

Offre de stage Année universitaire 2021-2022

1. Sujet

Que peut-on attendre de l'apprentissage statistique pour la modélisation hydrologique empirique des bassins versants ?

2. Type de stage

Stage de Master 2 ou de dernière année d'école d'ingénieur

3. Période – Durée

Six mois à partir de février-mars 2022

4. Organisme d'accueil et encadrant

Organisme d'accueil :

INRAE
UR Hydrosystèmes continentaux anthropisés (HYCAR)
Equipe Hydrologie des bassins versant (HYDRO)
1, rue Pierre-Gilles de Gennes
CS 10030
92761 Antony Cedex
Web :
<https://www.inrae.fr>
<https://www6.jouy.inrae.fr/hycar/>

Encadrant :

François Bourgin
Tel : 01 40 96 61 21
Email : francois.bourgin@inrae.fr
Web : <https://webgr.inrae.fr>

Les candidatures (CV + lettre de motivation) sont à adresser de préférence par mail à l'encadrant dont les coordonnées sont détaillées ci-dessus.

5. Indemnité de stage

Indemnité mensuelle d'environ 550 €

6. Profil du candidat

- Notions d'hydrologie, intérêt pour la modélisation et les sciences environnementales
- Bonnes connaissances en mathématiques appliquées et en apprentissage statistique
- Aisance en programmation (R ou Python)
- Aisance rédactionnelle et à l'oral

7. Poursuite éventuelle en thèse

Non, mais possibilités de projets de thèse sur d'autres sujets au sein de l'équipe d'accueil (sous réserve de financement disponible).

8. Description du sujet

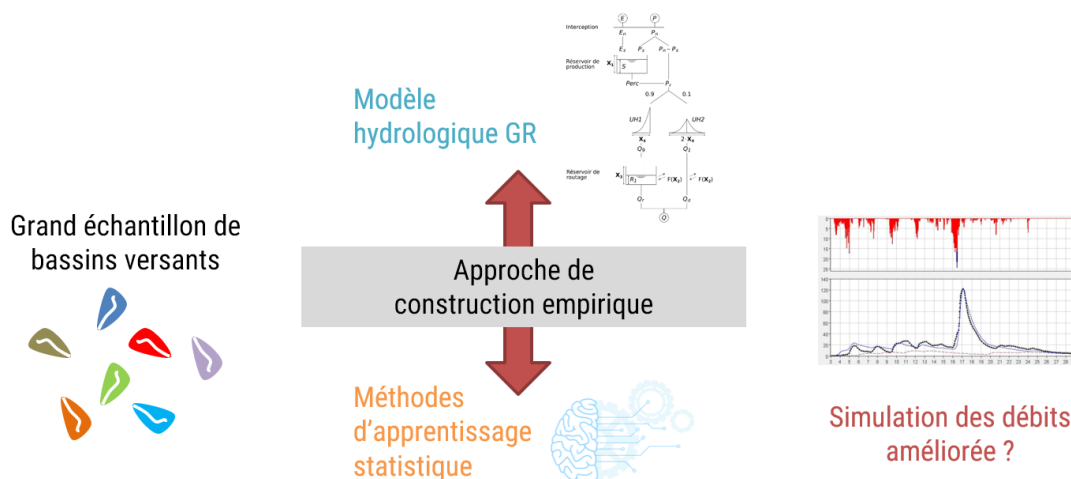
- **Contexte**

La modélisation hydrologique permet de simuler la réponse d'un bassin versant (notamment en débit dans un cours d'eau) à des conditions climatiques données (en particulier en termes de pluie et de température). La construction de ces modèles est souvent une démarche difficile du fait de la complexité naturelle des bassins versants, qui présentent des caractéristiques physiques très diversifiées et très variables dans l'espace et dans le temps. De ce fait, il existe un large spectre d'approches pour construire ces modèles, allant d'approches orientées sur la physique des processus à des approches beaucoup plus empiriques (basées sur les données) en passant par des démarches conceptuelles. Les modèles hydrologiques du Génie Rural (GR) développés à INRAE et très largement utilisés en conditions opérationnelles, par exemple pour la prévision des crues et des étiages, ont été construits depuis les années 1990 en suivant une démarche plutôt empirique, s'appuyant sur de larges bases de données regroupant de très nombreux bassins versants.

Les méthodes liées à l'apprentissage statistique, entendues comme des méthodes visant à rechercher une fonction prédictive en se basant sur les données, connaissent aujourd'hui un intérêt croissant et représentent une formidable opportunité pour le développement des géosciences et des sciences de l'eau (Reichstein et al., 2019; Shen, 2018). Des résultats encourageants ont notamment été obtenus sur de grands échantillons de bassins versants pour la simulation des débits, avec des réseaux de neurones récurrents (LSTM, GRU) qui permettent de représenter des dynamiques temporelles (Kratzert et al., 2019; Althoff et al., 2021).

- **Objectifs du stage**

Le stage a pour objectif d'explorer la question suivante : les méthodes issues de l'apprentissage statistique appliquées à de grands échantillons de bassins versants permettent-elles de compléter (selon une approche hybride), généraliser les approches empiriques existantes de modélisation pluie-débit (de type GR) et améliorer la caractérisation du fonctionnement des bassins versants et la généralité des approches de modélisation ?



- **Méthodologie / Etapes de travail**

Pour répondre à cette question, les travaux s'organiseront autour des étapes suivantes :

- Sélection d'un échantillon de bassins versants variés en France métropolitaine.
- Prise en main des bibliothèques spécialisées (torch, keras, tensorflow).
- Mise en place progressive de méthodes liées à l'apprentissage statistique (notamment réseaux de neurones) pour reproduire les différentes fonctions utilisées dans les modèles GR (interception, production, échange, routage).

- Utilisation de méthodes liées à l'apprentissage statistique pour généraliser le fonctionnement des modèles GR et identifier les marges de progression de la représentation globale de la transformation pluie-débit.
- Interprétation « hydrologique » des variables internes simulées.

Le calendrier suivant sera envisagé :

- Mois 1-3: Recherche bibliographique, prise en main des codes informatiques et préparation d'une base de données.
- Mois 3-5: Application progressive des méthodes liées à l'apprentissage statistique pour représenter le comportement global des bassins versants.
- Mois 6: Rédaction d'un mémoire.

- **Références bibliographiques et Internet**

Althoff, D., Rodrigues, L. N., and Silva, D. D. da: Addressing hydrological modeling in watersheds under land cover change with deep learning, *Adv. Water Resour.*, 154, 103965, <https://doi.org/10.1016/j.advwatres.2021.103965>, 2021.

Kratzert, F., Klotz, D., Herrnegger, M., Sampson, A. K., Hochreiter, S., and Nearing, G. S.: Toward Improved Predictions in Ungauged Basins: Exploiting the Power of Machine Learning, *Water Resour. Res.*, 55, 11344–11354, <https://doi.org/10.1029/2019WR026065>, 2019.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., and Prabhat: Deep learning and process understanding for data-driven Earth system science, *Nature*, 566, 195–204, <https://doi.org/10.1038/s41586-019-0912-1>, 2019.

Shen, C.: A Transdisciplinary Review of Deep Learning Research and Its Relevance for Water Resources Scientists, *Water Resour. Res.*, 54, 8558–8593, <https://doi.org/10.1029/2018WR022643>, 2018.