

ECOLE DOCTORALE DES SCIENCES DE L'ENVIRONNEMENT
D'ILE DE France N° 129
Proposition de sujet de thèse pour la rentrée 2019

Nom du Laboratoire d'accueil : **Laboratoire de Météorologie Dynamique**

N° UMR : **8539**

Nom du Directeur du laboratoire : Philippe DROBINSKI

Adresse complète du laboratoire : LMD/Sorbonne Université/IPSL,
Tr 45-55, 3eme étage, B99, Jussieu,
57252 Paris Cedex 05

Nom de l'Equipe d'accueil et adresse si différente de celle du laboratoire : **EMC3**

Nom du Directeur de thèse **HDR** :

Frédéric Hourdin

Téléphone :01 44 27 21 67

Mail : hourdin@lmd.jussieu.fr

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

Balaji Venkatramani

Mail : balaji@princeton.edu

• **Titre de la thèse en Français :**

Nouvelles approches pour la calibration et la quantification des incertitudes en modélisation du climat

• **Titre de la thèse en Anglais :**

New approaches for calibration and uncertainty quantification in climate modeling

• **Type de financement autre que ED 129, précisez si envisageable ou acquis (CNES,**

CEA, ADEME etc...) : ED129

Envisagé : CDSN ENS-Cachan

• **Encadrement :**

. Liste des autres doctorants que vous encadrez ou co-encadrez au 1^{er} janvier 2019

(Nom, Université d'inscription, type de financement, date de soutenance envisagée)

Néant

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

Les projections du changement climatique 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'intercomparaison CMIP (Coupled Model Intercomparison Project), qui rythment aujourd'hui les cycles de développement des modèles climatiques globaux de « classe mondiale », ont cet objectif : évaluer la dispersion des projections climatiques liées 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, d'échelle généralement beaucoup plus petite que la maille des modèles globaux, sont représentés au travers de «paramétrisations », représentation sous forme de 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 ou de calibration — 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. Jusque-là, ce tuning était réalisé de façon relativement artisanale sur la base de quelques tests de sensibilité réalisés sur la version courante du modèle, afin d'essayer d'ajuster le modèle pour corriger certains défauts. 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), permettent aujourd'hui 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 — pour N paramètres à ajuster — sur la base de la réalisation de quelques centaines de simulations climatiques seulement. Avec ces techniques, on a l'espoir de rendre beaucoup plus efficace l'étape de tuning des modèles, et d'accélérer beaucoup leur processus de développement. En effet, jusque-là, ce processus était entravé par le fait que chaque modification significative nécessitait une phase de retuning, un peu « à l'aveugle », souvent pour compenser des phases d'overtuning précédentes — où l'on a poussé les paramètres vers des valeurs erronées pour compenser une erreur structurelle.

Le cadre général de cette thèse sera de revisiter la question du tuning et de l'estimation de l'incertitude paramétrique des modèles de climat grâce à ces approches UQ, ainsi qu'avec des approches plus nouvelles de machine learning (ML). On s'attaquera en particulier pendant cette thèse à l'incertitude sur l'ampleur du réchauffement global provenant de l'incertitude paramétrique et à des questions de transfert au travers de configurations du modèle : transfert entre l'échelle du processus en configuration unicolonne et des simulations 3D globales, entre différentes résolutions spatiales du modèle, entre modèles forcés par les températures océaniques et modèles couplés à un modèle d'océan. 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 —, mais elle est aussi une façon de quantifier les différentes 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 en effectuant quelques simulations de sensibilité à un ou quelques paramètres clés. On pourra se baser ici sur l'analyse statistique d'un ensemble beaucoup plus large de simulations qu'on ne le faisait dans cette approche classique.

• Résumé Sujet en Anglais (1 page maximum) :

Climate change projections are based on global simulations carried out with physical-digital models, improved over the years, but which remain - like any model - approximate. Quantifying the uncertainty of the approximate nature of these models is a fundamental question, but difficult to address.

The successive phases of the international Coupled Model Intercomparison Project (CMIP), which are now punctuating the development cycles of global "global" climate models, have the following objective: to assess the dispersion of climate projections related to modeling choices made by the different teams, applying to the different models common experimental protocols. The analyzes of these simulations emphasize the

importance of the representation of the cloudy and convective processes, both for the dispersion of the projections and for the good representation of the current climate. These processes, generally of a much smaller scale than the mesh of the global models, are represented through "parametrizations", representation in the form of simplified models of the collective behavior of a set of convective plumes, storm cells, cumulus, or other . All these parameterizations include adjustable parameters more or less known.

Calibration of these free parameters - often referred to as tuning or calibration - is done in such a way as to reduce the distance to certain observations (global radiative fluxes, latitudinal variations, cloud effects, etc.), with objectives that vary from one modeling group to another. Until then, this tuning was done in a relatively inefficient way based on some sensitivity tests performed on the current version of the model, in order to try to adjust the model to correct some defects. The approaches and tools from the uncertainty quantification community - UQ - recently acquired in the team that develops the global atmospheric model LMDZ (atmospheric component of the IPSL model), allow today to realize this tuning much more automatically and systematically than before. The "emulators", "metamodels" or "surrogate models" make it possible to explore the entire N-dimensional space - for N parameters to be adjusted - on the basis of the realization of only a few hundred simulations. With these techniques, we hope to make the tuning stage of the models much more efficient, and to speed up the development process a lot. Indeed, until then, this process was slowed down by the fact that each significant change required a retuning phase, a little "blind", often to compensate for previous overtuning phases - where we pushed the parameters to erroneous values to compensate for a structural error.

The general framework of this thesis will be to revisit the question of tuning and estimating the parametric uncertainty of climate models through these UQ approaches, as well as with the newer approaches of machine learning (ML). In particular, this thesis will address the uncertainty about the magnitude of global warming from parametric uncertainty and transfer issues through model configurations: transfer between the scale of the process in unicolon configuration and global 3D simulations, between different spatial resolutions of the model, between models forced by ocean temperatures and models coupled to an ocean model. This question of transfer is essential in practice, because of the calculation time of the different configurations - seconds in unicolon mode, a day to a few days in forced 3D, a couple of weeks coupled by the slow drifts of the ocean - but it is also a way to quantify the different sources of uncertainty in climate models and to revisit fundamental questions of modeling and understanding of the climate system, classically approached by performing some simulations of sensitivity to one or a few key parameters. The statistical analysis of a much larger set of simulations can be used here than was done in this classical approach.