct^b$$

Estimation du paramètre $\boldsymbol{\theta} = (u, c, b)$ par méthodes numériques, noté $\hat{\boldsymbol{\theta}} = (\hat{u}, \hat{c}, \hat{b})$

Peu de données disponibles



Nécessité d'évaluer les incertitudes d'estimation des paramètres + temps d'atteinte τ_ρ

Evaluation des incertitudes d'estimation

Intervalle de confiance bilatéral au niveau $1 - \alpha$ de θ_i construit à l'aide des propriétés asymptotiques du MLE :

$$\left[\hat{\theta}_i - s \sqrt{\left(J(\hat{\boldsymbol{\theta}})^{-1} \right)_{ii}} ; \hat{\theta}_i + s \sqrt{\left(J(\hat{\boldsymbol{\theta}})^{-1} \right)_{ii}} \right], \quad i = 1 \dots 3$$

avec $J(\hat{\boldsymbol{\theta}})$ l'information de Fisher observée et $s = \Phi \left(1 - \frac{\alpha}{2} \right)$

Incertitude sur la distribution de τ_ρ évaluée avec la Delta méthode

Peu de données disponibles



Conditions asymptotiques non respectées. L'approche par *bootstrap* semble plus pertinente dans ce cas.

PRINCIPE DU BOOTSTRAP

Principe du *bootstrap* (rééchantillonnage)

Objectif :

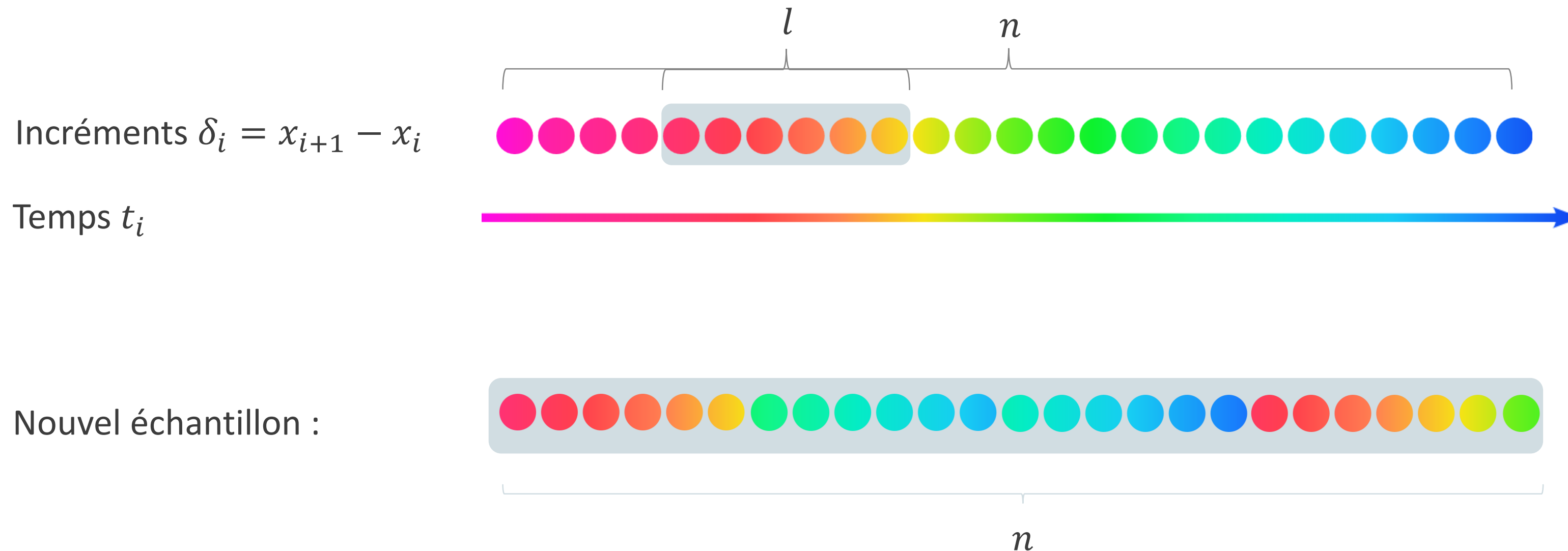
- évaluer par rééchantillonnage des données disponibles les incertitudes d'estimation des paramètres d'un modèle (ici : processus stochastique gamma) ou de tout autre indicateur d'intérêt (ici : caractéristiques du temps d'atteinte τ_ρ)

