



Évaluation des couches de normalisation conditionnelles probabilistes pour des modèles de prévision en géoscience

La génération d'ensembles pour les modèles de prévision météorologique est un des éléments clefs pour pouvoir évaluer l'incertitude intrinsèque des prévisions. Cette incertitude, en grande partie due à la nature chaotique du système physique, fait que la prévision d'un état moyen ne suffit pas pour prédire des événements extrêmes.

Les modèles météorologiques numériques sont désormais challengés par des réseaux neuronaux, dont certains sont des modèles génératifs. Les méthodes génératives qui utilisent Flow Matching échantillonnent (théoriquement) la distribution des évolutions possibles conditionnées sur les états passés du système [genCast] et/ou d'une prévision de l'état moyen attendu [ArchesWeather]. Ces méthodes produisent des prévisions d'une grande qualité, ainsi qu'une incertitude, mais leur coût d'inférence est très grand, rendant difficile une utilisation opérationnelle.

D'autres approches qui utilisent des couches de normalisation conditionnelles probabilistes [FGN; AIFS, AIFsb] peuvent générer naturellement des ensembles et utilisent une métrique de prévision d'ensembles pour s'évaluer (CRPS - cumulative rank probability score). Cette métrique permet d'obtenir de bonnes performances mais n'offre pas de garantie théorique sur la distribution échantillonnée.

Étant donné la taille des modèles sur lesquels ces couches de normalisation conditionnelles ont été appliquées, une intercomparaison est difficile. Nous voulons, pour mieux comprendre leur impact en prévision, évaluer leurs performances sur des expériences jumelles.

Durant le stage, l'étudiant appliquera une approche de type flow matching ainsi que des réseaux à couches de normalisation conditionnelles probabilistes sur deux problèmes bien maîtrisés: un système 1-D de Lorenz95 ainsi qu'un modèle 2-D de circulation de surface océanique simplifié (shallow water) avec l'objectif de comparer la qualité des sorties obtenues, notamment en termes de prévision des événements extrêmes.

Références bibliographiques:

[genCast] Price, I. et al. (2023). *GenCast: Diffusion-based ensemble forecasting for medium-range weather*. arXiv preprint.

[FGN] Ferran, A. et al. (2025) *Skillful joint probabilistic weather forecasting from marginals*. arXiv preprint.

[ArchesWeather] Couairon et al. (2024). *ArchesWeather & ArchesWeatherGen: a deterministic and generative model for efficient ML weather forecasting*. arXiv preprint.

[AIFS] Lang, S., Alexe, M., Chantry, M. et al. (2024). *AIFS: a new ECMWF forecasting system*. ECMWF Newsletter.

[AIFsb] Lang, S. et al. (2025). *An update to ECMWF's machine-learned weather forecast model AIFS*. arXiv preprint.

Prérequis:

Bonne connaissance de deep learning, pytorch, méthodes de diffusion/flow matching.

Langues: français ou anglais.

Encadrement:

- Dominique Béréziat (dominique.bereziat@LIP6.fr) est enseignant chercheur à Sorbonne Université dans le laboratoire d'informatique LIP6 est spécialiste en assimilation des images couplé avec l'apprentissage profond
- Anastase Charantonis (anastase.charantonis@INRIA.fr) est titulaire d'une chair de professeur junior à Inria Paris / Sorbonne Université. Il travaille à l'interface entre le deep learning et les géosciences.

Localisation:

Le stage, d'une durée de 6 mois, aura lieu au LIP6.