

## Mineral–nutrient relationships in African soils assessed using cluster analysis of X-ray powder diffraction data

Benjamin Butler (1), Keith Shepherd (2), Mercy Nyambura (3), Andrew Sila (2), Javier Palarea-Albaladejo (4), Stephen Hillier (1,5)

(1) The James Hutton Institute, Craigiebuckler, Aberdeen AB15 8QH, UK, (2) World Agroforestry Centre (ICRAF), P.O. Box 30677-00100 GPO, Nairobi, Kenya, (3) International Fertiliser Development Centre, c/o icipe Campus, P.O. Box 30772-00100, Nairobi, Kenya, (4) Biomathematics and Statistics Scotland, JCMB, The King's Buildings, Edinburgh EH9 3FD, UK, (5) Department of Soil and Environment, Swedish University of Agricultural Sciences (SLU), SE-75007 Uppsala, Sweden

Minerals are the major component of most soils. Through direct inheritance from the parent material and subsequent alteration by chemical weathering, the soil mineral composition can be spatially diverse - reflecting the local geology, climate, and age of the soil. Each mineral present within the soil environment exhibits a characteristic chemical composition, crystal structure and solubility, thus soil mineralogy is a key component affecting the sources and availability of plant nutrients.

X-ray powder diffraction (XRPD) is consistently found to be the most accurate analytical technique for identifying and quantifying minerals in soil. As part of the Africa Soil Information Service (AfSIS) dataset, approximately 2000 soils from sixty 10 x 10 km 'Sentinel' sites across sub-Saharan have been analysed by XRPD, each with corresponding measurements of total (X-ray fluorescence analysis) and/or extractable (Mehlich-3 extraction) nutrient concentrations (B, Mg, K, Ca, Mn, Fe, Ni, Cu and Zn). This combination of soil mineral and nutrient data can be used to understand the sources of nutrients in soil and their availability to plants.

The identification and quantification of minerals from soil XRPD data is often found to be challenging and time consuming, which creates difficulty when handling large datasets such as that of AfSIS. However, if the diffractograms from XRPD are treated as reproducible digital mineral signatures, data-driven approaches such as cluster analysis can be applied that partition the data into a manageable number of mineralogically similar groups, facilitating easier interpretation of mineral–nutrient relationships in soil.

In this investigation, XRPD measurements from 940 AfSIS sub-soils covering all sixty Sentinel sites were classified into mineralogically similar groups by applying the k-means clustering algorithm to pre-treated data (aligned, square-root transformed and mean-centred). Use of nine clusters was found to provide optimum partitioning of the data. Subsequent quantitative analysis of the mean XRPD pattern from each of these nine clusters yielded approximations of the mineral composition that characterised each group. In doing so, clusters were found to represent soils that ranged from 9–97 % quartz, with significant mineralogical diversity with respect to plagioclase, K-feldspar, illitic clays, kaolinitic clays, expandable clays, carbonates and iron oxides. Crucially this mineralogical variation across the nine clusters is reflected in all total and exchangeable nutrient concentrations with exception only to extractable Zn. Here the contribution of minerals to nutrient availability and cycling in soils will be outlined using the geochemical data and mineralogically defined clusters, and the potential of cluster analysis to be applied to large soil XRPD datasets will be summarised.