

## Contrat doctoral Etablissement 2025/2028 Formulaire de candidature

|                                      |  |
|--------------------------------------|--|
| Titre de la thèse                    | <b>Découverte et étude assistées par apprentissage automatique des comportements des structures solaires</b> |
| Ecole Doctorale                      | ED548  |
| Laboratoire                          | LIS UMR 7020   |
| Discipline                           | Informatique   |
| Directeur(s) de Thèse & Encadrant(s) | Directrice : Adeline PAIEMENT (LIS)<br>Encadrants : Jean ABOUDARHAM & Pascal DEMOULIN (astronomes, LESIA)    |

### Description du sujet de recherche

(3 pages maximum - contexte scientifique, objectifs, mots clé, références)

#### Contexte, originalité :

Les événements d'activité solaire (AS) semblent être fortement associés à l'évolution des structures solaires (SS, par exemple régions actives (RA), filaments, taches solaires et trous coronaux), qui sont des objets de l'atmosphère solaire qui diffèrent de l'atmosphère « calme » et qui apparaissent, évoluent et disparaissent sur une période de quelques jours, voire quelques mois. Les mécanismes exacts de l'AS et les liens entre les événements d'AS et les SS sont encore mal compris.

Nous émettons l'hypothèse que les différentes SS peuvent avoir des comportements types, à la fois isolés et en interaction avec d'autres SS et avec les événements d'AS, y compris le cycle solaire de 11 ans. De tels modèles de comportement n'ont jamais été recherchés.

#### Objectifs :

Notre objectif est de les découvrir, s'ils existent, à l'aide de méthodes d'apprentissage automatique (machine learning ML). Ces comportements, ainsi que d'éventuelles anomalies observées dans les comportements, constitueraient de nouveaux indicateurs de haut niveau de l'AS, ouvrant ainsi de nouvelles perspectives pour en approfondir notre compréhension.

#### Pertinence par rapport à l'état de l'art :

Les mouvements et changements de forme ou de luminosité des SS ont été résumés et enregistrés dans des cartes synoptiques depuis le début du XXe siècle [1] par diverses organisations dont le partenaire du projet. Des changements dans la topologie magnétique ont été associés aux événements d'AS [2-10]. Certains travaux ont prédit des éruptions solaires en combinant les propriétés magnétiques des RA et leur luminosité [11] ou leur apparence [12,13]. Ces études prometteuses se sont concentrées sur un nombre limité d'indicateurs à la fois et n'ont pris en compte aucun comportement émergent de l'évolution temporelle des SS.

Une analyse à long terme des comportements des SS, en relation avec l'AS, est un nouveau concept qui offrira un point de vue inédit dans l'étude des mécanismes de l'AS. Alors que de nombreuses études visaient à reproduire des observations individuelles via des simulations numériques [14], très peu ont modélisé l'évolution générale des SS uniquement à partir de données observationnelles. L'évolution de certaines RA individuelles a été étudiée du point de vue de diverses propriétés [14]. Une étude de l'émergence et de la désintégration de 10 RA [15] a révélé une relation entre la durée de vie et la luminosité maximale. Une observation de la hauteur et de la décroissance magnétique de 12 filaments [16] a permis (visuellement) d'identifier deux comportements possiblement liés à des mécanismes d'éruption distincts. Ces études reposaient exclusivement sur une découverte visuelle, donc limitée à un certain nombre d'objets et de comportements à la fois.

Des algorithmes automatisés de détection et de suivi des SS ont été développés, certains par nos partenaires [17-19]. Les physiciens solaires les utilisent pour suivre les évolutions temporelles des emplacements et des propriétés physiques des SS. Cependant, ils ne proposent pas d'analyse de leur évolution temporelle et de leurs interactions. Nous proposons une étude systématique des comportements utilisant le ML et de grands ensembles de données. Nous examinerons la durée de vie des SS, les interactions possibles entre elles et des indicateurs globaux de leurs propriétés, notamment géométriques, optiques et magnétiques.

