

**ECOLE DOCTORALE DES SCIENCES DE L'ENVIRONNEMENT
D'ILE DE France N° 129**

Proposition de sujet de thèse pour la rentrée 2020

Nom du Laboratoire d'accueil : LSCE/LMD

N° UMR : 8212 (LSCE)

Nom du Directeur du laboratoire : Philippe Bousquet

Adresse complète du laboratoire : CE Saclay, l'Orme des Merisiers, Bat 714, 91191 Gif-sur-Yvette Cedex

Nom de l'Equipe d'accueil et adresse si différente de celle du laboratoire : CLIM – modélisation du climat

Nom du Directeur de thèse **HDR** : KAGEYAMA Masa

Téléphone : 01 69 08 87 06

Mail : Masa.Kageyama@lsce.ipsl.fr

OU

Nom du co-directeur **HDR** : DUBOS Thomas

Téléphone : 169 335 143

Mail : dubos@lmd.polytechnique.fr

• Titre de la thèse en Français :

MACHINE LEARNING POUR LA MODÉLISATION DU CLIMAT: POTENTIEL ET ÉCUEILS

• Titre de la thèse en Anglais :

POTENTIAL AND PITFALLS OF USING MACHINE LEARNING IN CLIMATE MODELLING

• Résumé Sujet en Français (1 page maximum) :

L'anticipation du climat futur est un enjeu vital pour nos sociétés. La modélisation numérique du climat joue un rôle central dans cette anticipation. Les modèles de climat sont d'abord construits pour représenter le climat récent. Leur capacité éventuelle à représenter les climats passés froids ou chauds contribue fortement à asseoir leur crédibilité pour anticiper des climats futurs jamais rencontrés par l'humanité. La modélisation numérique du climat s'appuie sur les principes fondamentaux de la physique, exprimés directement pour les phénomènes d'échelle spatio-temporelle suffisante pour être résolus par le modèle, ou statistiquement et semi-empiriquement pour les phénomènes sous-maille. Le machine learning est une technique de modélisation statistique 'sans physique' qui a connu récemment de remarquables progrès. Des travaux pionniers [Gentine et al, 2018] suggèrent que certaines parties du modèle, notamment tout ou partie des modules sous-maille de la composante atmosphérique, pourraient reposer sur des techniques de machine learning et ainsi potentiellement gagner en efficacité numérique. La thèse explorera l'apport potentiel des techniques de machine learning pour la modélisation des climats présents passés et futurs. Le projet s'appuiera le modèle de climat de l'IPSL et particulièrement sur le jeu de paramétrisations physiques de sa composante atmosphérique, LMDZ. Les difficultés seront abordées progressivement dans le cadre de protocoles expérimentaux de plus en plus réalistes: atmosphère seule en configuration idéalisée dite "aquaplanète", atmosphère en configuration réalistes, avec couplage aux surfaces continentales, et si possible atmosphère couplée à un modèle d'océan.

• Résumé Sujet en Anglais (1 page maximum) :

Anticipating future climate change is vital for our societies. Numerical climate modelling plays an essential role for this anticipation. Climate models are first designed to represent recent climate. Showing their capacity to represent past cold or warm climates is also essential for increasing their credibility in representing climates different from present, never experienced by humankind. Climate modelling is based on fundamental physical principles, which are explicitly resolved in the case of phenomena of large-enough spatio-temporal scales, or represented statistically or semi-empirically for smaller scale phenomena. Machine learning is a 'physics-free' statistical modelling technique which has recently greatly improved in various domains. Pioneering work [Gentine et al, 2018] suggests that part of a climate model, in particular sub-grid-scale parameterizations, could be represented via results of analyses based on machine learning and hence help the models to gain computational efficiency. The objectives of this PhD is to explore the potential and pitfalls in building such new models partly based on machine learning techniques, for representing present, past and future climates. The project will make use of the LMDZ atmospheric general circulation model. The work will first examine idealized configurations (aquaplanets, no land) and then progress towards more realistic protocols, first including the coupling to the land surface, and if possible coupling to the ocean.

• Type de financement autre que ED 129, précisez si envisageable ou acquis (CNES, CEA, ADEME etc...) :
CEA NUMERICS (conditions d'éligibilité de type ETN européen)

• Encadrement :

. Liste des autres doctorants que vous encadrez ou co-encadrez au 1^{er} janvier 2020
(Nom, Université d'inscription, type de financement, date de soutenance envisagée)

Masa Kageyama : Marie Sicard, UVSQ ED129, CEA/convention services climatiques, soutenance dernier trimestre 2021

Thomas Dubos : Sakina Takache, (co-encadrement avec S. Mailler, LMD) ED IP Paris (ED626?), soutenance fin 2021

Cette thèse bénéficiera également de l'encadrement de Sébastien Fromang et Yann Meurdesoif (sur le modèle atmosphérique) et de l'équipe ESTIMR (Extrêmes : Statistiques, Impacts et Régionalisation) au LSCE (sur les méthodes de machine learning).