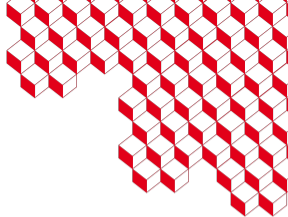




isas



Calage bayésien pour la quantification de l'incertitude conditionnelle des paramètres d'entrée dans les modèles numériques chaînés

VVF Lège Cap-Ferret, 12 Octobre 2023

O. BALDÉ (oumar.balde@cea.fr), G. DAMBLIN, A. MARREL, A. BOULORÉ and L. GIRALDI
CEA CENTRE SACLAY

Plan de la présentation

Calage d'un modèle numérique

- Calage de modèle

- Estimation bayésienne

Calage des modèles numériques interconnectés

- Calage de modèles numériques chaînés

- Cas applicatif : application ALCYONE

- Calage bayésien conditionnel

Nouvelle méthode GP-LinCC

- Principe

- En pratique

Conclusion et perspectives



Plan de la présentation

Calage d'un modèle numérique

Calage des modèles numériques interconnectés

Nouvelle méthode GP-LinCC

Conclusion et perspectives





Calage de modèle

- **Modèles numériques** : outils essentiels pour mimer des systèmes physiques (ex : les centrales nucléaires).
- **Nombre important de paramètres d'entrée**, caractérisant les systèmes physiques étudiés.
- Paramètres d'entrée le plus souvent **incertains** ainsi que les **sorties** \implies besoin de la **quantification d'incertitude** [de Rocquigny et al., 2008].
- **Calage de modèle** :
 - **But** : quantifier l'incertitude des paramètres en fonction des données expérimentales disponibles du système physique étudié.
 - Deux types de calage : calage déterministe et **calage bayésien**.

Calage bayésien d'un modèle numérique

- **But** : quantifier l'incertitude paramétrique par des distributions de probabilité [Kennedy and O'Hagan, 2001] :
 - La distribution *a priori* des paramètres incertains $\theta \in \mathcal{D}_\theta \subset \mathbb{R}^p$: $\pi(\theta)$.
 - La vraisemblance des données observées z du système physique : $\mathcal{L}(z | \theta)$
 - **La distribution *a posteriori*** $\pi(\theta | z)$ est obtenue à partir de la formule de Bayes :

$$\pi(\theta | z) \propto \mathcal{L}(z | \theta) \pi(\theta).$$

- $\pi(\theta | z)$: mise à jour de $\pi(\theta)$ en utilisant l'information fournie par z .
- Un estimateur de paramètres largement utilisé est **maximum a posteriori (MAP)** : $\hat{\theta} := \operatorname{argmax}_{\theta \in \mathcal{D}_\theta} \pi(\theta | z)$.

Plan de la présentation

Calage d'un modèle numérique

Calage des modèles numériques interconnectés

Nouvelle méthode GP-LinCC

Conclusion et perspectives



Calage de modèles chaînés

- Focus sur les deux modèles chaînés M1 et M2 de la figure 1 :
 - $x \in \mathbb{R}^d$ ($d \geq 1$) : variables de contrôle,
 - $\theta \in \mathcal{D}_\theta \subset \mathbb{R}^p$ ($p \geq 1$) et $\lambda \in \mathcal{D}_\lambda \subset \mathbb{R}^q$ ($q \geq 1$) : paramètres incertains respectifs de M1 et M2.
- Comment caler θ et λ en fonction des données observées ($W = w$, $Z = z$) ?

1 Calage joint :

$$\pi_{\text{full}}(\theta, \lambda \mid z, w) = \pi(\theta \mid \lambda, z)\pi(\lambda \mid w, z)$$

[Marque-Pucheu et al., 2016].

2 Calage basé sur des modèles de coupure (ligne rouge en pointillé) :

$$\pi_{\text{cut}}(\theta, \lambda \mid z, w) := \pi(\theta \mid \lambda, z)\pi(\lambda \mid w)$$

[Plummer, 2015, Jacob et al., 2017].