### Méthodes :

Étape 1 : La première étape de ce projet visera à fusionner ou co-exploiter des observations multimodales et hétérogènes du soleil qui en révèlent différentes facettes. Elles donnent accès à diverses propriétés des SS (optiques à différentes longueurs d'onde, magnétiques...). Une représentation réduite des propriétés des SS sera construite par apprentissage de manifold, en s'inspirant de nos travaux précédents [20,21], et en utilisant toutes les informations pertinentes, y compris la forme 3D des SS (produite par le projet ANR JCJC PRESAGE en cours), la luminosité et les propriétés magnétiques. A ce stade, les propriétés les plus pertinentes pour modéliser l'évolution temporelle des SS sont mal connues. Certaines propriétés magnétiques ont été identifiées [6,9] comme étant plus étroitement liées aux événements d'AS, et elles serviront de point de départ. Tout au long de la thèse, plusieurs combinaisons d'observations et de propriétés seront expérimentées pour obtenir un espace informatif où l'évolution temporelle des SS forme des trajectoires clairement identifiables. Le choix des propriétés sera en outre éclairé par une analyse de corrélation avec les événements d'AS, en utilisant des bases de données de CME et d'éruptions solaires historiques.

Étape 2 : Les trajectoires identifiées à l'étape 1 seront regroupées par similarité à l'aide d'algorithmes de clustering, chaque cluster identifiant un modèle de comportement donné. Des modèles statistiques pour chaque comportement seront construits à partir d'un grand nombre de SS individuelles. Au cours de cette étape, nous adapterons une méthode de modélisation statistique issue de nos travaux antérieurs [20,21]. Cette méthode a été initialement développée pour modéliser le mouvement de patients en orthopédie, dans l'optique de détecter des anomalies de mouvement se traduisant par une déviation par rapport au modèle de mouvement « normal ». Cette méthode se transcrit assez directement à la description des évolutions des SS et à la détection d'anomalies dans ces évolutions.

Étape 3 : L'entropie de transfert [22] sera calculée à partir des trajectoires de SS voisines pour tester si leurs interactions impactent les comportements. Sa mesure du flux d'information fournira des indices sur d'éventuelles relations de causalité entre deux comportements ou sur l'apparition d'une anomalie dans un comportement.

Étape 4 : Les comportements découverts aux étapes 2 (individuels) et 3 (en interaction) seront examinés à travers les propriétés physiques correspondantes et sur un grand nombre de SS observées, afin d'identifier des interprétations physiques pour ces modèles de comportement. Les partenaires en physique solaire de l'Observatoire de Paris (LESIA) apporteront leur expertise pour cette étape.

Étape 5 : Les comportements, ainsi que les anomalies ponctuelles dans les comportements, seront aussi examinés à travers les événements d'AS qui leur sont éventuellement associés. Ainsi, les modèles de comportement et les détections d'anomalies pourront servir à entraîner un réseau de neurones prédisant les événements d'AS, en lien avec la météorologie spatiale.

Au cours de cette thèse, nous pourrions nous concentrer sur un seul type de SS, à savoir les RA qui jouent un rôle central dans l'AS tout en ayant l'avantage de présenter des détections, des formes 3D et des évolutions temporelles relativement plus simples que d'autres SS comme les filaments. Si le temps le permet, l'étude pourra être étendue à d'autres SS dans le cadre du doctorat. Le projet ANR JCJC PRESAGE en cours pourra contribuer à ce travail de généralisation.

#### Retombées attendues :

Les résultats attendus de ce projet exploratoire sont : 1) obtenir des informations sur l'existence de comportements typiques ainsi que d'anomalies de comportement pour les SS, 2) modéliser ces comportements et en fournir une interprétation physique, 3) relier les comportements et leurs anomalies à des événements d'AS. Ces modèles de comportement ouvriront la perspective d'une meilleure prédiction des événements d'AS, avec des applications notamment pour la météorologie spatiale.

#### Mots clés :

Apprentissage automatique, physique solaire, imagerie, multimodalité, modélisation de comportements

