Geophysical Research Abstracts
Vol. 17, EGU2015-8431-1, 2015
EGU General Assembly 2015

Massive processing of pyro-chromatogram mass spectra (py-GCMS) of soil samples using the PARAFAC2 algorithm

Lauric Cécillon (1), Katell Quénéa (2), Christelle Anquetil (2), and Pierre Barré (3)

(1) Irstea, UR EMGR Ecosystèmes Montagnards, Université de Grenoble, France (lauric.cecillon@irstea.fr), (2) UMR METIS, UPMC, UMR7619, Paris, France (katell.quenea@upmc.fr; christelle.anquetil@upmc.fr), (3) Laboratoire de Géologie de l'ENS, PSL Research University, UMR8538 CNRS, Paris, France (barre@geologie.ens.fr)

Due to its large heterogeneity at all scales (from soil core to the globe), several measurements are often mandatory to get a meaningful value of a measured soil property. A large number of measurements can therefore be needed to study a soil property whatever the scale of the study. Moreover, several soil investigation techniques produce large and complex datasets, such as pyrolysis-gas chromatography-mass spectrometry (Py-GC-MS) which produces complex 3-way data. In this context, straightforward methods designed to speed up data treatments are needed to deal with large datasets.

GC-MS pyrolysis (py-GCMS) is a powerful and frequently used tool to characterize soil organic matter (SOM). However, the treatment of the results of a py-GCMS analysis of soil sample is time consuming (number of peaks, co-elution, etc.) and the treatment of large data set of py-GCMS results is rather laborious. Moreover, peak position shifts and baseline drifts between analyses make the automation of GCMS programs data treatment difficult. These problems can be fixed using the Parallel Factor Analysis 2 (PARAFAC 2, Kiers et al., 1999; Bro et al., 1999). This algorithm has been applied frequently on chromatography data but has never been applied to analyses of SOM.

We developed a Matlab routine based on existing Matlab packages dedicated to the simultaneous treatment of dozens of pyro-chromatograms mass spectra. We applied this routine on 40 soil samples. The benefits and expected improvements of our method will be discussed in our poster.

References

Kiers et al. (1999) PARAFAC2 – PartI. A direct fitting algorithm for the PARAFAC2 model. Journal of Chemometrics, 13: 275-294.

Bro et al. (1999) PARAFAC2 – PartII. Modeling chromatographic data with retention time shifts. Journal of Chemometrics, 13: 295-309.