

**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 : LMD

N° UMR :8539

Nom du Directeur du laboratoire :Philippe Drobinski

Adresse complète du laboratoire : Sorbonne Université, Tour 45-55
4 Place Jussieu, 75252 PARIS Cedex 05

Nom de l'Equipe d'accueil et adresse si différente de celle du laboratoire :
EMC3 (Etude et Modélisation du Climat et du changement Climatique)

Nom du Directeur de thèse **HDR** :Frédéric Hourdin Nom du co-directeur de thèse
HDR :

Téléphone : 06 73 92 67 90 Téléphone :

Mail : hourdin@lmd.jussieu.fr

Mail :

OU

Nom du co-encadrant **non HDR** :

Téléphone :

Mail :

• **Titre de la thèse en Français :**

Approches statistiques pour l'automatisation de la calibration des paramètres des modèles climatiques et la quantification des incertitudes associées

• **Titre de la thèse en Anglais :**

Statistical approaches for automating the calibration of climate model parameters and the quantification of associated uncertainties

• Résumé Sujet en Français (1 page maximum) :

Les prévisions de l'amplitude de ce réchauffement et l'anticipation de ses conséquences reposent sur des simulations globales effectuées avec des modèles physico-numériques, améliorés au fil des années, mais qui restent — comme tout modèle — approchés. La quantification de l'incertitude liée à la nature approchée de ces modèles est une question fondamentale, mais difficile à aborder. Les grands programmes d'inter-comparaison CMIP (Coupled Model Intercomparison Project), qui rythment aujourd'hui les cycles de développement de ces modèles climatiques globaux de classe mondiale, ont cet objectif : évaluer la dispersion des projections climatiques liée aux choix de modélisation faits par les différentes équipes, en appliquant aux différents modèles des protocoles expérimentaux communs. Les analyses de ces simulations soulignent l'importance de la représentation des processus nuageux et convectifs, et ce à la fois pour la dispersion des projections et pour la bonne représentation du climat actuel. Ces processus sont représentés au travers de "paramétrisations", modèles simplifiés du comportement collectif d'un ensemble de panaches convectifs, cellules orageuses, cumulus, ou autre. Toutes ces paramétrisations incluent un certain nombre de paramètres ajustables plus ou moins connus. La calibration de ces paramètres libres — on parle souvent de tuning — est réalisée de façon à réduire la distance à certaines observations (flux radiatifs globaux, variations latitudinales, effets des nuages. . .), avec des objectifs qui varient d'un groupe de modélisation à l'autre [Hourdin et al., 2013]. L'augmentation de la puissance de calcul et l'arrivée à maturité de méthodes statistiques d'apprentissage automatique doit permettre de passer un cap sur ces questions.

La thèse portera sur la revisite du tuning et de la quantification de l'incertitude associée en utilisant les approches et outils issus de la communauté de la quantification des incertitudes — UQ pour Uncertainty Quantification —, acquises récemment dans l'équipe qui développe le modèle atmosphérique global LMDZ (composante atmosphérique du modèle de l'IPSL), et qui permettent de réaliser ce tuning de façon beaucoup plus automatique et systématique qu'auparavant. Les "émulateurs", "métamodèles" ou "modèles de substitution" (surrogate models) permettent d'explorer l'ensemble de l'espace à N dimensions sur la base de la réalisation de $\sim 10^N$ simulations climatiques seulement. Le développement des paramétrisations nuageuses est aujourd'hui souvent réalisé dans des versions uni-colonne des modèles de climat [Rio et al., 2010; Jam et al., 2013], dans lesquelles on tente de reproduire avec une intégration de quelques secondes une simulation explicite tri-dimensionnelle de la même scène nuageuse. Les nouveaux développements sont ensuite testés dans des intégrations tri-dimensionnelles globales dans lesquelles on impose les températures observées à la surface des océans comme condition aux limites. Dans ces simulations, plus contraintes que des simulations couplées océan-atmosphère, on évalue la simulation des nuages et on ajuste les flux radiatifs (en jouant sur l'impact sur ces flux des nuages) de façons à s'approcher des climatologies construites à partir des observations satellites. On vérifie enfin l'amélioration de ces modifications de contenu physique et d'ajustement des paramètres dans le modèle couplé océan-atmosphère. Le tuning des paramètres du modèle peut se faire à chaque étape de ce processus. On abordera en particulier la question du transfert, qui consiste notamment à essayer d'ajuster des configurations plus coûteuses du modèle sur la base de simulations plus légères : réglage des flux dans le modèle atmosphérique tri-dimensionnel à partir d'une série de cas de référence uni-colonne ; réglage des températures de l'océan dans le modèle couplé à partir du réglage des flux dans des simulations atmosphériques forcées par les températures océaniques [pour s'attaquer notamment aux classiques "biais chauds" des bords Est des océans tropicaux Hourdin et al., 2015] ; réglage d'une configuration haute résolution à partir de simulations plus grossières. Cette question du transfert est essentielle en pratique, du fait des temps de calcul des différentes configurations — secondes en mode unicolonne, un jour à quelques jours en 3D forcé, quelques semaines en couplé du fait des dérives lentes de l'océan —. Elle permet aussi de quantifier les sources d'incertitudes dans les modèles climatiques et de revisiter des questions fondamentales de la modélisation et de la compréhension du système climatique, classiquement abordées de façon beaucoup moins systématique en effectuant des simulations de sensibilité à un ou quelques paramètres clés.

