Goldschmidt 2025 Abstract
https://doi.org/10.7185/gold2025.31918

GeoChemNet: An Interactive Tool for Visualizing and Interpreting Outliers in Geochemical Data Using Networks

TIMOTHY CHEE CHENG LUI[1], ANIRUDH PRABHU[2], ANNA BIDGOOD[3], SHAUNNA MORRISON[2] AND JEF CAERS[1]

[1]Stanford University
[2]Carnegie Institution for Science
[3]Kobold Metals

We introduce our open-access tool, *GeoChemNet*, that visualizes and analyzes geochemical data while also plotting geospatial information. Its key feature is the use of networks to uncover relationships and groupings between elements and samples. The use of force-directed graphs to represent geochemical datasets allows for a non-linear investigation of geochemistry. And the tool is not dimensionally constrained, allowing for better representation of the high multidimensionality that exists in geochemical datasets. These, plus the ability to concurrently plot results on a map, make *GeoChemNet* a great tool for domain experts, geochemists, and data scientists to visualize, discuss, and understand their datasets.

Traditional visualizations like scatter plots and pairplots are limited to 2 or 3 dimensional elemental relationships and cannot easily handle the ~50 element datasets modern analytical methods obtain. Data science methods like principal component analysis (PCA) rely on linear relationships and focus on variance explanation at the cost of exploring outlying data. *GeoChemNet* tackles these issues from a different perspective, leading to a non-linear, dimensionally unbounded understanding of the geochemical dataset, while also creating interactive visualizations that highlight spatial relations. Hence, *GeoChemNet* can be another tool in the toolbelt to analyze geochemical data.

We show how *GeoChemNet* works, highlighting how figures are created, how to read the plots, and how to use the interactive widgets for live updating and interrogating of the dataset. We compare how well different techniques (pairplot, PCA, *GeoChemNet*) can represent an artificial dataset with non-linear structures. As well as how *GeoChemNet* can be used on real datasets in synergy with other data science methods such as clustering in a case study using Zambian soil geochemistry.