Concept général :

- B nouveaux échantillons sont construits à partir de l'échantillon de référence $\mathbf{D}_{obs}$
- les paramètres du processus et les indicateurs d'intérêt sont estimés/calculés pour chaque nouvel échantillon

Principe du *bootstrap* (rééchantillonnage)

Moving Block Bootstrap (MBB) – non paramétrique



Si on prend $l = 1 \rightarrow$ Efron bootstrap (EB)

Principe du *bootstrap* (rééchantillonnage)

Moving Block Bootstrap (MBB) – non paramétrique

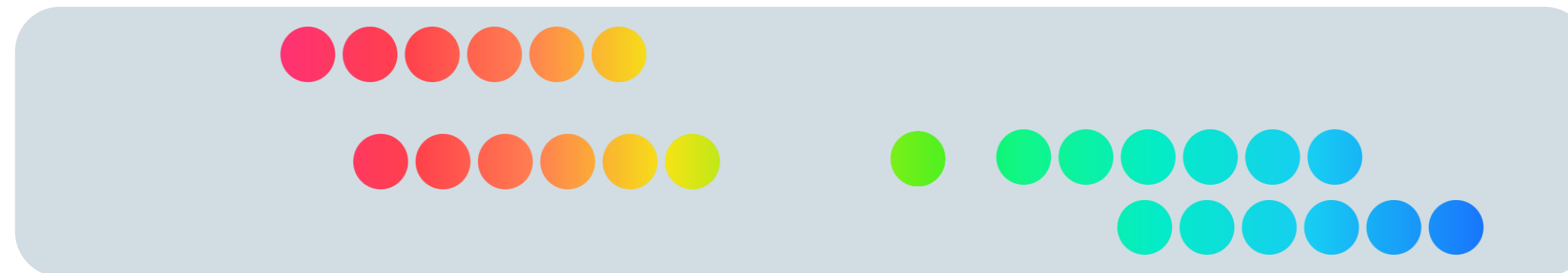
Incréments $\delta_i = x_{i+1} - x_i$



Temps t_i



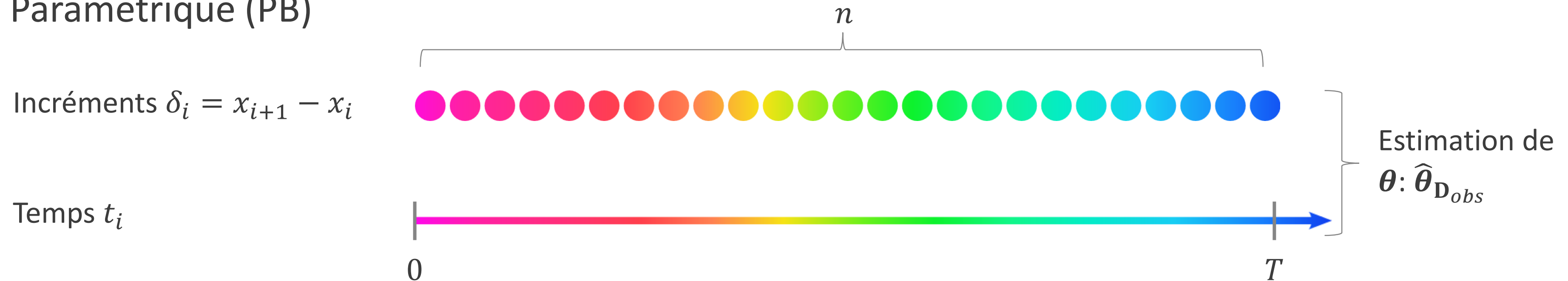
Nouvel échantillon :



