

PROPOSITION DE STAGE EN COURS D'ETUDES

Référence : **DOTA-2026-12**
(à rappeler dans toute correspondance)

Lieu : Palaiseau

Département/Dir./Serv. : DOTA/MPSO

Tél. : 01 80 38 63 76

Responsable(s) du stage :
Sidonie Lefebvre et Patrick Chervet

Email : Sidonie.lefebvre@onera.fr

DESCRIPTION DU STAGE

Thématique(s) : réduction de dimension, machine learning, transfert radiatif

Type de stage : ☒ Fin d'études bac+5 ☒ Master 2 ☐ Bac+2 à bac+4 ☐ Autres

Intitulé : Accélération des simulations radiatives du code MATISSE par métamodélisation

Sujet :

Depuis plus de 20 ans, le département optique de l'ONERA développe le code de transfert radiatif MATISSE (*Modélisation Avancée de la Terre pour l'Imagerie et la Simulation des Scènes et de leur Environnement* - <https://matisse.onera.fr>), un outil de référence pour la simulation d'images de fonds naturels et de transmission et luminances en ligne de visée de l'ultraviolet à l'infrarouge. Ce simulateur est utilisé notamment pour la conception de senseurs optroniques et le développement d'algorithmes de détection avancés.

Cependant, les calculs réalisés par MATISSE sont coûteux en temps de calcul, en particulier pour des études multi-paramètres. Les approches issues de l'intelligence artificielle et de la métamodélisation permettent aujourd'hui d'envisager des approximations précises à faible coût.

Ainsi [Havemann2018] propose dans la méthode HT-FRTC de constituer une base de luminances spectrales associées à un calcul de transfert radiatif effectué sur une base de profils verticaux de températures, pressions, fractions massiques de gaz, couplés à des spectres d'émissivité de sol et des angles solaires. Une Analyse en Composante principale (ACP) permet de réduire la complexité des luminances spectrales en capturant leur variabilité sur une base réduite de composantes principales. Pour un nouveau profil non présent dans la base, un métamodèle de type processus Gaussien permet de remonter aux luminances spectrales à partir des scores de projection sur la base des composantes principales. Pour réduire encore les temps de calculs, le choix des longueurs d'onde sur lesquelles les calculs monochromatiques sont effectués est réalisé par clustering. Sur le domaine UV, [Aguila2019] propose de s'appuyer sur un ensemble de données représentatives de profils verticaux d'albedo et d'épaisseur optique et d'effectuer deux réductions de dimension par ACP séparément sur les données d'entrée, avec une correction à l'aide d'un modèle polynomial, et sur la dimension spectrale des sorties en luminance.

Nous souhaitons évaluer le potentiel et les limites de ces méthodes sur le domaine spectral couvert par MATISSE. Ainsi, à partir d'une base de données générée à l'aide de calculs MATISSE au début du stage en collaboration avec l'équipe MATISSE, le ou la stagiaire commencera par mettre en œuvre la stratégie la plus simple : deux réductions de dimension par ACP à la fois dans le domaine des entrées et celui des sorties. Une analyse de sensibilité sera effectuée, pour déterminer quelles variables d'entrée conduisent au plus de variabilité sur les sorties d'intérêt : et doivent donc être échantillonnées plus finement, et proposer un enrichissement de la base de données. On s'intéressera aussi à la méthode PLS (Partial Least Square) qui pourrait constituer une alternative intéressante à l'ACP, la PLS ayant l'avantage de prendre en compte la covariance entre les variables d'entrée et de sortie.

La deuxième étape portera sur la construction d'un métamodèle, de type processus Gaussien (GP - Gaussian Process) sur l'espace de dimension réduite, en s'inspirant des travaux de [Havemann2018]. L'apport de méthodes du type KPLS (Gaussian Process + PLS) et les Sparse Gaussian Process, qui optimisent le choix de points « fictifs » servant à résumer l'information contenue dans un ensemble de données de grande dimension pour l'entraînement du GP.

Les performances seront évaluées et comparées sur quelques cas d'intérêt, dont un cas en ciel clair et un avec nuages.

Le stage s'appuiera sur un environnement Python, avec les bibliothèques :

[SMT](#) – Surrogate Modeling Toolbox ;

[OpenTURNS](#) – outils d'optimisation et d'incertitude ;

GPFlow – Gaussian Process basé sur TensorFlow.

Références bibliographiques :

[Havemann2018] Stephan Havemann, Jean-Claude Thelen, Jonathan P. Taylor, R. Chawn Harlow, *The Havemann-Taylor Fast Radiative Transfer Code (HT-FRTC): A multipurpose code based on principal components*, Journal of Quantitative Spectroscopy and Radiative Transfer, **2018**, 220, 180-192, <https://doi.org/10.1016/j.jqsrt.2018.09.008>.

[Aguila2019] del Águila, A.; Efremenko, D.S.; Molina García, V.; Xu, J. *Analysis of Two Dimensionality Reduction Techniques for Fast Simulation of the Spectral Radiances in the Hartley-Huggins Band*. Atmosphere **2019**, 10, 142. <https://doi.org/10.3390/atmos10030142>.

Est-il possible d'envisager un travail en binôme ? Non

Méthodes à mettre en oeuvre :

☐ Recherche théorique

☐ Travail de synthèse

☒ Recherche appliquée

☒ Travail de documentation

☐ Recherche expérimentale

☐ Participation à une réalisation

Possibilité de prolongation en thèse : Non

Durée du stage :

Minimum : 4 mois

Maximum : 5 mois (6 mois sur dérogation uniquement)

Période souhaitée : début courant du premier semestre 2026

PROFIL DU STAGIAIRE

Connaissances et niveau requis :

Vous êtes issu(e) d'une formation en :

- Physique, mathématiques appliquées, modélisation numérique ;
- Data science ou machine learning ;
- Traitement du signal ou optronique, avec appétence pour l'IA.

Et vous avez :

- Des bases solides en Python et en calcul scientifique ;
- Un intérêt pour la modélisation physique et les données complexes ;
- Une curiosité pour les techniques de machine learning appliquées.

Ecoles ou établissements souhaités :

Master 2 ou école d'ingénieur