



## **Assessment of the behavior of potentially toxic elements (PTEs) in soil from the Sarno River Basin through a compositional data analysis**

Thiombane Matar (1), De Vivo Benedetto (1), Stefano Albanese (1), Josep-Antoni Martín-Fernández (2), Annamaria Lima (1), and Angela Doherty (3)

(1) Department of Earth Sciences, Environment and Resources, University Federico II of Neapel (Italy) (thiombane.matar@unina.it), (2) Dept. d'Informàtica, Matemàtica Aplicada i Estadística, University of Girona (Spain) (josepantoni.martin@udg.edu), (3) Bay of Plenty Regional Council, 5 Quay St, Whakatane (New Zealand) (ilovevolcanoes@gmail.com)

“Assessment of the behavior of potentially toxic elements (PTEs) in soil from the Sarno River Basin through a compositional data analysis”

Thiombane Matar\*, De Vivo Benedetto, Albanese Stefano, Josep-Antoni Martín-Fernández, Lima Annamaria, Angela Doherty

\*Corresponding author: Email:thiombane.matar@unina.it

### Abstract

The Sarno River Basin (south-west Italy), nestled between the Somma–Vesuvius volcanic complex and the limestone formations of the Campania–Apennine Chain, is one of the most polluted river basins in Europe due to a high rate of industrialization and intensive agriculture. Water from the Sarno River, which is heavily contaminated by the discharge of human and industrial waste, is partially used for irrigation on the agricultural fields surrounding it.

We apply compositional data analysis on 319 samples collected during two field campaigns along the river course, and throughout the basin, to determine the level and potential origin (anthropogenic and/or geogenic) of the potentially toxic elements (PTEs).

The concentrations of 53 elements determined by ICP-MS, and were subsequently log-transformed. Using a clr-biplot and principal factor analysis, the variability and the correlations between a subset of extracted variables (26 elements) were identified. Using both normalized raw data and clr-transformed coordinates, factor association interpolated maps were generated to better visualize the distribution and potential sources of the PTEs in the Sarno Basin.

The underlying geology substrata appear to be associated with raised levels of Na, K, P, Rb, Ba, V, Co, B, Zr, and Li, due to the presence of pyroclastic rocks from Mt. Somma–Vesuvius. Similarly, elevated Pb, Zn, Cd, and Hg concentrations are most likely related to both geological and anthropogenic sources, the underlying volcanic rocks and contamination from fossil fuel combustion associated with urban centers.

Interpolated factors score maps and clr-biplot indicate a clear correlation between Ni and Cr in samples taken along the Sarno River, and Ca and Mg near the Solofra district. After considering nearby anthropogenic sources, the Ni and Cr are PTEs from the Solofra tannery industry, while Ca and Mg correlate to the underlying limestone-rich soils of the area.

This study shows the applicability of the compositional data analysis transformations, which relates perfectly relationships and dependencies between elements which can be lost when univariate and classical multivariate analyses are employed on normal data.

Keywords: Sarno basin, PTEs, compositional data analysis, centered-log Transformation (clr), Biplot, Factor analysis, ArcGIS