

Estimation des paramètres et apprentissage statistique en grande dimension : application à la modélisation du transport de solutés dans des écoulements souterrains

Stage de fin d'études 2026, proposé par EDF R&D, Département LNHE, groupe QE3S

Contexte

EDF R&D a pour missions principales de contribuer à l'amélioration de la performance des unités opérationnelles du groupe EDF et à la préparation de l'avenir énergétique grâce à ses chercheurs et ses partenariats, ses moyens d'essais et sa capacité numérique. Le **LNHE** (Laboratoire National d'Hydraulique et Environnement) porte les compétences nécessaires à la réalisation de ces missions en hydrogéologie, hydrologie, hydraulique, sédimentologie, hydrobiogéochimie et chimie environnementale. Il étudie les interactions des ouvrages de production d'EDF avec l'eau et l'environnement, notamment pour ce qui concerne la sécurité et le dimensionnement des installations, la performance des ouvrages de production et l'acceptabilité des ouvrages de production et leur intégration dans l'environnement. Le groupe de recherche "Qualité des eaux de surface, des eaux souterraines et des sols" (**QE3S**) contribue à la maîtrise et à la réduction des impacts environnementaux des ouvrages de production du groupe EDF sur toutes les matrices environnementales, à court et long terme. Le groupe s'intéresse pour cela aux transferts dans le milieu naturel, dont celui imposé par les écoulements souterrains, au droit de ses sites nucléaires.

L'écoulement des eaux souterraines et le transport d'espèces chimiques en nappe sont conditionnés par les propriétés physiques du sous-sol, milieu complexe et mal connu. Le processus de modélisation doit alors valoriser au mieux les données dont on dispose, tout en rendant compte avec transparence de notre connaissance incertaine du milieu. Le stage vise à explorer et comparer deux méthodologies de modélisation numérique appliquées au suivi environnemental de la qualité des eaux souterraines.

Les méthodes évaluées ont pour objectif de faciliter l'intégration et la valorisation de données souvent peu ou pas exploitées dans la chaîne complète de modélisation. Elles permettent, dans un cadre global d'utilisation de techniques géostatistiques, la quantification des incertitudes en complément de la résolution du problème étudié. Les approches comparées présentent des logiques distinctes :

- Une approche de modélisation inverse, qui vise à estimer les paramètres du modèle en assimilant les données observées (*Iterative Ensemble Smoother*, White, 2018). Différentes stratégies de paramétrisation du modèle peuvent être considérées, comme la méthode des points pilotes (de Marsily, 1978).
- Une approche bayésienne, ne reposant pas sur l'estimation directe d'un ensemble de paramètres, mais sur l'apprentissage de la relation entre un jeu de données mesurées et des quantités d'intérêt (Delottier, 2023 ; Scheidt, 2015).

Des travaux antérieurs consacrés à la calibration de modèles hydrogéologiques fourniront un point de départ pour la prise en main de la démarche globale de modélisation et des outils numériques associés. Le stage consistera ensuite à comparer les deux approches en étudiant leur applicabilité et leur robustesse dans un contexte hydrogéologique spécifique à l'environnement des installations d'EDF. Il s'inscrit dans le cadre d'une thèse en cours sur le sujet.

Le code hydrogéologique [MODFLOW](#) ainsi que l'utilitaire [PEST++](#) seront utilisés dans le cadre de ce stage, leur connaissance est un plus mais n'est en aucun cas un prérequis nécessaire pour ce stage, un temps de formation étant prévu. La connaissance du langage Python est souhaitée.

Profil recherché

Etudiant(e) en Master 2 / Ecole d'ingénieur (en stage de fin d'études), avec un attrait pour les sujets de R&D

- Formation en modélisation numérique / mathématiques appliqués / analyse de données
- Connaissances en mécanique des fluides, intérêt pour la compréhension des processus physiques rattachés aux sciences de l'environnement, des connaissances en géostatistiques sont un plus
- Connaissance de l'algorithmique, bon niveau en programmation informatique (python)
- Capacité de recherche et de synthèse documentaires (en français et en anglais)

- Capacité à travailler de manière autonome et en équipe.

Durée et lieu du stage

Lieu : 6 quai Watier - 78401 Chatou (RER A : station Rueil Malmaison)

Durée : 5-6 mois (typiquement mars/avril-août/septembre 2026)

Indemnité de stage : précisée lors de l'entretien

Pour postuler

Raphaël Lamouroux : raphael.lamouroux@edf.fr

Marc Kham : marc.kham@edf.fr

Mathieu Couplet : mathieu.couplet@edf.fr

Zoé Petitjean : zoe.petitjean@edf.fr

Bibliographie

[White, 2018] J. T. White A model-independent iterative ensemble smoother for efficient history-matching and uncertainty quantification in very high dimensions, Environ. Model. Softw., Volume 109 (2018), pp. 191-201

[de Marsily, 1978] de Marsily, G. De l'identification des systèmes hydro-géologiques. *Thèse de doctorat d'Etat de l'Université Pierre et Marie Curie*, Paris. (1978)

[Scheidt, 2015] Scheidt, C., Renard, P. & Caers, J. Prediction-Focused Subsurface Modeling: Investigating the Need for Accuracy in Flow-Based Inverse Modeling. Math Geosci 47, 173–191 (2015).

[Delottier, 2023] Delottier, H., Doherty, J., and Brunner, P.: Data space inversion for efficient uncertainty quantification using an integrated surface and sub-surface hydrologic model, Geosci. Model Dev., 16, 4213–4231