- Cas applicatif \implies estimation de $\pi_{\text{cut}}(\theta, \lambda \mid z, w)$.

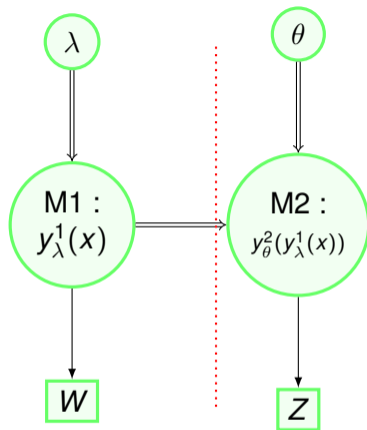


Figure 1 – Deux modèles chaînés.

Cas applicatif : ALCYONE

- Focus sur **les modèles thermique et de comportement des gaz de fission**.
- Modèle thermique (MT) : calcul de l'évolution de la température T dans le cœur du réacteur à eau pressurisée :
 - T fonction du paramètre incertain $\lambda \in \mathcal{D}_\lambda \mathbb{R}$ (**conductivité thermique** du combustible).
 - $\pi(\lambda | w)$ bien estimée à partir des mesures directes w de MT [Bouloré et al., 2023].
- Modèle de comportement des gaz de fission : représentation continue du comportement des produits de fission gazeux.
- $\theta \in \mathcal{D}_\theta$: paramètres incertains du modèle de comportement des gaz de fission.

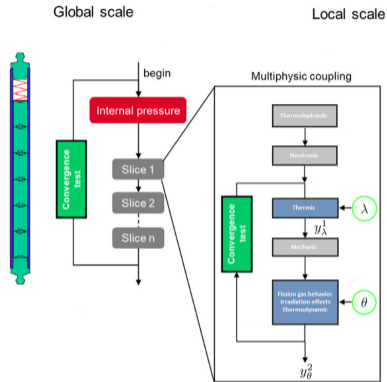


Figure 2 – Chaînage des modèles thermiques et de comportement des gaz de fission de l'application ALCYONE [Michel et al., 2021].

Modèle de comportement des gaz de fission

- Sortie d'intérêt : **Fraction de gaz relâchée** (RGF)

$$y_{\theta, \lambda}(x) := y_{\theta}^2(y_{\lambda}^1(x)).$$

- Relation entre les mesures expérimentales z et les sorties du modèle de comportement des gaz de fission :

$$z_i = y_{\theta, \lambda}(x_i) + \epsilon_{\text{exp}, i}, \quad i = 1, \dots, n,$$
$$\epsilon_{\text{exp}, i} \sim \mathcal{N}(0, \sigma_{\epsilon_{\text{exp}, i}}^2).$$

- **Objectif : estimer** $\pi_{\text{cut}}(\theta, \lambda | z, w) \iff$ **estimer** $\pi(\theta | \lambda, z)$ **car** $\pi(\lambda | w)$ **bien estimée :**

$$\pi(\theta | \lambda, z) \propto \mathcal{L}(z | \theta, \lambda)\pi(\theta | \lambda).$$

En réécrivant par :

$$\pi(\theta(\lambda) | z) \propto \mathcal{L}(z | \theta(\lambda))\pi(\theta(\lambda)).$$



Calage bayésien conditionnel

- En général, une distribution *a posteriori* est connue jusqu'à une constante \implies besoin des algorithmes MCMC.
- **Approche standard** : exécuter autant de MCMC indépendants que le nombre d'échantillons d'intérêt de $\lambda \implies$ pose deux problèmes :
 - Le calcul des RGF étant coûteux \implies exécution des MCMC faite uniquement que pour quelques valeurs de λ .
 - La non prise en compte du fait que $\pi(\theta(\lambda) | z)$ peut donner des informations sur $\pi(\theta(\lambda') | z)$ si λ n'est pas trop éloigné de λ' .
- **La méthode proposée, appelée GP-LinCC** (pour **Gaussian Process and Linearization-based Conditional Calibration**), fournit une distribution de probabilité prédictive *a posteriori* de θ conditionnellement à tout λ .

