Paper Number: 3650

Determining elemental associations by multivariate analysis of regional-scale geochemical data

Grunsky, E[1], McKinley J.M.[2] and Mueller U.A.[3]

[1] Department of Earth and Environmental Sciences, University of Waterloo, Ontario, Canada, N2L 3G1
[2]School of Geography, Archaeology and Palaeoecology, Queen's University Belfast, BT7 1NN, UK, j.mckinley@qub.ac.uk
[3] School of Science, Edith Cowan University, Joondalup, Western Australia, WA 6027

Northern Ireland has a varied mineral prospecting history dating from the 18th to early 20th century. Previous geochemical surveys have revealed anomalies of lead, barium and silver together with galena and barite in heavy mineral samples. However, areas of thick glacial till, alluvium and peat cover make it difficult to identify subtle geochemical anomalies. A regional scale soil geochemical data obtained as part of the Tellus Project by the Geological Survey Northern Ireland (GSNI) provides a multi-elemental soil geochemical data set. The soil geochemical survey potentially reflects changes in underlying bedrock, soil type, anthropogenic contamination and anomalies associated with precious and base metal mineralisation. This paper uses a compositional multivariate approach to analyse total concentration data comprising XRF analyses of 6862 rural soil sample sites sampled at 20cm depths and collected on a non-aligned grid at one site per 2 km^2. Using published detection limits, censored data were imputed and the elements S, Ag, In, Sb, Ta, Te were excluded from further analysis. Using these imputed values for 46 elements, each soil sample site was assigned to the regional geology map provided by GSNI using the dominant lithology for the map polygon. This provided 30 lithologies with a suitable number of sample sites. The geochemical data were transformed using centered log ratios (clr). Compositional biplots indicated that element associations linked to underlying lithologies with 72% of the variation determined by the first four principal components (PCs). An analysis of variance showed that 10 PCs were necessary to classify the soil geochemical data. A robust and non-robust approach was used to undertake an allocation procedure (based on the 10 dominant PCs). To allow a comparison, minimum/maximum autocorrelation (MAF) factors based on the clr data, according to the formula given in Bandarian and Mueller (2008), were then computed from the 46 variables. This produced typicalities, in essence, an allocation as a measure of the Mahalanobis distance (MD) of a given observation to the centroid of each class. Based on the F distribution and a chosen threshold of 0.95, an observation is allocated to a specific class if the MD is less than the critical value of the F distribution for the given degrees of freedom. The associated probability (typicality) is determined according to Garrett (1989; 1991). If the MD exceeds the critical value then there is a typicality of zero. PCA typicalities (robust) and posterior probabilities were plotted and compared with MAF typicalities (robust) and posterior probabilities.

Classification results of the posterior probabilities and typicalities showed a prediction accuracy of 55.43%, using 10 principal components, compared to 58.73% using 8 MA factors. Multi-dimensional scaling (MDS) was used to create a 2D plot of the mean MDS scores for each of the 30 groups. These

plots summarised the best discrimination between the groups without introducing a classification bias. Dominant geological domains such as Palaeogene flood basalts that cover the north eastern region of Northern Ireland are easily identified in the mean MDS scores. Other natural groupings of limestone, argillaceous rocks and mudstones are also recognisable, indicating Carboniferous limestone, interbedded shales and mudstones.

References:

[1] Bandarian E and Mueller U (2008) Geostats: 1173-1178

[2] Garrett (1989) Geological Survey of Canada 89:309-318

[3] Garrett (1991) Newsletter of the Association of Exploration Geochemists 81:9-14