



SUJET DE STAGE MASTER 2 ANNEE 2020

TITRE DU PROJET : MATRIX

Des matrices organiques aux matrices numériques : Analyses multivariées pour l'identification des sources de pollution par Fluorescence 3D.

MOTS CLEFS : Fluorescence 3D, HAP, environnement, *chemometrics*, analyse de données.

ENCADREMENT :

- Coralie Biache
Tél : 03 72 74 52 29, e-mail : coralie.biache@univ-lorraine.fr
- Marc Offroy
Tél : 03 72 74 47 06, e-mail : marc.offroy@univ-lorraine.fr

LIEU DU STAGE : Université de Lorraine, CNRS, Laboratoires LIEC et GeoRessources, F-54000 Nancy, France.

CONTEXTE :

Parmi les 2,5 millions de sites contaminés identifiés en Europe [1], 332 500 sites sont impactés par des contaminations aux composés aromatiques polycycliques (CAP) et plus spécifiquement aux hydrocarbures aromatiques polycycliques (HAP). Ces HAP sont des molécules constituées d'au moins deux noyaux benzéniques, présentant des propriétés hydrophobes. Il en existe des centaines mais seules 16 d'entre elles sont listées comme polluants prioritaires par l'US-EPA du fait de leur propriétés toxiques, mutagènes et cancérigènes [2]. Les HAP sont issus de différentes sources, principalement du fait des activités anthropiques [3]. De par leur large diffusion dans l'environnement il est difficile de leur attribuer une origine précise. Dans certains cas, les teneurs et surtout les distributions des HAP peuvent renseigner sur l'origine de ces HAP [4] certaines sources de contaminations présentant une signature particulière. A l'heure actuelle, la caractérisation de ces sources de contaminations nécessite de réaliser une extraction par solvants organiques (classiquement dichlorométhane [5]) puis l'identification et le dosage des composés présents par des techniques chromatographiques (GC-MS, HPLC-DAD, LC-MS). Cependant, ces analyses sont longues et coûteuses ce qui limite l'acquisition et l'étude d'importantes séries de données.

La spectroscopie de fluorescence est une autre technique analytique permettant de concilier rapidité d'analyse et signature de source. En effet, de par leurs structures moléculaires, les HAP présentent la propriété de fluorescer. Initialement utilisée en mode Emission-Excitation synchrone, elle a été appliquée pour étudier les matrices organiques complexes comme les charbons et les bitumes routier [6] afin de déterminer le degré de condensation des structures aromatiques. Ce mode d'analyse, développé dans les années 90, permettait de réaliser des analyses sur des durées acceptables (15 minutes) alors que faire une analyse sur l'ensemble du spectre d'Emission-Excitation (Fluorescence 3D : excitation [nm] x émission [nm] x intensité [u.a.]) à cette époque prenait plusieurs heures. Cette technique a fortement évolué grâce aux perfectionnements de l'optique et de l'électronique permettant d'acquérir en un minimum de temps un très grand nombre de données (*e.g.* 10-20 minutes d'analyse par échantillon). La fluorescence 3D est aujourd'hui principalement utilisée pour l'analyse des échantillons aqueux (eaux de rivière et de mer) afin d'identifier les sources de matière organique dissoute d'origine naturelle [7]. En revanche, elle est peu utilisée pour l'analyse d'extraits organiques (solvants organiques) et dans un contexte de contamination anthropique. Pourtant, la grande quantité d'information contenue dans une matrice de fluorescence 3D doit permettre de compléter les informations structurales pour conduire à l'identification de source de HAP. Des outils de post-processing doivent donc être développés pour extraire les informations les plus pertinentes contenues dans ces matrices de fluorescence. Dans les études physico-chimiques, une discipline appelée *chemometric* a été développée pour justement acquérir par des outils mathématiques [8] une connaissance plus approfondie et une interprétation plus complète des échantillons analysés. Ces approches emploient l'analyse multivariée (MCR) et le *data mining* en chimie analytique.

