

Stage niveau M2/Ecole d'ingénieur  
**Combinaison de simulations climatiques par « machine learning »**  
LSCE-IPSL (CNRS/CEA) & Univ. De Lausanne

Les modèles de climat globaux (« Global Climate Models », GCM, en anglais) ou régionaux (« Regional Climate Models, RCMs) sont les outils de base pour projeter l'évolution du climat dans les décennies à venir (IPCC 2014). Les simulations (par exemple de température, de précipitation, etc.) issues de ces modèles sont couramment utilisées pour contraindre diverses études d'impacts, telles qu'en hydrologie ou agronomie entre autres, en contexte de changements climatiques.

Or, bien que leurs apports soient indéniables, ces modèles ont parfois des biais (c-à-d., des écarts à des observations) qui sont différents d'un modèle à un autre, d'une région d'étude à une autre, et d'une variable climatique à une autre (Christensen et al. 2008). Si de nombreuses approches de correction statistique de ces biais ont vu le jour ces dernières années (e.g. Piani et al. 2010; Vrac 2018), celles-ci présentent des limites (Ehret et al., 2012). D'autres méthodologies dites « ensemblistes » cherchent à combiner les simulations de différents modèles en les moyennant (Knutti et al. 2010). L'idée sous-jacente en contexte de changement climatique est que cette moyenne (bien que lissant le signal) ne gardera que la partie « robuste » des simulations d'évolution et que, d'une manière générale, la moyenne des simulations diminuera les biais.

Ce stage de recherche propose d'explorer une approche alternative basée sur une méthode de « machine learning » nommée le « graph cut » (Boykov et al., 2001, 2006). L'idée principale est tout d'abord d'identifier, pour chaque modèle, les régions présentant le minimum de biais par rapport à des données de référence. Puis, deux modèles (par exemple avec des régions de minimisation de biais distinctes) sont alors combinés, non pas en les moyennant, mais en « découpant » et « collant » les simulations suivant un chemin optimal de minimisation d'écarts entre les simulations. L'hypothèse à tester est que cette composition de simulations présente alors des biais minima sur l'ensemble des régions identifiées précédemment. Cette technique de « graph cut », déjà employée en contexte de synthèse d'images (Kwatra et al. 2003) et en géostatistique (Li et al. 2016), n'a jamais été testée (et donc jamais évaluée) dans un cadre de simulations climatiques.

Objectifs du stage :

1. Mettre en place et appliquer la méthodologie de graph cut. Ceci sera fait sur un ensemble de simulations climatiques provenant des exercices internationaux CMIP et EURO-CORDEX
  - a. Sur une période passée (dite « historique »)
  - b. Sur une période future (par ex., la fin du 21<sup>e</sup> siècle)
2. Analyser les résultats et les comparer à des approches ensemblistes par moyennes.

Compétences requises : Le stagiaire devra avoir un niveau M2 ou Ecole d'ingénieur avec des connaissances en statistiques et/ou en machine learning et/ou en climatologie. Une bonne capacité de programmation informatique (idéalement en langage R) est fortement souhaitée.

Période/durée : Idéalement, un minimum de 5 mois, à partir de mars 2019.

Lieux : Le stage se déroulera au Laboratoire des Sciences du Climat et de l'Environnement (LSCE, Gif-sur-Yvette) dans l'équipe « Extrêmes – Statistiques – Impacts – Régionalisation » (ESTIMR). Des échanges (mail/téléphone/Skype) réguliers avec Prof. Grégoire Mariéthoz dans l'équipe GAIA de l'Université de Lausanne auront également lieu et une visite de quelques jours sur place est envisageable.

Candidature : Le dossier de candidature devra être envoyé par mail à Mathieu Vrac ([mathieu.vrac@lsce.ipsl.fr](mailto:mathieu.vrac@lsce.ipsl.fr)), Grégoire Mariéthoz ([gregoire.mariethoz@unil.ch](mailto:gregoire.mariethoz@unil.ch)) et Soulivanh Thao ([soulivanh.thao@lsce.ipsl.fr](mailto:soulivanh.thao@lsce.ipsl.fr))

et devra comporter :

- une lettre de motivation décrivant l'adéquation des envies et du profil du candidat au sujet
- Un relevé de notes du M1 (voire du M2 si disponible)
- Un CV détaillé
- Les coordonnées (adresse/institut/téléphone/mail) d'une personne référence (ne pas envoyer de lettre de recommandations, la personne référence sera contactée si la candidature est retenue pour approfondissement).

#### Références:

- Boykov, Y., O. Veksler and R. Zabih (2001). "Fast approximate energy minimization via graph cuts." IEEE Transactions on Pattern Analysis and Machine Intelligence 23(11): 1222-1239.
- Boykov, Y. and G. Funka-Lea (2006). "Graph cuts and efficient N-D image segmentation." International Journal of Computer Vision 70(2): 109-131.
- Christensen, J. H., F. Boberg, O. B. Christensen and P. Lucas-Picher (2008). "On the need for bias correction of regional climate change projections of temperature and precipitation." Geophysical Research Letters 35(20).
- Ehret, U., E. Zehe, V. Wulfmeyer, K. Warrach-Sagi and J. Liebert (2012): HESS Opinions "Should we apply bias correction to global and regional climate model data?" Hydrol. Earth Syst. Sci. 16 (9), 3391-3404, 10.5194/hess-16-3391-2012.
- IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.
- Knutti, R., R. Furrer, C. Tebaldi, J. Cermak and G. A. Meehl (2010). "Challenges in combining projections from multiple climate models." Journal of Climate 23(10): 2739-2758.
- Kwatra, N., A. Schödl, I. Essa, G. Turk and A. Bobick (2003). "Graphcut textures: Image and video synthesis using graph cuts." ACM transactions on Graphics 22(3): 277-286.
- Li, X., G. Mariéthoz, D. Lu and N. Linde (2016). "Patch-based iterative conditional geostatistical simulation using graph cuts." Water Resources Research.
- Piani, C., J. O. Haerter and E. Coppola (2010). "Statistical bias correction for daily precipitation in regional climate models over Europe." Theoretical and Applied Climatology 99(1-2): 187-192.
- Vrac, M. (2018) Multivariate bias adjustment of high-dimensional climate simulations: the Rank Resampling for Distributions and Dependences (R2D2) Bias Correction. Hydrol. Earth Syst. Sci., 22, 3175-3196, <https://doi.org/10.5194/hess-22-3175-2018>