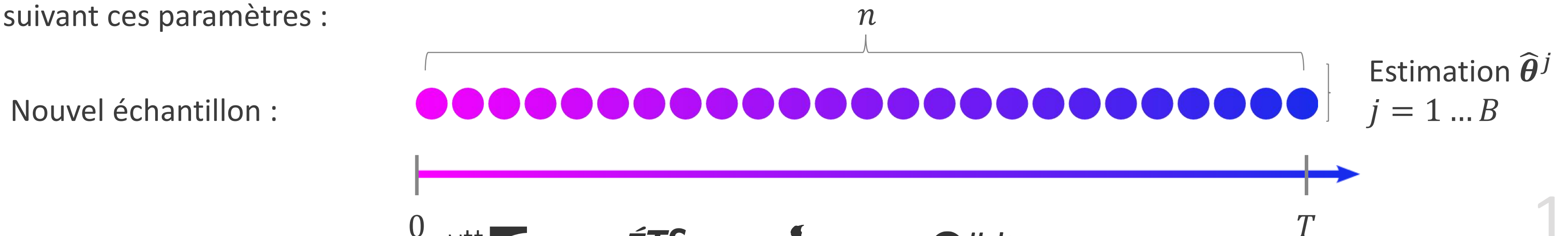
Estimation $\hat{\theta}^j$
 $j = 1 \dots B$

Principe du *bootstrap* (rééchantillonnage)

Paramétrique (PB)