• **Résumé Sujet en Anglais (1 page maximum) :**

Predictions of the magnitude of this warming and the anticipation of its consequences are based on global simulations carried out with models, which have been improved over the years, but which - like any model - remain approximate. Quantification of the uncertainty associated with the approximate nature of these models is a fundamental but difficult issue to address. The CMIP (Coupled Model Intercomparison Project) intercomparison programs, which now punctuate the development cycles of these world-class global climate models, have this objective: to assess the dispersion of climate projections linked to the modelling choices made by the different teams, by applying common experimental protocols to the different models. The analyses of these simulations highlight the importance of the representation of cloudy and convective processes, both for the dispersion of projections and for the good representation of the current climate. These processes are represented through "parameterizations", simplified models of the collective behavior of a set of convective plumes, storm cells, cumulus clouds, or other. All these parameterizations include a certain number of more or less known adjustable parameters. The calibration of these free parameters - often referred to as tuning - is carried out so as to reduce the distance to certain observations (global radiative fluxes, latitudinal variations, cloud effects, etc.), with objectives that vary from one modelling group to another [Hourdin et al., 2013]. The increase in computing power and the maturity of statistical methods for machine learning should make it possible to move forward on these issues.

The thesis will focus on revisiting tuning and the quantification of the associated uncertainty using approaches and tools from the uncertainty quantification community - UQ - recently acquired in the team developing the LMDZ global atmospheric model (atmospheric component of the IPSL model), and which make it possible to perform this tuning in a much more automatic and systematic way than before. The "emulators", "metamodels" or "surrogate models" (surrogate models) make it possible to explore the entire N-dimensional space on the basis of only ~10 N climate simulations. The development of cloud parameterizations is nowadays often carried out in one-column versions of climate models [Rio et al., 2010; Jam et al., 2013], in which an attempt has been made to reproduce an explicit three-dimensional simulation of the same cloud scene with an integration of a few seconds. The new developments are then tested in three-dimensional global integrations in which the observed ocean surface temperatures were imposed as boundary conditions. In these simulations, which are more constrained than coupled ocean-atmosphere simulations, the cloud simulation is evaluated and the radiative fluxes are adjusted (by playing on the impact on these cloud fluxes) to approximate the climatologies constructed from satellite observations. Finally, the improvement of these changes in physical content and parameter adjustment in the coupled ocean-atmosphere model is verified. Tuning of the model parameters can be done at each step of this process. In particular, the transfer issue will be addressed, which consists in particular of trying to adjust more expensive model configurations on the basis of lighter simulations: adjustment of fluxes in the three-dimensional atmospheric model based on a series of single-column reference cases; adjustment of ocean temperatures in the coupled model based on the adjustment of fluxes in atmospheric simulations forced by ocean temperatures [to tackle, in particular, the classic "warm bias" of the eastern edges of the tropical oceans Hourdin et al, 2015]; setting a high-resolution configuration from coarser simulations. This question of transfer is essential in practice, because of the computation times of the different configurations - seconds in single-column mode, a day to a few days in forced 3D, a few weeks in coupled mode due to slow ocean drift -. It also makes it possible to quantify the sources of uncertainty in climate models and to revisit fundamental issues in modelling and understanding the climate system, which are traditionally addressed much less systematically by carrying out sensitivity simulations to one or a few key parameters.

• **Type de financement autre que ED 129, précisez si envisageable ou acquis (CNES, CEA, ADEME etc...) :**

Aucun

• **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)

Aucun