

Plans d'expériences séquentiels pour la calibration des modèles numériques coûteux

14 février 2014

Guillaume Damblin ¹ & Pierre Barbillon ² & Merlin Keller ¹ & Alberto Pasanisi ¹ & Eric Parent ²

¹ EDF R&D, 6 quai Watier BP 49 78 001 Chatou,
guillaume.damblin@edf.fr

² AgroParisTech, 16 rue Claude Bernard 75 005 Paris,
pierre.barbillon@agroparistech.fr

Résumé. La calibration consiste à ajuster un modèle numérique aux expériences physiques disponibles. Pour exécuter une simulation, il faut renseigner en entrée du modèle les valeurs des variables contrôlées du phénomène physique, mais aussi les valeurs des paramètres intrinsèques au modèle. En pratique ces dernières sont souvent entachées de beaucoup d'incertitude. La calibration statistique consiste à les estimer pour faire coïncider au mieux les sorties du modèle numérique avec les expériences physiques disponibles. La statistique bayésienne permet de quantifier aisément l'incertitude sur les valeurs estimées mais également de tenir compte d'éventuelles connaissances a priori sur les valeurs de ces paramètres. L'inférence repose alors souvent sur des méthodes de Monte Carlo nécessitant un nombre conséquent d'appels au modèle numérique.

En pratique, le modèle est souvent coûteux, une simulation nécessitant parfois plusieurs heures, voir plusieurs jours de calcul. Une approximation du modèle (émulateur) basée sur un *a priori* de type processus gaussien peut être calculée à partir d'un nombre restreint de simulations appelé plan d'expériences. La qualité de l'approximation dépend fortement du choix de celui-ci. Nous proposons dans ce travail des méthodes séquentielles pour construire un plan d'expériences minimisant l'erreur de calibration imputée à l'approximation du modèle numérique par l'émulateur. Une première stratégie consiste à adapter le critère EI (Expected Improvement) à la problématique de la calibration. La seconde consiste à minimiser séquentiellement un coût quantifiant l'incertitude de calibration (stratégie SUR, Stepwise Uncertainty Reduction). Des exemples numériques sont présentés pour illustrer la pertinence des méthodes proposées.

Mots-clés. émulateur processus gaussien, calibration, expériences numériques, plans d'expériences séquentiels.

Abstract. Calibration is the confrontation between real field experiments and numerical experiments i.e. runs of a computer model. To run the computer model, a set of inputs has to be specified. Some inputs are aligned on the field experiments, the other inputs usually called parameters are specific to the computer model and unknown although prior information may be available. The goal of statistical calibration consists of tuning these parameters to make the outputs of the computer model as close as possible to the field data. The uncertainties on these parameters and the prior information can be taken into account thanks to a Bayesian analysis. The inference is then conducted through MCMC techniques which need lots of runs of the model.

However, the model is usually time-consuming. Thus, an emulator (cheap approximation) is built by setting a Gaussian process prior on the computer model. The design of numerical experiments, *i.e.* selected inputs wherein the model is evaluated, strongly impacts the calibration process when replacing the model by an emulator. That is why we aim at adapting sequentially this design in order to reduce the uncertainties for calibration due to the emulator. We propose two different strategies to achieve this goal. One strategy is based on the Expected Improvement criterion which is adapted to calibration and the other is a stepwise uncertainty reduction (SUR) strategy. Numerical illustrations are provided to illustrate the efficiency of these strategies.

Keywords. Gaussian process emulator, calibration, computer experiments, adaptive design of experiment.

Résumé long

En ingénierie, la simulation occupe désormais une place prépondérante dans l'étude des phénomènes physiques complexes. Lorsqu'une expérience physique est irréalisable, car trop coûteuse ou impossible à mettre en œuvre, l'expérience numérique constitue un outil précieux se substituant à l'expérience physique. La confiance attribuée à celle-ci dépend bien-sûr de la qualité du modèle mathématique sous-jacent (décrivant le système physique) mais aussi de la précision du calcul numérique. Ces incertitudes doivent être traitées attentivement dans le cadre de la vérification et de la validation du modèle [7], afin de pouvoir quantifier la capacité de celui-ci à prédire correctement la quantité physique d'intérêt.