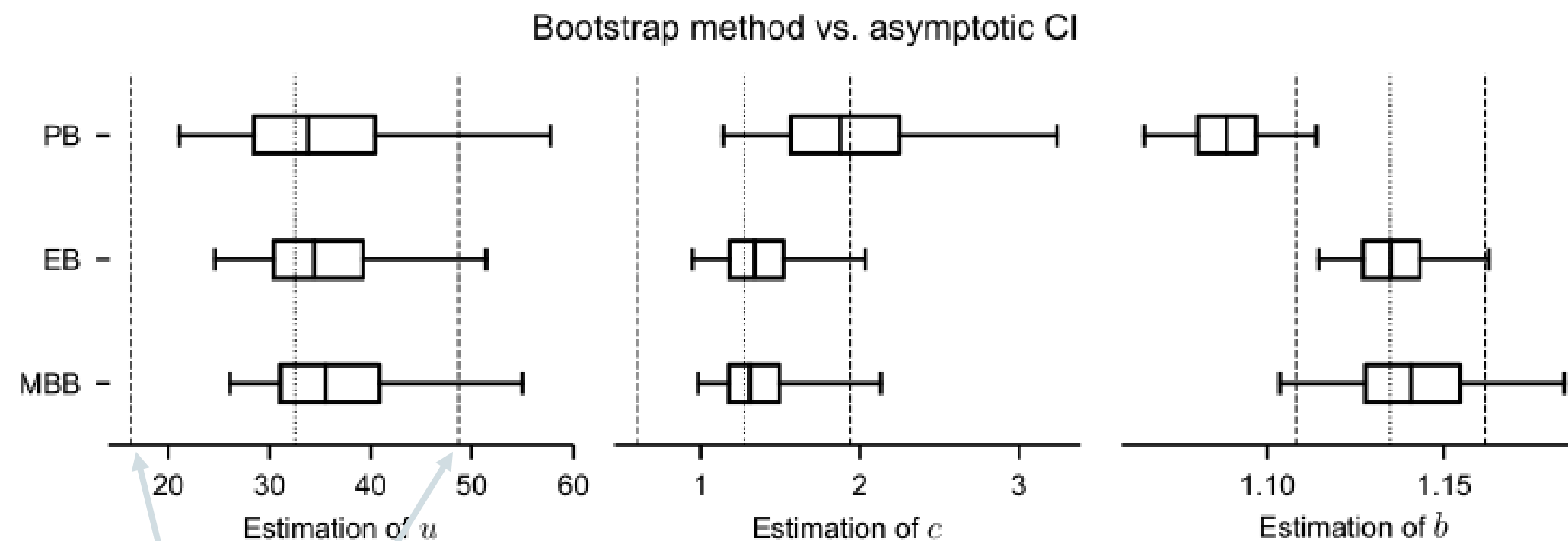


On suppose les paramètres du modèle connus $\hat{\theta}_{D_{obs}}$ et on simule une trajectoire de dégradation suivant ces paramètres :



MISE EN ŒUVRE

Applications aux données de dégradation



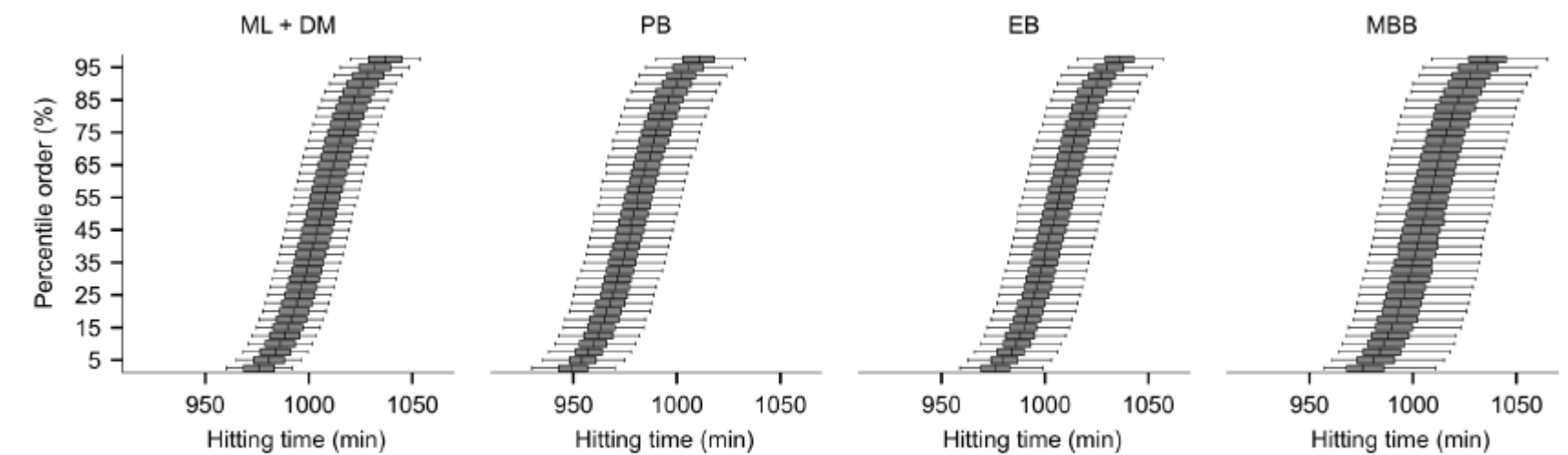
Comparison of bootstrap and asymptotic CI.

