



Sujet de stage: Calibration bayésienne de modèles physiques de thermohydraulique

Clément Gauchy*, Pierre Barbillon†

Description du sujet

Dans le cadre du développement d'outils de simulation en physique des réacteurs, il est nécessaire de calibrer des paramètres de modèles physiques sur des données expérimentales en tenant compte des différentes sources d'incertitudes. Une première source d'incertitude vient de l'aspect expérimental des mesures physiques, et une deuxième source d'incertitude provient d'un écart entre le phénomène physique d'intérêt et le modèle physique choisi, que l'on appelle alors *erreur de modèle*. La prise en compte et la détermination de cette erreur de modèle est un sujet de recherche actif avec plusieurs modélisations possibles. En toute généralité, on considère qu'on dispose de n mesures du phénomène physique d'intérêt notées : $(y_i^{\text{mes}})_{1 \leq i \leq n} \in \mathbb{R}$. À ces mesures physiques correspondent des observations des conditions expérimentales notées $(x_i)_{1 \leq i \leq n} \in \mathbb{R}^d$ qui sont des entrées du modèle physique noté g . Ce modèle dépend également de paramètres $\theta \in \mathbb{R}^p$ à calibrer.

Dans [5], un modèle statistique est proposé pour mettre en relation les observations physiques et les sorties du modèle physique :

$$y_i^{\text{mes}} = g(x_i, \theta) + \delta(x_i) + \varepsilon_i, \quad (1)$$

où les ε_i représentent un bruit d'observation et sont supposés être des variables aléatoires indépendantes identiquement distribuées de loi gaussienne centrée de variance notée σ_ε^2 ($\varepsilon_i \stackrel{iid}{\sim} \mathcal{N}(0, \sigma^2)$) et δ représente l'erreur de modèle et est modélisée par un processus gaussien de moyenne nulle et de noyau de covariance C .

L'erreur de modèle peut également être modélisée par une structure latente sur le paramètre θ [6]. Ainsi le modèle statistique se réécrit comme

$$\begin{array}{l} y_i^{\text{mes}} \\ \theta_i \end{array} \stackrel{iid}{\sim} \begin{array}{l} g(x_i, \theta_i) + \varepsilon_i \\ p_\alpha \end{array} \quad (2)$$

où p_α représente une famille paramétrique de lois de probabilités. L'estimation du paramètre α peut s'effectuer par un algorithme Expectation Maximization stochastique [1] ou par une inférence bayésienne [2].

On peut également utiliser la méthodologie CIRCE (*Calcul des Incertitudes Relatives aux Corrélations Élémentaires*) développée au CEA, qui consiste en une approximation

*CEA Saclay, ISAS/DM2S/SGLS/LIAD, clement.gauchy@cea.fr

†AgroParisTech, UMR MIA, pierre.barbillon@agroparistech.fr

linéaire du problème d'inversion stochastique présenté dans l'équation 2.

$$y_i^{\text{mes}} = g(x_i, \theta_0) + \nabla_{\theta} g(x_i, \theta_0)^T \lambda_i + \varepsilon_i \quad (3)$$

où $\lambda_i = \theta_i - \theta_0$ avec $\lambda_i \stackrel{iid}{\sim} \mathcal{N}(\mu, \Sigma)$. Les paramètres à estimer dans ce modèle sont (μ, Σ) via un algorithme ECME (*Expectation/Conditional Maximisation Either*) [3].

Axes de travail Les objectifs du stage sont

- Mettre en œuvre et comparer les 3 méthodologies de calibration évoquées ci-dessus sur des modèles physiques simples de transfert thermique entre un canal aux parois chauffées et un mélange diphasique d'eau et de vapeur.
- Proposer une méthodologie de sélection de modèles de type BMA (*Bayesian Model Averaging*) ([4]) dans le cas où plusieurs modèles physiques g sont disponibles.

En fonction des résultats obtenus, la rédaction d'un article scientifique sur les travaux menés pendant le stage est envisageable.

Encadrement Le sujet de stage est proposé par le Groupement d'Intérêt Scientifique (GIS) [LARTISSTE](#) et se déroulera au laboratoire [MIA Paris-Saclay](#) sur le campus d'AgroParisTech. Il sera encadré par [Pierre Barbillon](#), professeur en statistiques à AgroParisTech, et [Clément Gauchy](#), ingénieur-chercheur au CEA Saclay. Des réunions hebdomadaires seront programmées durant le stage. La durée de stage sera de 6 mois et commencera au deuxième trimestre 2025.

Profil Le profil recherché est celui d'un étudiant ou d'une étudiante ayant un niveau M2 de mathématiques appliquées avec une spécialisation en statistiques/machine learning. Des connaissances en physiques seront appréciées mais non nécessaires.

Contact Un CV, une lettre de motivation, et un relevé de notes de M1 sont à envoyer à pierre.barbillon@agroparistech.fr, clement.gauchy@cea.fr

Références

- [1] P. Barbillon, G. Celeux, G. Grimaud, Y. Lefebvre, and E. De Rocquigny. Nonlinear methods for inverse statistical problems. *Computational Statistics & Data Analysis*, 55(1) :132–142, 2011. ISSN 0167-9473. doi:doi.org/10.1016/j.csda.2010.05.030.
- [2] Shuai Fu, Gilles Celeux, Nicolas Bousquet, and Mathieu Couplet. Bayesian inference for inverse problems occurring in uncertainty analysis. *International Journal for Uncertainty Quantification*, 5(1), 2015.
- [3] Y. Lefebvre G. Celeux, A. Grimaud and E. de Rocquigny. Identifying intrinsic variability in multivariate systems through linearized inverse methods. *Inverse Problems in Science and Engineering*, 18(3) :401–415, 2010. doi:[10.1080/17415971003624330](https://doi.org/10.1080/17415971003624330).
- [4] Jennifer A Hoeting, David Madigan, Adrian E Raftery, and Chris T Volinsky. Bayesian model averaging : a tutorial (with comments by m. clyde, david draper and ei george, and a rejoinder by the authors. *Statistical science*, 14(4) :382–417, 1999.
- [5] Marc C. Kennedy and Anthony O'Hagan. Bayesian calibration of computer models. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 63(3) : 425–464, 2001. doi:[10.1111/1467-9868.00294](https://doi.org/10.1111/1467-9868.00294).

- [6] Khachik Sargsyan, Xun Huan, and Habib N. Najm. Embedded model error representation for bayesian model calibration. *International Journal for Uncertainty Quantification*, 9(4) :365–394, 2019. ISSN 2152-5080.