

Abstract

To understand and to be able to forecast natural phenomena is increasingly important nowadays, as those predictions are often the basis of many decisions, whether economical or ecological. In order to do so, mathematical models are introduced to represent the reality at a specific scale, and are then implemented numerically. However in this process of modelling, many complex and subscale phenomena have to be simplified, often through the introduction of additional parameters in order to account for those unresolved processes, but those parameters need to be properly estimated. Classical methods of estimation usually involve an objective function, that measures the distance between the simulations and some observations, which is then optimised. Such an optimisation require many runs of the numerical model and possibly the computation of its gradient, thus can be expensive to evaluate computational-wise.

However, some other uncertainties can also be present, which represent some uncontrollable and external factors that affect the modelling. Those variables will be qualified as *environmental*. By modelling them with a random variable, the objective function is then a random variable as well, that we wish to minimise in some sense. Omitting the random nature of the environmental variable can lead to localised optimisation, and thus a value of the parameters that is optimal only for the fixed nominal value. To overcome this, the minimisation of the expected value of the objective function is often considered in the field of optimisation under uncertainty for instance.

In this thesis, we focus instead on the notion of regret, that measures the deviation of the objective function from its optimal value given a realisation of the environmental variable. This regret (either additive or relative) translates a notion of robustness through its probability of exceeding a specified threshold. So, by either controlling the threshold or the probability, we can define a family of estimators based on this regret.

The regret can quickly become expensive to evaluate since it requires an optimisation of the objective for every realisation of the environmental variable. We then propose to use Gaussian Processes (GP) in order to reduce the computational burden of this evaluation. In addition to that, we propose a few adaptive methods in order to improve the estimation : the next points to evaluate are chosen sequentially according to a specific criterion, in a Stepwise Uncertainty Reduction (SUR) strategy.

Finally, we will apply some of the methods introduced in this thesis on an academic problem of parameter estimation. We will study the calibration of the bottom friction of a model of the Atlantic ocean near the French coasts, while introducing some uncertainties in the forcing of the tide, and get a robust estimation of this friction parameter in a twin experiment setting.

Keywords : Optimisation under uncertainties ; Robust calibration ; Gaussian Processes ; Ocean modelling ; Regret

* * *

Résumé

De nombreux phénomènes physiques sont modélisés afin d'en mieux connaître les comportements ou de pouvoir les prévoir. Cependant pour représenter la réalité, de nombreux processus doivent être simplifiés, car ils sont souvent trop complexes, ou apparaissent à échelle bien inférieure à celle du maillage. Au lieu de complètement omettre, les effets des processus sont souvent retranscrits dans les modèles à l'aide de paramétrisations, c'est-à-dire en introduisant des termes représentant des caractéristiques physiques, qui doivent être ensuite estimées. Les méthodes classiques d'estimation se basent sur la définition d'une fonction objectif qui mesure l'écart entre le modèle numérique et la réalité, qui est ensuite optimisée.

Cependant, au delà de l'incertitude sur la valeur du paramètre à estimer, un autre type d'incertitude peut aussi être présent. Cela permet de représenter la variabilité intrinsèque de certains processus externes, qui vont avoir un effet sur la modélisation. Ces variables vont être qualifiées d'*environnementales*. En les modélisant à l'aide d'une variable aléatoire, la fonction objectif devient à son tour une variable aléatoire, que l'on va chercher à minimiser dans un certain sens. Si on omet ce caractère aléatoire, on peut se retrouver avec un paramètre optimal uniquement pour la valeur nominale du paramètre environnemental, et le modèle peut s'éloigner de la réalité pour d'autres réalisations. Ce problème d'optimisation sous incertitudes est souvent abordé en optimisant les premiers moments de la variable aléatoire, l'espérance en particulier.

Dans cette thèse, nous nous intéressons plutôt à la notion de regret, qui mesure l'écart entre la fonction objectif et la valeur optimale qu'elle peut atteindre, pour la réalisation de la variable environnementale donnée. Cette idée de regret (additif ou bien relatif), nous permet de proposer une notion de robustesse à travers l'étude de sa probabilité de dépasser un certain seuil, ou inversement à travers le calcul de ses quantiles. À l'aide de ce seuil, ou de l'ordre du quantile choisi, on peut donc définir une famille d'estimateurs basés sur le regret.

Néanmoins, le calcul du regret, et donc des quantités dérivées peut vite devenir très coûteux, car il nécessite une optimisation par rapport au paramètre de contrôle. Nous proposons donc d'utiliser des processus Gaussiens (GP) afin de construire un modèle de substitution, et donc de réduire cette contrainte en pratique. Nous proposons aussi des méthodes itératives basées notamment sur la stratégie SUR (Stepwise Uncertainty Reduction, Réduction d'incertitudes séquentielle) : le point à évaluer ensuite est choisi selon un critère permettant d'améliorer au mieux des quantités associées au regret-relatif.

Enfin, nous appliquerons les outils présentés dans cette thèse à un problème académique d'estimation de paramètre. Nous étudierons ainsi la calibration sous incertitudes du paramètre de friction de fond d'un modèle océanique, représentant la façade atlantique des côtes françaises, ainsi que la Manche dans un cadre d'expériences jumelles.

Mots-Clés : Optimisation sous incertitudes ; calibration robuste ; Processus Gaussiens ; Modélisation de l'océan ; Regret