DEVELOPPEMENT DU SUJET DE STAGE :

L'objectif du stage est donc d'appliquer et de développer des outils mathématiques (*chemometrics*), notamment les résolutions multivariées (e.g PARAFACE, MCR-ALRS, etc. [9,10]), pour exploiter de façon optimale les matrices de Fluorescence 3D issues d'extraits organiques d'échantillons contaminés aux HAP afin d'identifier des signatures caractéristiques de sources de contaminations.

- **Première étape : Partie expérimentale.**

Acquisition de spectres Fluo 3D en nombre suffisant pour permettre une utilisation robuste des analyses multivariées (250 spectres). Pour ce faire, six HAP ont été sélectionnés (*i.e.* naphtalène, fluorène, phénanthrène, fluoranthène, benzo[*a*]pyrène et benzo[*ghi*]pérylène) en fonction de leurs abondances dans les sols contaminés et leur nombre de cycles. Ces HAP seront analysés par Fluo 3D, purs et mélangés à des proportions et des concentrations variables, représentatives de celles retrouvées dans des extraits de sols. Cette première étape permettra d'optimiser les paramètres d'analyses, de mieux comprendre les interactions se déroulant au cours de la mesure (quenching) et d'identifier des différences de signaux en fonction de la composition des solutions.

- **Deuxième étape : Chemometric.**

La deuxième étape consistera en la réalisation d'analyses multivariées sur les données de ces échantillons représentatifs. En effet, ils permettront, d'une part, de déterminer quels sont les meilleurs prétraitements mathématiques adaptés à la technique expérimentale (*i.e.* corriger des phénomènes de fluorescence « parasites ») et d'autre part, de choisir quelle méthode *chemometrics* permet d'extraire le maximum de sources de nos mélanges de HAP (*i.e.* ACP, MCR-ALS ou encore PARAFAC, etc.).

- **Troisième étape : Test sur des échantillons réels.**

Des extraits organiques de matrices contaminées de différents origines (DNAPL, goudrons de houille, charbon, sols contaminés, sédiments de rivières...) seront analysés et viendront compléter la banque de données de spectres déjà obtenus. Ces échantillons ayant déjà été caractérisés au niveau moléculaire (HAP), cette dernière étape permettra ainsi de s'assurer de la validité de la méthode mathématique développée pour caractériser la signature d'échantillons complexes où de nombreuses interactions entre molécules existent.

RÉFÉRENCES :

- [1] European Environment Agency, (2012).
- [2] L.H. Keith, W.A. Telliard. Environ. Sci. Technol. 13 (1979) 416-423.
- [3] G. Witt. Mar. Pollut. Bull. 31 (1995) 237-248.
- [4] M.B. Yunker, R.W. Macdonald, R. Vingarzan, R.H. Mitchell, D. Goyette, S. Sylvestre. Org. Geochem. 33 (2002) 489-515.
- [5] Y. Li, R. Michels, L. Mansuy, S. Fleck, P. Faure. Fuel. 81 (2002) 747-755.
- [6] J. Kister, N. Pieri, R. Alvarez, M. A. Diez, J.J. Pis. Energy Fuels. 10 (1996) 948-957.
- [7] P.G. Coble. Mar. Chem. 51 (1996) 325-346.
- [8] M. Bosco, M.P. Callao and M.S. Larrechi. Talanta 72 (20017) 800-807.
- [9] S. Piqueras, L. Duponchel, R. Tauler, A. de Juan. Anal. Chim. Acta 705 (2011) 182-192.
- [10] S. Piqueras, L. Duponchel, M. Offroy, F. Jamme, R. Tauler, A. de Juan, Anal. Chem. 85 (2013) 6303-6311.

COMPETENCES REQUISES :

La candidate ou le candidat devra avoir des compétences solides en physico-chimie (notamment en spectroscopie de Fluorescence), et en mathématique (notamment statistiques, calcul matriciel). Elle ou il devra également avoir un goût particulier et une certaine autonomie en programmation (MATLAB, VBA, PYTHON, R ou tout autre langage).

En fonction de son intérêt pour le sujet et du déroulement du stage, l'étudiant.e retenu.e pour le stage aura la possibilité de poursuivre en thèse (ContratDoc) sur un sujet dans la continuité de celui du stage (sous réserve de réunir les critères de sélection imposés par l'école doctorale de rattachement de la thèse, SIRENa)