Plan de la présentation

Calage d'un modèle numérique

Calage des modèles numériques interconnectés

Nouvelle méthode GP-LinCC

Conclusion et perspectives



Méthodologie :

- Pour estimer $\theta(\lambda)$, elle s'appuie sur :
 - la représentation *a priori* de chaque composante de $\theta(\lambda)$ par une trajectoire d'un processus gaussien indépendant [Rasmussen et al., 2006].
 - La linéarité du modèle de comportement des gaz de fission par rapport à $\theta(\lambda)$.
 - L'hypothèse de compensation, avis d'expert, stipule que z est peu informative (comparée à w) sur l'incertitude de λ . Ceci est équivalent à dire que

$$\pi_{\text{cut}}(\theta, \lambda | z, w) \approx \pi_{\text{full}}(\theta, \lambda | z, w)$$

au sens d'une dissimilarité :

- 1 relation entre $\pi_{\text{cut}}(\theta, \lambda | z, w)$ et $\pi_{\text{full}}(\theta, \lambda | z, w)$:

$$\pi_{\text{cut}}(\theta, \lambda | z, w) = \pi_{\text{full}}(\theta, \lambda | z, w) \frac{\pi(z | w)}{\pi(z | \lambda)}.$$

- 2 L'hypothèse de compensation sera satisfaite si $\mathcal{L}(z | \theta, \lambda)$ est non identifiable en (θ, λ) .

En pratique

Ingrédients :

- Un ensemble de données expérimentales $z \in \mathbb{R}^n$ à utiliser pour le calage.
- Un ensemble de simulations du modèle numérique $\{y_{\theta^{(l)}, \lambda_k}(x_i)\}_{l=1}^{n_{sim}} \in \mathbb{R}^{n_{sim}}$ pour estimer tous les coefficients des différents modèles linéaires correspondant à chaque valeur de $D_m := \{\lambda_k\}_{k=1}^m$ et à chaque x_i ($1 \leq i \leq n$) :

$$y_{\theta^{(l)}, \lambda_k}(x_i) = \beta_0(x_i, \lambda_k) + \beta_1^t(x_i, \lambda_k)\theta^{(l)} + \epsilon_{x_i, \lambda_k},$$

où $\epsilon_{x_i, \lambda_k} \sim \mathcal{N}(0, \sigma^2(x_i, \lambda_k))$.

Différentes étapes :

- Étape 1 : Analyse de sensibilité globale : réduction du nombre de paramètres d'entrée du modèle de comportement des gaz de fission
 $\implies p = 11$ à $p = 4$.
- Étape 2 : construction de modèles linéaires.
- Étape 3 : Application de la méthode GP-LinCC pour calibrer $\theta(\lambda) = (\theta_1(\lambda), \dots, \theta_p(\lambda))^t \in \mathbb{R}^p$.
- **Budget total des simulations utilisées :**
 $n \times m \times n_{sim} = 40 \times 20 \times 200 = 160000$ (2 mois).

Comparaison des trois densités à une nouvelle valeur de la conductivité thermique $\lambda^* = 1.099$

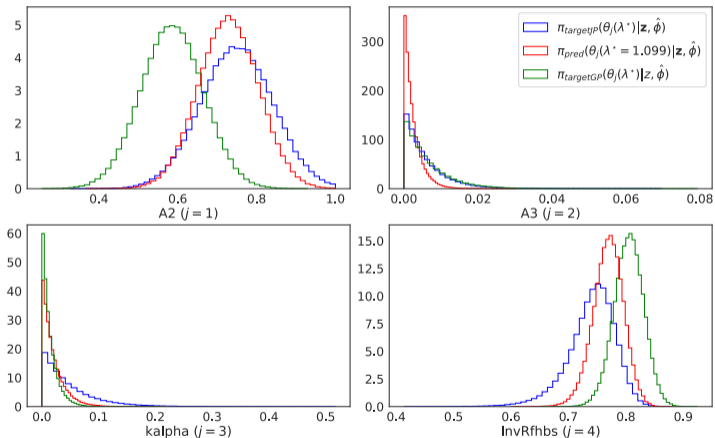


Figure 3 – Comparaison des trois densités pour une nouvelle valeur de la conductivité $\lambda^* = 1.099$.