Nous considérons dans notre présentation le cas d'un modèle numérique $Y : (x, \theta) \in \mathcal{X} \times \Theta \rightarrow Y(x, \theta)$, x étant un vecteur de variables contrôlées et θ un vecteur de constantes du modèle, indépendant de x . La calibration de Y consiste à estimer θ de telle façon que Y reproduise le plus précisément possible le phénomène physique qu'il modélise. À partir des mesures physiques disponibles $z = (z_1, \dots, z_n)$, le vecteur θ peut être estimé par la méthode déterministe des moindres carrés. Une justification possible de cette méthode consiste à faire l'hypothèse que le modèle Y représente exactement le phénomène physique, modulo l'erreur de mesure des observations. En effet, lorsque les erreurs d'observations notées ϵ_i sont gaussiennes, centrées et identiquement distribuées, l'estimateur du maximum de vraisemblance de θ dans le modèle statistique

$$z_i = Y(x_i, \theta) + \epsilon_i \quad (1)$$

avec $\epsilon_i \sim \mathcal{N}(0, \lambda)$, est équivalent à celui des moindres carrés :

$$\hat{\theta} = \operatorname{argmin}_{\theta} \sum_{i=1}^n (z_i - Y(x_i, \theta))^2. \quad (2)$$

Pour résoudre ce problème d'optimisation, il faut pouvoir évaluer un grand nombre de fois le modèle numérique Y sur $\mathcal{X} \times \Theta$ (cela revient à considérer l'exécution d'une simulation négligeable en temps CPU), ce qui nous permet d'obtenir en un temps réduit la valeur de $Y(x, \theta)$ pour n'importe quelle configuration $(x, \theta) \in \mathcal{X} \times \Theta$. Malheureusement, ce n'est pas souvent possible en pratique étant donné la complexité des systèmes physiques étudiés et donc des modèles utilisés. Dans ce type de situation, un budget limité de N simulations $\{Y(x_1, \theta_1), \dots, Y(x_N, \theta_N)\}$ est alloué pour la construction d'un méta-modèle (émulateur) \hat{Y} de Y qui fournit une prédiction de la valeur de Y pour tout $(x, \theta) \in (\mathcal{X}, \Theta)$. Le jeu de simulations $\{(x_1, \theta_1), \dots, (x_N, \theta_N)\}$ est appelé plan d'expériences. Habituellement, \hat{Y} est modélisé comme un processus Gaussien stationnaire conditionné par $\{Y(x_1, \theta_1), \dots, Y(x_N, \theta_N)\}$ et noté \hat{Y} :

$$\hat{Y}(x, \theta) \sim \mathcal{N}(m_{(x,\theta)}(\beta), \Sigma_{(x,\theta)}(\Phi))$$

avec (β, Φ) un vecteur d'hyper-paramètres caractérisant la loi du processus. Cette technique appelée *krigeage* tire son origine de la géostatistique et a été par la suite mise en œuvre en simulation numérique [6]. Le problème d'optimisation (2) se réécrit maintenant :

$$\hat{\theta} = \operatorname{argmin}_{\theta} \sum_{i=1}^n (z_i - m_{(x,\theta)}(\beta))^2 \quad (3)$$

Une limite importante de l'estimateur (3) (méthode ANLS pour *Approximated Non Linear Square*) est la non prise en compte de l'erreur de prédiction (quantifiée par $\Sigma_{(x,\theta)}$) dans la résolution du problème. Elle peut être prise en compte dans la vraisemblance du modèle statistique (1). En effet, lorsque Y est coûteux, la vraisemblance n'est plus fonction uniquement du paramètre de calibration et de la variance λ du bruit de mesure, mais également des paramètres du processus gaussien modélisant Y . Notons $d = (z, \{Y(x_1, \theta_1), \dots, Y(x_N, \theta_N)\})$ et L la vraisemblance du modèle. Alors :

$$\log L(d|\beta, \Phi, \theta, \lambda) = -\frac{n}{2} \times (d - m_{(x,\theta)}(\beta) \mathbf{1}_{n+N})^T \Sigma_d^{-1} (d - m_{(x,\theta)}(\beta) \mathbf{1}_{n+N}) - \frac{1}{2} \log |\Sigma_d| - \frac{n}{2} \log 2\Pi \quad (4)$$

avec,

$$\Sigma_d = \Sigma_{(x,\theta)}(\Phi)(d) + \begin{pmatrix} \lambda \operatorname{Id}_n & 0 \\ 0 & 0 \end{pmatrix} \quad (5)$$

En pratique une estimation en deux étapes est préconisée [2] : d'abord, les paramètres du processus Gaussien (β, Φ) sont estimés par maximum de vraisemblance à partir des simulations $\{Y(x_1, \theta_1), \dots, Y(x_N, \theta_N)\}$, puis injectés dans l'expression de la vraisemblance des mesures physiques (*separated mle*) :

$$\log L(z|\theta, \lambda) = -\frac{n}{2} \times (z - m_{(x,\theta)}(\hat{\beta}) \mathbf{1}_n)^T \Sigma_z^{-1} (z - m_{(x,\theta)}(\hat{\beta}) \mathbf{1}_n) - \frac{1}{2} \log |\Sigma_z| - \frac{n}{2} \log 2\Pi \quad (6)$$