#### Références :

- [1] *Cartes synoptiques de la chromosphère solaire. Annales de l'Observatoire de Paris, Section de Meudon*
- [2] Messerotti et al.. *Space Sci. Rev.* (2009) doi:10.1007/s11214-009-9574-x
- [3] Domijan et al.. *Sol. Phys.* (2019) arxiv:1812.02652
- [4] Park et al.. *Astrophys. J.* (2018) doi:10.3847/1538-4357/aaed40
- [5] Huang et al.. *Astrophys. J.* (2018) doi:10.3847/1538-4357/aaae00
- [6] Bobra & Couvidat. *Astrophys. J.* (2015) doi:10.1088/0004-637X/798/2/135
- [7] Florios et al.. *Sol. Phys.* (2018) arXiv:1801.05744
- [8] Boucheron et al.. *Astrophys. J.* (2015) doi:10.1088/0004-637X/812/1/51
- [9] Bobra & Ilonidis. *Astrophys. J.* (2016) doi:10.3847/0004-637X/821/2/127
- [10] Inceoglu et al.. *Astrophys. J.* (2018) doi:10.3847/1538-4357/aac81e
- [11] Nishizuka et al.. *Astrophys. J.* (2017) doi:10.3847/1538-4357/835/2/156
- [12] Jonas et al.. *Sol. Phys.* (2018) arxiv:1708.01323
- [13] Nishizuka et al.. *Astrophys. J.* (2018) doi:10.3847/1538-4357/aab9a7
- [14] van Driel-Gesztelyi & Green. *Living Rev. Sol. Phys.* (2015) doi:10.1007/lrsp-2015-1
- [15] Ugarte-Urra et al.. *Astrophys. J.* **815** (2015) doi:10.1088/0004-637X/815/2/90
- [16] Xing et al.. *Astrophys. J. Lett.* **857** (2018) arXiv:1804.01232
- [17] Fuller, **Abouadarham** & Bentley.. *Sol. Phys.* (2005) <https://previ.obspm.fr/articles/SolPhy-227-61.pdf>
- [18] Bonnin, **Abouadarham** et al.. *Sol. Phys.* (2013) arXiv:1202.2072
- [19] Zharkova et al. Solar Feature Catalogues In EGSO. *Sol. Phys.* (2005).
- [20] **Paiement** et al. in *BMVC* (2014) doi:10.5244/C.28.79

[21] Tao, **Paiement** et al.. *Comput. Vis. Image Underst.* (2016) doi:10.1016/j.cviu.2015.11.016

[22] Schreiber. *Phys. Rev. Lett.* (2000) arXiv:nlin/0001042

## **Encadrement et conditions matérielles pour le doctorant**

→ La thèse sera dirigée par Adeline Paiement, MdC HDR au LIS, et experte en science des données et apprentissage automatique pour la physique et l'astrophysique. Mme Paiement a déjà encadré une thèse appliquée à la physique solaire dans le passé. Elle encadre actuellement deux doctorants sur des sujets d'astrophysique, qui doivent tous les deux soutenir en 2026. Ces encadrements sont à hauteur de 50% chacun, car co-encadrés par des physiciens. Elle est aussi impliquée marginalement dans l'encadrement de trois thèses liées à la physique des océans, se terminant en 2026 et 2027, pour un taux d'encadrement total de 130%.

Ce projet s'inscrit dans un partenariat existant et très actif avec le laboratoire LESIA à l'Observatoire de Paris. Jean Aboudarham (astronome senior, expert en observation du soleil) et Pascal Démoulin (astronome, expert en observation du soleil et sur l'évolution des SS et AS) seront co-encadrants. Le LESIA a fourni de grandes quantités de données essentielles à cette thèse.

Cette thèse s'intégrera dans et complétera le projet ANR JCJC PRESAGE (PREdicting Solar Activity using machine learning on hétéroGÉneous data) (oct. 2021 – fév. 2027, partenariat avec Jean Aboudarham). Elle exploitera et adaptera les méthodes de ML en cours de développement dans le cadre de PRESAGE pour l'analyse des observations multimodales.

La thèse se déroulera dans le laboratoire LIS Toulon. Le doctorant aura accès aux ressources de calcul du LIS, et notamment à son cluster calcul haute performance. Le projet ANR JCJC PRESAGE de Mme Paiement a permis l'ajout d'un nœud de calcul dédié aux projets liés à la physique solaire, et le doctorant y aura donc un accès prioritaire.

Il existe une possibilité d'environnement de la thèse via le projet ANR PRESAGE, au moins pour la première année de thèse. Le post-doctorant de ce projet a quitté son poste plus tôt que prévu pour rentrer en Angleterre et y prendre un poste permanent. Il reste donc 35 643,59 € qui sera perdu si non utilisé pour une thèse. Cela permettra de demander une extension de manière à environner la thèse sur toute sa durée.

Un double diplôme ingénieur-M2 astrophysique existe à Strasbourg et peut fournir des candidats ayant un profil idéal (déjà un recrutement de doctorant en 2023).

## **Compétences attendues et personnes à contacter**

---

### **Compétences attendues :**

→ Master 2 ou diplôme d'ingénieur en informatique, science des données, ou physique, avec expérience forte en programmation exigée, et expérience souhaitable en apprentissage automatique et/ou physique solaire.

### **Personne(s) à contacter :**

→ Adeline Paiement, [adeline.paiement@univ-tln.fr](mailto:adeline.paiement@univ-tln.fr)