

## Improved chemometric methodologies for the assessment of soil carbon sequestration mechanisms

Marco A. Jiménez-González (1), Gonzalo Almendros (1), Ana M. Álvarez (2), and Francisco J. González-Vila (3)  
(1) National Museum of Natural Sciences, MNCN-CSIC, Madrid (Spain), (2) Autonomous University of Madrid (Spain), (3) Institute of Natural Resources and Agrobiological Sciences, IRNAS-CSIC, Sevilla (Spain)

The factors involved soil C sequestration, which is reflected in the highly variable content of organic matter in the soils, are not yet well defined. Therefore, their identification is crucial for understanding Earth's biogeochemical cycle and global change. The main objective of this work is to contribute to a better qualitative and quantitative assessment of the mechanisms of organic C sequestration in the soil, using omic approaches not requiring the detailed knowledge of the structure of the material under study. With this purpose, we have carried out a series of chemometric approaches on a set of widely differing soils (35 representative ecosystems).

In an exploratory phase, we used multivariate statistical models (e.g., multidimensional scaling, discriminant analysis with automatic backward variable selection...) to analyze arrays of more than 200 independent soil variables (physicochemical, spectroscopic, pyrolytic...) in order to select those factors (descriptors or proxies) that explain most of the total system variance (content and stability of the different C forms). These models showed that the factors determining the stabilization of organic material are greatly dependent on the soil type. In some cases, the molecular structure of organic matter seemed strongly correlated with their resilience, while in other soil types the organo-mineral interactions played a significant bearing on the accumulation of selectively preserved C forms.

In any case, it was clear that the factors driving the resilience of organic matter are manifold and not exclusive. Consequently, in a second stage, prediction models of the soil C content and their biodegradability (laboratory incubation experiments) were carried out by massive data processing by partial least squares (PLS) regression of data from Py-GC-MS and Py-MS. In some models, PLS was applied to a matrix of 150 independent variables corresponding to major pyrolysis compounds (peak areas) from the 35 samples of whole soils. The variable importance in the projection (VIP) histogram obtained from this treatment (total C and total mineralization coefficients as dependent variables) illustrated the contribution of the individual compounds to the total inertia of the models (e.g., carbohydrate-derived compounds, methoxyphenols, or specific alkylbenzenes were relevant in explaining the total quality and the biodegradation rates of the organic matter). Further simplified models consisting of direct PLS analysis of the debugged ion matrix calculated by averaging all ions (45 – 250 amu) in the whole chromatographic area in the 5–60 min range (here referred to as 'rebuilt MS spectra' or 'Py-MS spectra' when obtained connecting directly the pyrolyser to the MS detector through suitable interfaces) were carried out.

The above three approaches coincided in pointing out that C sequestration behave as an emergent soil property depending on the complexity of its progressive molecular levels. Most of the total variance is explained by specific assemblages of variables, strongly depending on the soil types. On the other hand, chemical biodiversity (e.g., Shannon indices or coefficients from multivariate data models) behaved as a common background in the prediction models including very different soil types. In fact, assessment of chemodiversity of the pyrolytic compound assemblages (or the Py-MS ion data) would represent a valid clue for the assessment of the extent to which the original biomass has been diagenetically reworked into chaotic structures with non-repeatable units, providing a useful proxy to forecast at least a portion of the total variance in the soil organic matter biodegradability.