

# Plan de travail prévisionnel

## Application de l'intelligence artificielle à la chimie du sol africain

**Objectif :** Aider le développement du secteur agricole africain par le biais de l'intelligence artificielle appliqué à la chimie du sol.

### Introduction

Les progrès dans l'analyse rapide et à faible coût d'échantillons de sol, le géo-référencement d'échantillons de sol et une plus grande disponibilité des données de télédétection terrestre offrent de nouvelles opportunités pour prédire les propriétés fonctionnelles du sol à des emplacements non échantillonnés. Par contre, pour la mesure de la concentration des éléments du sol liées à la fertilité, actuellement est nécessaire utiliser des test coûteux en termes d'argent et de temps. Les techniques principales sont l'analyse d'éléments organiques CHNSO et la chromatographie en phase liquide pour l'analyse des éléments à faible concentration (P, K, Ca). L'analyse CHNS, analyse la plus courante en analyse élémentaire, est basée sur un procédé par combustion dont la température peut atteindre un maximum de 1150°C pour la mesure du carbone, de l'hydrogène et de l'azote. La proportion de ces éléments détermine le degré d'évolution de la matière organique, c'est-à-dire de son aptitude à se décomposer plus ou moins rapidement dans le sol et à fertiliser les cultures. La chromatographie ionique (ICP) est une méthode particulièrement bien adaptée à l'analyse des anions et cations majeurs des solutions du sol, par contre demande une procédure longue de préparation des échantillons et des instruments coûteux. Cette contraintes sont encore plus importantes pour des pays en voie de développement telles que ces de l'Afrique subsaharienne. Des analyses comme la mesure en fluorescence des rayons-X (XRF) et la spectroscopie infrarouge sont plus rapides et économiques mais ne donnent pas une mesure directe de la concentration des éléments importants comme le carbone, l'hydrogène et l'azote. D'ici l'intérêt de prédire ces variables par une modélisation mathématique et de trouver une corrélation entre les mesures obtenues par analyse conventionnelle et ces obtenues par des techniques plus rapides et économiques.

### Le jeux des données

5GB des données des mesures réalisées pour des échantillons de sol géo-référencés qui ont été collectés dans le cadre du projet Africa Soil Information Service (AfSIS).

- données de spectroscopie infrarouge
- analyses de carbone, azote et hydrogène
- analyses des éléments par XRF
- analyses des éléments par chromatographie ICP

disponibles dans le Amazon Web Services S3 bucket publique à l'adresse <https://registry.opendata.aws/afsis/>

### Les modèles existants et les démarches pour les améliorer

Une première modélisation était faite pour prédire la concentration du calcium mesurée par chromatographie liquide ICP (AWS AfSIS tutorial python code, no date)

mais aucune prédiction directe de la fertilité du sol était réalisée.

Ce modélisation peut être amélioré avec l'information supplémentaire des caractéristiques du sol, comme la profondeur et la présence éventuelle d'une culture agricole.

Les données de spectroscopie infrarouge ont été modélisées par des algorithmes classiques de régression (Ridge, Random Forest..) (*African-Soil-Analysis - FTIR spectroscopy*, no date) pour prédire des variables chimique du terrain comme le pH ou la concentration du phosphore. Une publication académique récente (Xu *et al.*, 2019) indique que la modélisation avec l'algorithme "Partial Least Square" (PLS) régression, utilisé depuis longtemps pour le développement des nouveaux médicaments (Deeb *et al.*, 2007), peut être appliqué avec succès à la chimie du sol. Dans ce travail on appliquera la PLS pour la prédiction d'une ou plus caractéristiques chimiques des échantillons de sol analysés à partir des respectives mesures de spectroscopie infrarouge. Les résultats seront comparés avec ceux obtenu par une réseaux de neurones, que représente une évolution de la modélisation PLS (Bjerrum, Glahder and Skov, 2017) (Padarian, Minasny and McBratney, 2019).

## Bibliographie

*African-Soil-Analysis - FTIR spectroscopy* (no date). Available at:

[https://github.com/pcohen89/African-Soil-Analysis/blob/master/PC soil analysis.py](https://github.com/pcohen89/African-Soil-Analysis/blob/master/PC%20soil%20analysis.py) (Accessed: 15 November 2020).

*AWS Afsis tutorial python code* (no date). Available at: <https://github.com/qedsoftware/afsis-soil-chem-tutorial/blob/master/afsis-soil-chem-tutorial.ipynb> (Accessed: 15 November 2020).

Bjerrum, E. J., Glahder, M. and Skov, T. (2017) 'Data Augmentation of Spectral Data for Convolutional Neural Network (CNN) Based Deep Chemometrics'. Available at: <http://arxiv.org/abs/1710.01927> (Accessed: 18 November 2020).

Deeb, O. *et al.* (2007) 'Effect of the electronic and physicochemical parameters on the carcinogenesis activity of some sulfa drugs using QSAR analysis based on genetic-MLR and genetic-PLS', *Chemosphere*, 67(11), pp. 2122–2130. doi: 10.1016/j.chemosphere.2006.12.098.

Padarian, J., Minasny, B. and McBratney, A. B. (2019) 'Using deep learning to predict soil properties from regional spectral data', *Geoderma Regional*, 16, p. e00198. doi: 10.1016/j.geodrs.2018.e00198.

Xu, X. *et al.* (2019) 'Detection of soil organic matter from laser-induced breakdown spectroscopy (LIBS) and mid-infrared spectroscopy (FTIR-ATR) coupled with multivariate techniques', *Geoderma*, 355. doi: 10.1016/j.geoderma.2019.113905.