



Offre de stage au Laboratoire Atmosphères Observations Spatiales (LATMOS)

Intitulé : Estimation de la distribution des gouttes de pluie comme un processus dynamique neuronal

Niveau : master 2 et 3ème année d'école d'ingénieur

Durée : 4/6 mois à partir du mois de mars

Lieu du stage : LATMOS, 11 Boulevard d'Alembert, 78280 Guyancourt (<http://www.latmos.ipsl.fr>)

Nom des encadrants : Aymeric Chazottes (aymeric.chazottes@latmos.ipsl.fr) et Cécile Mallet (cecile.mallet@latmos.ipsl.fr)

Indemnité de stage : Gratification autour de 550€/mois + indemnité de transport

Possibilité de poursuivre en thèse : Oui

Compétences requises :

- **Programmation :** Python/pytorch (C/C++ est un plus)
- **Aptitude à l'abstraction**
- **Bases techniques solides** en optimisation/edp/statistiques et/ou techniques d'apprentissage
- **Capacité d'abstraction et pragmatisme** (le contexte du stage est ancré dans le machine learning ainsi que sur des domaines allant des mathématiques appliquées à la microphysique de la pluie)

Description du sujet :

La pluie joue un rôle dans de nombreux domaines allant des télécommunications à l'hydrologie. La pluie y est, le plus souvent, considérée comme une grandeur continue mais elle peut aussi être vue comme un ensemble discret de gouttes (distribuées dans l'espace et le temps). L'interaction des ondes électromagnétiques avec l'atmosphère précipitante est influencée par la microphysique; la mesure radar ou l'affaiblissement du signal lors d'une liaison de télécommunication sera donc liée à la répartition des gouttes dans l'atmosphère. Une hydrologue voudra connaître des volumes d'eau mesurés au niveau du sol. Ainsi la pluie sera résumée à un processus continu caractérisé par un taux précipitant qui correspond au volume d'eau de l'ensemble des gouttes intégré par pas de temps.

La pluie est un phénomène complexe, "non stationnaire" et "extrême". Considérer la pluie comme un processus discret permet de la représenter plus finement mais aussi de mieux contrôler/expliciter/interpréter les paramètres du modèle retenu.

Il existe de nombreux appareils de mesure de la pluie, nous nous intéresserons plus spécifiquement aux mesures fournies par un disdromètre. Cet appareil mesure les gouttes une à une dans une zone très localisée de l'espace. Notre jeu de données de départ sera donc constitué de triplets correspondant à une taille (d_i), une vitesse (v_i) et un temps d'arrivée (t_i). Généralement ce genre de série temporelle est modélisé par un processus ponctuel temporel avec marque (m_i). La marque est l'information qui vient en plus du temps (t_i), ici la taille et/ou la vitesse de la goutte.

Il existe une littérature conséquente sur la modélisation stochastique de la pluie et de ses gouttes (De Michele and Ignaccolo, 2013). Il en découle une multitude de modèles dont certains incorporent un a priori sur la nature de la pluie (Lavernat and P. Golé, 1998) alors que d'autres sont purement stochastiques (Waymire and Gupta, 1981a)(Waymire and Gupta, 1981b)(Waymire and Gupta, 1981c). Ces différents modèles sont basés sur des hypothèses fortes qui conviennent mal à un phénomène aussi versatile que la pluie. Il semble important de proposer une modélisation à même de prendre en compte les connaissances a priori tout en restant fidèle à la mesure de la pluie. L'objectif du stage est de développer un modèle dont on puisse valider et comprendre les paramètres.

Récemment des modèles neuronaux adaptés aux processus ponctuels inhomogènes avec marques (Shchur et al., 2021) ont été proposés. Ces modèles ont une architecture leur permettant de prendre en compte les spécificités de la pluie tout en ajustant les propriétés propres aux séries temporelles apprises.

Le stage se focalisera sur ces modèles et sur leur capacité à apprendre des séries temporelles de gouttes de pluie (correspondant à plusieurs années de pluie en différentes zones du globe.)

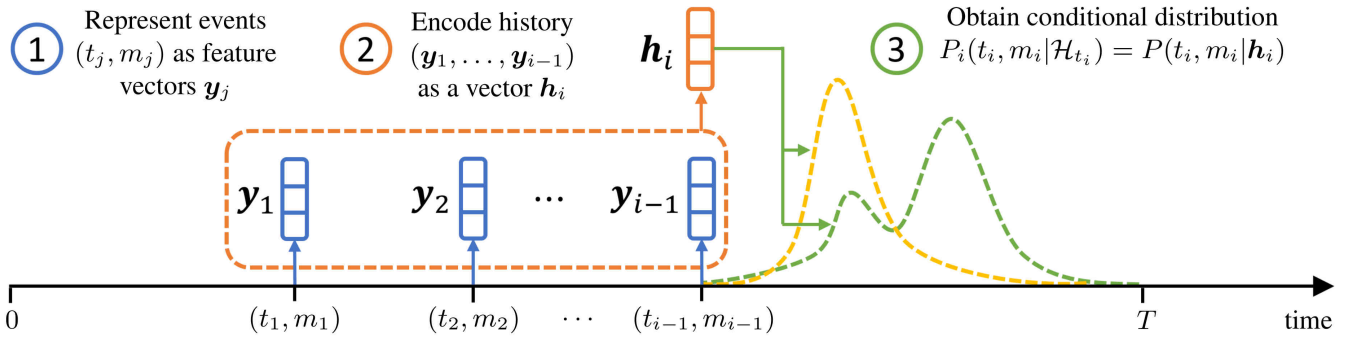


FIGURE 1 – Modélisation neuronale d’un processus ponctuel temporel (Shchur et al., 2021). Le modèle apprend des probabilités conditionnelles (3) caractérisant les données à partir d’une série temporelle d’événements (1) i.e. les couples (t_i, m_i) . Pour cela, il doit connaître l’historique des événement précédents ainsi que l’état courant du modèle (2).

Références

- C. De Michele and M. Ignaccolo. New perspectives on rainfall from a discrete view. *Hydrological Processes*, 2013.
- J. Lavergnat and P. P. Golé. A Stochastic Raindrop Time Distribution Model. *Journal of Applied Meteorology*, 37(8) : 805–818, 1998.
- Oleksandr Shchur, Ali Caner Türkmen, Tim Januschowski, and Stephan Günemann. Neural temporal point processes : A review. *arXiv preprint arXiv :2104.03528*, 2021.
- Ed Waymire and Vijay K. Gupta. The mathematical structure of rainfall representations : 1. a review of the stochastic rainfall models. *Water Resources Research*, 17(5) :1261–1272, 1981a. doi : 10.1029/WR017i005p01261. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/WR017i005p01261>.
- Ed Waymire and Vijay K. Gupta. The mathematical structure of rainfall representations : 2. a review of the theory of point processes. *Water Resources Research*, 17(5) :1273–1285, 1981b. doi : 10.1029/WR017i005p01273. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/WR017i005p01273>.
- Ed Waymire and Vijay K. Gupta. The mathematical structure of rainfall representations : 3. some applications of the point process theory to rainfall processes. *Water Resources Research*, 17(5) :1287–1294, 1981c. doi : 10.1029/WR017i005p01287. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/WR017i005p01287>.