

Summary

In the geosciences, inversion problems emerge from the need to estimate environmental variables from indirect measurements. This thesis concentrates on the inference of geophysical or hydrogeological properties using geophysical or hydrological measurements. To enable uncertainty quantification, we employ a probabilistic inversion approach within a Bayesian framework, focusing on the posterior probability density function of the model parameters given the observed data. We address the following two research questions: (1) How can we efficiently solve Bayesian inverse problems involving intractable likelihood functions? (2) How can the combination of inversion and risk assessment methods enhance the accuracy of rare event probability estimation? The likelihood function is intractable in latent variable models, where the relationship between target model parameters and measurements is obscured by an intermediate (latent) variable. We consider cases involving uncertain petrophysical relationships and hyperparameter estimation, respectively. While the former is concerned with estimating hydrogeological parameters from geophysical data by treating the intermediate geophysical properties as latent variables, the latter targets hyperparameters (such as mean, standard deviation and integral scales) by considering the local properties of the field as latent variables. Performing inversion in both situations necessitates the estimation of an intractable likelihood function. To address this challenge, we employ two methods: the correlated pseudo-marginal method, which involves Monte Carlo averaging over samples of the latent variable, and a Gaussian approximation based on local linearization of the geophysical forward operator. In the context of petrophysical uncertainty, we find that the correlated pseudo-marginal method allows for accurate estimation, even in scenarios with high petrophysical uncertainty. In contrast, the less computationally intensive linearized Gaussian approach becomes gradually less accurate as the uncertainty in the petrophysical relationship increases. In the realm of hyperparameter estimation, we concentrate on the correlated pseudo-marginal method, showcasing its capacity to enhance the accuracy of hyperparameter estimation. Our second research question focuses on rare event estimation, for which the main objective is not to determine the posterior distribution of model parameters, but to characterize the distribution of a quantity of interest depending on these parameters through a non-linear relationship. Specifically, we aim to calculate the probability of this quantity assuming critical values, representing a rare event. To handle the associated challenges of rare event estimation, we employ two approaches: Sequential

Monte Carlo combined with subset sampling and an energy-based model approach utilizing a bias potential. We evaluate the performance of the proposed methods using problems related to groundwater hazards and illustrative examples from the engineering literature. Encouragingly, both approaches demonstrate the ability to accurately estimate rare event probabilities smaller than one in a million. In summary, this thesis presents methodological advances to improve the computational efficiency and realism of probabilistic Bayesian inversion approaches for targeted estimation of properties and attributes of environmental systems.

Résumé

Dans le domaine des géosciences, les problèmes inverses émergent de la nécessité d'estimer des variables environnementales à partir de mesures indirectes. Cette thèse se concentre sur l'inférence de propriétés géophysiques ou hydrogéologiques à partir de mesures géophysiques ou hydrologiques. Pour permettre une quantification des incertitudes associées nous employons une approche d'inversion probabiliste dans un cadre bayésien, en nous concentrant sur la distribution a posteriori des paramètres du modèle compte tenu des données observées. Nous abordons les deux questions de recherche suivantes: (1) Comment adapter efficacement des méthodes d'inversion bayésiennes en présence de variables latentes? (2) Comment combiner les approches d'inversion et d'évaluation des risques pour améliorer l'exactitude dans l'estimation de probabilités d'événements rares? La fonction de vraisemblance est difficile à cerner dans les modèles à variables latentes, où la relation entre les paramètres du modèle cible et les mesures est opacifiée par les incertitudes résiduelles sur lesdites variables latentes. Nous examinons des cas impliquant respectivement des relations pétrophysiques incertaines et l'estimation d'hyperparamètres. Alors que le premier cas concerne l'estimation de paramètres hydrogéologiques à partir de données géophysiques en traitant les propriétés géophysiques intermédiaires comme des variables latentes, le second cas cible des hyperparamètres (tels que la moyenne, l'écart-type et les échelles d'intégrales) en considérant cette fois les propriétés locales du champ comme des variables latentes. L'inversion nécessite dans les deux situations l'estimation d'une fonction de vraisemblance complexe. Pour relever ce défi, nous utilisons deux méthodes: la méthode pseudo-marginale corrélée, qui implique une moyenne de Monte Carlo sur des échantillons de la variable latente, et une approximation gaussienne basée sur la linéarisation locale de la réponse géophysique ciblée. Dans le contexte de l'inversions en présence de relations pétrophysiques incertaines, nous constatons que la méthode pseudo-marginale corrélée permet une estimation précise, même dans des scénarios à

forte incertitude pétrophysique. En revanche, l'approche gaussienne linéarisée, moins gourmande en ressources computationnelles, perd progressivement en précision à mesure que l'incertitude de la relation pétrophysique augmente. Dans le cas de l'estimation des hyperparamètres, nous nous concentrons sur la méthode pseudo-marginale corrélée et constatons qu'elle peut permettre d'améliorer l'estimation des hyperparamètres. Notre deuxième question de recherche porte sur l'estimation d'événements rares, pour laquelle l'objectif principal n'est pas de déterminer la distribution postérieure des paramètres du modèle, mais plutôt de caractériser la distribution d'une quantité d'intérêt dépendant de manière non linéaire de ces paramètres. Plus précisément, nous cherchons à calculer la probabilité que cette quantité prenne une valeur critique, qui représente un événement rare. Pour relever les défis associés à l'estimation des événements rares, nous utilisons deux approches : d'une part, une méthode de Monte Carlo séquentielle combinée à un sous-échantillonnage et, d'autre part, une approche fondée sur des modèles génératifs reposant sur le concept d'énergie libre (Energy-based models; EBM) et utilisant un potentiel de biais. Nous évaluons les performances des méthodes proposées en utilisant des problèmes simplifiés liés aux risques de contamination des eaux souterraines et des exemples tirés de la littérature de l'estimation de probabilité d'événements rares en ingénierie. Il est encourageant de constater que les deux approches permettent d'estimer avec précision des probabilités d'événements rares d'occurrence inférieure à un sur un million. En résumé, cette thèse présente des avancées méthodologiques visant à améliorer l'efficacité computationnelle et le réalisme des approches d'inversion probabiliste bayésienne pour l'estimation ciblée de propriétés et attributs des systèmes environnementaux.