Hypothèse de compensation vérifiée pour $\lambda_{nom} = 1$, $\lambda_{nom} - 5\%$ et $\lambda_{nom} + 5\%$.

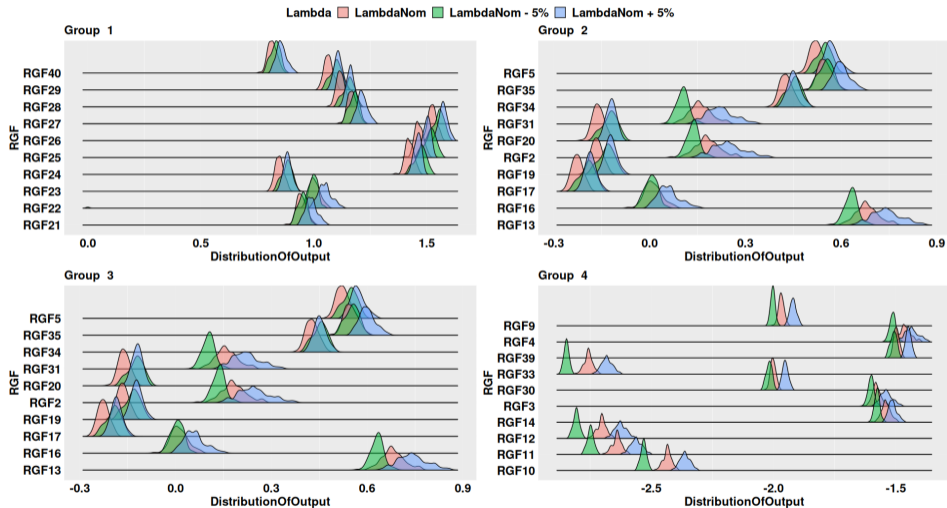


Figure 4 – Comparaison des distributions des 40 log-RGF pour ces trois valeurs de λ .

Comparaison des log-RGF mesurées et calculées en $\lambda_{nom} = 1$, $\lambda_{nom} - 5\%$ et $\lambda_{nom} + 5\%$.

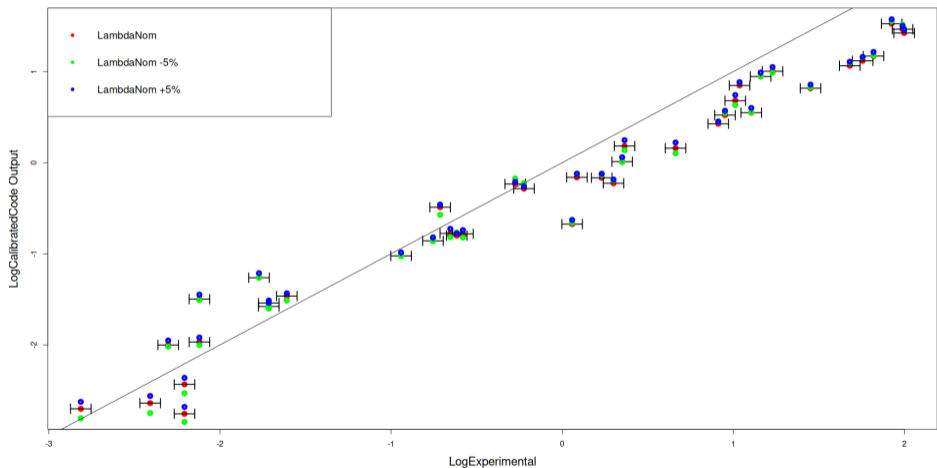


Figure 5 – Comparaison des log-RGF mesurées et calculées en trois valeurs de λ .