avec $\Sigma_z = \Sigma_{(x,\theta)}(\hat{\Phi})(z) + \lambda \text{Id}_n$. Ensuite, il reste à maximiser (6) en (θ, λ) . Évidemment, lorsque \hat{Y} prédit suffisamment bien Y , l'estimateur (3) tend vers l'estimateur (2). Beaucoup de travaux récents s'attachent à comprendre l'impact du plan d'expériences $\{(x_1, \theta_1), \dots, (x_N, \theta_N)\}$ pour la construction de \hat{Y} . L'idée la plus répandue consiste à construire un plan d'expériences occupant de façon optimale le domaine $\mathcal{X} \times \Theta$ au sens de critères géométriques et statistiques [5, 3, 4]. Dans cette optique, les plans LHS *maximin* sont souvent utilisés pour construire un méta-modèle. Pourvu que la fonction Y soit suffisamment régulière, les méta-modèles estimés à partir de ces plans ont la capacité de prédire Y suffisamment précisément pour tout $(x, \theta) \in \mathcal{X} \times \Theta$. Des critères globaux tel l'IMSE (*Integrated Mean Square Error*) permettent de quantifier les performances du méta-modèle en prédiction et se révèlent être des indicateurs intéressants pour juger de la qualité du méta-modèle et donc du plan d'expérience. Il est d'ailleurs possible d'enrichir séquentiellement un plan initial pour améliorer la valeurs de ces critères.

En outre, ces stratégies séquentielles sont d'autant plus pertinentes à mettre en œuvre que l'objectif à atteindre n'est plus directement la prédiction d'une fonction coûteuse sur son domaine de variation. En effet, des plans d'expériences séquentiels ont été proposés dans la littérature pour l'optimisation de fonctions coûteuses (algorithme EGO, *Efficient Global Optimization* basé sur le critère EI) mais aussi pour l'estimation de probabilités de défaillance (critère SUR [1]). En effet, au regard de ces deux objectifs, vouloir prédire précisément la fonction sur le domaine de variation (pour tout $(x, \theta) \in \mathcal{X} \times \Theta$) n'est plus l'option la plus judicieuse. Par exemple, en ce qui concerne l'optimisation d'une fonction, la capacité du méta-modèle à fournir une bonne prédiction est importante seulement dans les zones contenant potentiellement l'optimum de la fonction. L'algorithme séquentiel permet plus facilement la détection de ces zones d'intérêts.

Pour notre objectif de calibration basé sur le critère des moindres carrés, la vraisemblance (6) dépend uniquement des valeurs de Y aux points d'expériences (x_i, θ) pour tout $i = 1, \dots, n$. Dans ce travail, nous appréhendons le problème de calibration (2) comme un problème d'optimisation de la fonction $\theta \rightarrow \sum_{i=1}^n (z_k - Y(x_k, \theta))^2$ conditionnellement au fait que seul un nombre limité d'appel à Y est autorisé. Nous proposons un algorithme séquentiel de construction d'un plan d'expériences pour l'estimation de \hat{Y} en adaptant l'algorithme EGO et le critère SUR à ce problème. Les méthodes bayésiennes qui permettent de prendre en compte des informations *a priori* sur la valeur de θ sont mises en pratique avec \hat{Y} construit à partir de ces algorithmes séquentiels, puis comparées avec \hat{Y} construit à partir d'un plan LHS *maximin*. Sur des fonctions tests, nous illustrons les apports de la méthode en terme de robustesse des estimateurs calculés et relativement à la taille du plan d'expériences.

Références

- [1] J. Bect, D. Ginsbourger, L. Ling, V. Picheny, and E. Vaquez. Sequential design of computer experiments for the estimation of a probability of failure. *Statistics and Computing*, 22(3) :773–793, 2012.

- [2] D.D. Cox, J.S. Park, and E.S. Clifford. A statistical method for tuning a computer code to a data base. *Computational Statistics and Data Analysis*, 37 :77–92, 2001.
- [3] K. Fang, R. Li, and A. Sudijanto. *Design and modeling for computer experiments*. 2006.
- [4] J. Franco. *Planification d'expériences numériques en phase exploratoire pour la simulation des phénomènes complexes*. PhD thesis, Ecole Nationale Supérieure des Mines de Saint-Etienne, 2008.
- [5] J.R. Koehler and A.B. Owen. Computer experiments. *Handbook of Statistics, Vol.13*, 1996.
- [6] J. Sacks, Welch. W.J., T.J. Mitchell, and H.P Wynn. Design and analysis of computer experiments. *Technometrics*, 31 :41–47.
- [7] T.G. Trucano, L.P. Swiler, T. Igusa, W.L. Oberkampf, and M. Pilch. Calibration, validation and sensitivity analysis : What's what. *Reliability Engineering and System Safety*, 91 :1331–1357, 2006.