IC asymptotiques

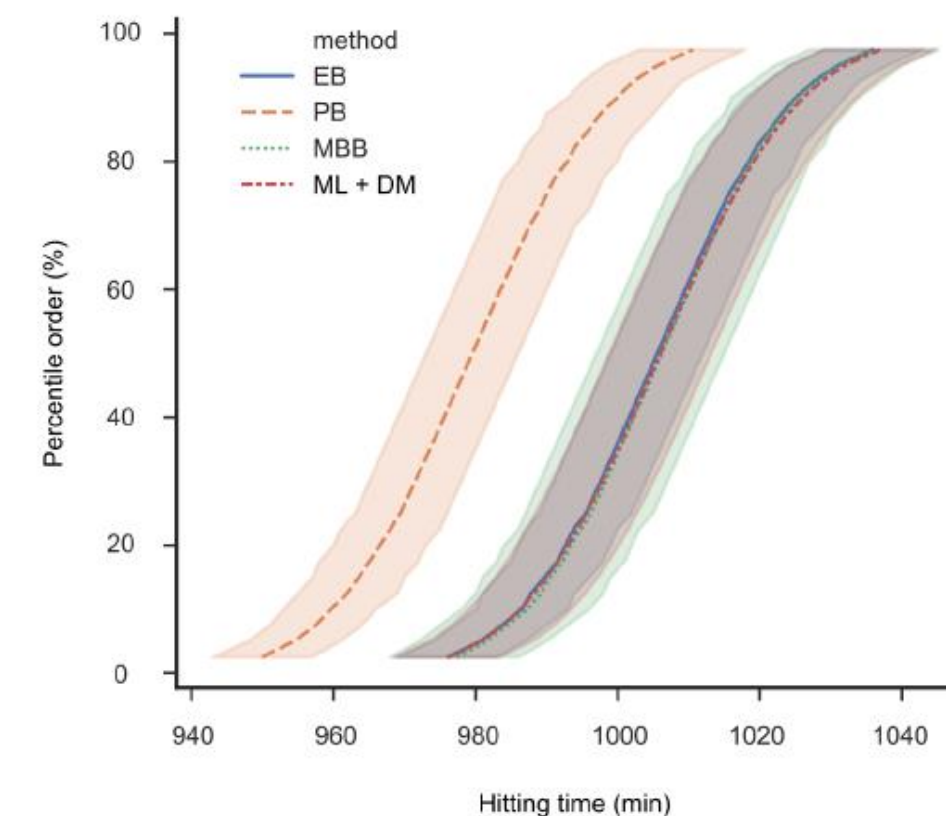
- Choix de gamma NH justifié
- **PB** : grands biais et dispersion → difficultés pour les « petites » trajectoires
- **EB** et **MBB** : bonne estimation du paramètre b (biais et dispersion relatifs < 1%) ; estimation correcte pour u et c (biais et dispersion relatifs $\approx 10\%$)
- Boîtes à moustaches non-symétriques

Applications aux données de dégradation

- **PB** donne des temps d'atteinte éloignés des autres méthodes
- Même dispersion quel que que soit l'ordre de percentile de τ_ρ
- **MBB** présente la plus forte dispersion sur les temps d'atteinte
- **ML + DM** donne des résultats proches de **EB**



Uncertainties propagation for hitting time ($\rho = 100 \text{ mg}$) distribution.



Hitting time distributions of T_ρ (for $\rho = 100 \text{ mg}$) for 3 bootstrap methods and ML + DM.

CONCLUSIONS ET PERSPECTIVES

Conclusions

- Mise en œuvre avec succès de méthodes *bootstrap* sur un processus de dégradation.
- En présence de peu de données, on préfère utiliser des méthodes *bootstrap* à ML + DM (cette dernière reposant sur l'hypothèse de normalité asymptotique).

Perspectives

- La liste des méthodes *bootstrap* présentée ici n'est pas exhaustive : d'autres méthodes pourraient être explorées.
- Évaluation par simulation intensive de la performance des techniques de *bootstrap* en fonction des valeurs de (θ, m, n_j) .
- En présence de connaissance d'experts sur le phénomène de dégradation, une approche bayésienne pourrait être appropriée.
- Exploitation des données de monitoring - beaucoup plus compliquées : bruit, fonction de transfert liée à la méthode d'acquisition, intensité de cavitation variable (environnement non contrôlé).

MERCI ! QUESTIONS ?