Plan de la présentation

Calage d'un modèle numérique

Calage des modèles numériques interconnectés

Nouvelle méthode GP-LinCC

Conclusion et perspectives





Conclusion

- La méthode GP-LinCC peut apprendre la relation entre θ et λ via la fonction de calage $\theta(\lambda)$ à partir des données expérimentales disponibles et d'un plan d'expériences numériques du modèle pour un ensemble de λ .
- La distribution prédictive fournie par GP-LinCC est proche de la distribution *a posteriori* de θ conditionnellement à de nouvelles réalisations λ^* de la conductivité thermique.
- Les intervalles crédibles de la distribution prédictive sont parfois plus petits que ceux des autres. Cela peut s'expliquer par l'estimation *Empirical Bayes* des hyperparamètres des p processus gaussiens [Reich and Ghosh, 2019] ainsi que le choix des fonctions de covariance pour chaque composante de $\theta(\lambda)$ (ici Matérn 5/2).
- Preprint de la méthode GP-LinCC <https://doi.org/10.48550/arXiv.2307.01111>.

Perspectives

- L'extension au cadre non linéaire.
- La prise en compte d'une éventuelle erreur de modèle.

References I

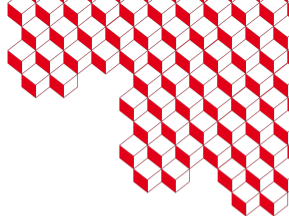
- A. Bouloré, C. Struzik, V. Bouineau, F. Gaudier, G. Damblin, and S. Bernaud. Modelling of UO_2 thermal conductivity : Improvement of the irradiation defects contribution and uncertainty quantification. *Nuclear Engineering and Design*, 407 :112304, 2023.
- E. de Rocquigny, N. Devictor, and S. Tarantola. *Uncertainty in industrial practice : a guide to quantitative uncertainty management*. John Wiley & Sons, 2008.
- P. E. Jacob, L. M. Murray, C. C. Holmes, and C. P. Robert. Better together ? Statistical learning in models made of modules. *arXiv preprint arXiv :1708.08719*, 2017.
- M. C. Kennedy and A. O'Hagan. Bayesian calibration of computer models. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 63(3) :425–464, 2001.
- S. Marque-Pucheu, G. Perrin, and J. Garnier. Calibration of Nested Computer Models. In *VII European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS Congress), Crete Island, Greece, 5-10 June 2016*, 2016.

References II

- B. Michel, I. Ramière, I. Viallard, C. Introini, M. Lainet, N. Chauvin, V. Marelle, A. Boulore, T. Helfer, R. Masson, et al. Two fuel performance codes of the PLEIADES platform : ALCYONE and GERMINAL. In *Nuclear Power Plant Design and Analysis Codes*, pages 207–233. Elsevier, 2021.
- M. Plummer. Cuts in Bayesian graphical models. *Statistics and Computing*, 25 :37–43, 2015.
- C. E. Rasmussen, C. K. Williams, et al. *Gaussian processes for machine learning*, volume 1. Springer, 2006.
- B. J. Reich and S. K. Ghosh. *Bayesian statistical methods*. Chapman and Hall/CRC, 2019.



isds



Merci pour votre attention. Des questions ?

Modélisation statistique de l'approche GP-LinCC

- Modèle statistique : $z_i = r(x_i) + \epsilon_i$, $1 \leq i \leq n$, $\epsilon_{\text{exp}, i} \sim \mathcal{N}(0, \sigma_{\epsilon_{\text{exp}, i}}^2)$ avec $\sigma_{\epsilon_i}^2$ connue.
- **Hypothèse de linéarité** : $\exists \theta(\lambda) \in \mathbb{R}^p$: $y_{\theta(\lambda)}(x_i) = \widehat{\beta}_0(x_i, \lambda) + \widehat{\beta}_1^t(x_i, \lambda)\theta(\lambda) = r(x_i)$, $\forall x_i \in \mathbb{R}^d$ et on a :

$$z_i = \widehat{\beta}_0(x_i, \lambda) + \widehat{\beta}_1^t(x_i, \lambda)\theta(\lambda) + \epsilon_{i, \lambda}. \quad (1)$$

