

Offre de Stage IPSL 2024

(Soutenu par le programme EUR IPSL-*Climate Graduate School*)

Titre du sujet de stage : Emulation du module de transfert radiatif par apprentissage automatique dans ORCHIDEE

Description du sujet (1 page maximum) :

Le modèle de surface continentale ORCHIDEE est l'un des composants d'ESM de l'IPSL. Son module de transfert radiatif peut calculer les lumières reflétées, absorbées et transmises à de multiples niveaux de la canopée. Or, ce module est aussi le plus coûteux en temps de calcul dans ORCHIDEE. C'est pourquoi nous sommes intéressés de développer un modèle statistique, qui pourrait représenter cette partie d'une manière précise et rapide.

Des méthodes d'apprentissage automatique sont indiqués pour cet objectif. Dans la première étude réalisée, on a trouvé qu'un émulateur avec Random Forest peut produire des résultats de bonne précision par rapport le modèle physique. Cela avec un gain de temps de ressources de calculs. Dans la continuité de cette étude, il est nécessaire de tester d'autres méthodes d'apprentissage pertinentes, par exemple ceux d'encodeur-décodeur, pour surmonter certaines limitations du RF. Cela serait aussi utile pour d'autre étude d'émulateur pour ORCHIDEE plus tard.

Récemment de nouveaux outils sont apparus pour réaliser l'intégration d'un modèle ML dans un modèle climatique. Ce genre d'outil établit le lien entre PYTHON et FORTRAN, et peut s'adapter tant aux supercalculateurs qu'à un ordinateur simple. Smartsim, développé par HPE, est l'un de ces outils. Dans notre première étude, SMARTSIM était installé sur un serveur de calcul au LSCE pour ORCHIDEE. L'implémentation préliminaire a été réalisée. Ce sera important de continuer et de finaliser cette implémentation avec ORCHIDEE.

La mission du stage peut inclure deux aspects, selon les compétences du candidat: 1) appliquer et tester des méthodes d'encodeur-decodeur pour émuler les données ORCHIDEE déjà préparés, 2) et/ou continuer l'implémentation et les tests des outils comme Smartsim, et les tester d'une manière plus complète sur un supercalculateur, afin d'obtenir une meilleure performance en temps et en précision, si possible.

Résumé en anglais (5 lignes) :

The transfer of radiations in canopy plants model is very time consuming in the current ORCHIDEE model. It is important to define an efficient emulator by using machine learning method which is to have good accuracy compared to the physical model. We have applied RF method and found good results. It will be important to continue this work in two aspects: 1) test other ML methods to

overcome some weakness of RF; 2) continue the integration of the machine learning model into ORCHIDEE by using recent tools such as smartsim on a supercomputer.

Responsable du stage (Nom/prénom/statut) :

Wang-Faivre, Xiaoni, Ingénieure de Recherche CNRS

Naveau, Philippe, Directeur de Recherche CNRS

Laboratoire concerné : LSCE

Adresse à laquelle a lieu le stage : LSCE

Equipe de recherche concernée (si pertinent) ou autre participant à l'encadrement du stage : Ordinateur et supercalculateur du LSCE

Niveau du stage (Licence, M1, M2, *internship*) : M2

Licence ou Master(s) où sera proposé le sujet :

Thème scientifique de l'IPSL concerné : SAMA

Durée du stage : 6 mois

Période : 02/2024       

Rémunération de l'ordre de 580 euros par mois.

Est-il prévu une thèse dans le prolongement du stage ?

Ce serait intéressant de proposer une thèse apres, selon la qualification du candidat et la confirmation du financement.