avec $\epsilon_{i, \lambda} := \epsilon_{\text{exp}, i} + \epsilon_{x_i, \lambda} \sim \mathcal{N}(0, \sigma_{\epsilon_{\text{exp}, i}}^2, \widehat{\sigma}_{i, \lambda}^2)$. L'équation (1) en format vectoriel :

$$z = \widehat{\beta}_0(x, \lambda) + \widehat{\beta}_1(x, \lambda)\theta(\lambda) + \epsilon_\lambda \quad (2)$$

où $z = (z_1, \dots, z_n)^t \in \mathbb{R}^n$, $\widehat{\beta}_1(x, \lambda) := (\widehat{\beta}_1^t(x_1, \lambda), \dots, \widehat{\beta}_1^t(x_n, \lambda))^t \in \mathbb{R}^{n \times p}$, $\widehat{\beta}_0(x, \lambda) := (\widehat{\beta}_0(x_1, \lambda), \dots, \widehat{\beta}_0(x_n, \lambda))$
et $\epsilon_\lambda = (\epsilon_{1, \lambda}, \dots, \epsilon_{n, \lambda})^t$.

- **Les quantités $\widehat{\beta}_0(x, \lambda)$ et $\widehat{\beta}_1(x, \lambda)$ sont connus uniquement pour $\lambda \in \{\lambda_1, \dots, \lambda_m\}$.**
- Le modèle numérique $y_{\theta(\lambda)}(x)$ est coûteux en temps de calcul
⇒ il ne peut être linéarisé que pour un nombre finis de λ .

Approche GP-LinCC

■ Étape d'estimation : apprentissage de la fonction $\theta(\lambda)$:

1 Distribution *a priori* : chaque $\theta_j(\lambda) \sim \mathcal{GP}(m_{\beta_j}(\lambda), \sigma_j^2 K_{\psi_j}(\lambda, \lambda'))$, $1 \leq j \leq p$, $\phi := (\beta_j, \sigma_j^2, \psi_j)_{j=1}^p$.

2 Hypothèse de compensation : propagation de m réalisations de λ dans (2) $\implies m$ équations :

$$z = \widehat{\beta}_0(\mathbf{x}, \lambda_k) + \widehat{\beta}_1(\mathbf{x}, \lambda_k)\theta(\lambda_k) + \epsilon_{\lambda_k}, \quad 1 \leq k \leq m$$

$$(z, \dots, z) = \left(\widehat{\beta}_0(\mathbf{x}, \lambda_1), \dots, \widehat{\beta}_0(\mathbf{x}, \lambda_m) \right) + \left(\widehat{\beta}_1(\mathbf{x}, \lambda_1)\theta(\lambda_1), \dots, \widehat{\beta}_1(\mathbf{x}, \lambda_m)\theta(\lambda_m) \right) + (\epsilon_{\lambda_1}, \dots, \epsilon_{\lambda_m}).$$

On pose $\mathbf{z} := (z, \dots, z) \in \mathbb{R}^{n \times m}$ et $\Theta_m := (\theta(\lambda_1), \dots, \theta(\lambda_m))^t \in \mathbb{R}^{m \times p}$.

3 Distribution *a posteriori* $\pi(\Theta_m | \mathbf{z}, \phi)$ par la **formule de Bayes** :

$$\pi(\Theta_m | \mathbf{z}, \phi) \propto \mathcal{L}(\mathbf{z} | \Theta_m) \pi(\Theta_m | \phi). \quad (3)$$

■ Étape de prédiction :

1 Distribution prédictive $\pi_{pred}(\theta(\lambda^*) | \mathbf{z}, \phi)$ pour de nouvelles réalisations λ^* :

$$\pi_{pred}(\theta(\lambda^*) | \mathbf{z}, \phi) = \int \pi(\theta(\lambda^*) | \Theta_m, \phi) \pi(\Theta_m | \mathbf{z}, \phi) d\Theta_m,$$

$$\widehat{\theta}_{pred}(\lambda^*) = \operatorname{argmax} \pi_{pred}(\theta(\lambda^*) | \mathbf{z}, \phi).$$