Multivariate data analysis and machine learning in mineral deposits of the Pacific Rim: leading to robust geological and geometallurgical domains

M.F. Gazley[1]*, N. Caciagli*[2] *and S.B. Hood*[3]

1.
Principal Geochemist, RSC Mining and Mineral Exploration, Wellington, New Zealand. MAusIMM. Email: m.gazley@rscmme.com

2.
Senior Manager Resource Geochemistry, Barrick Gold Corporation, Toronto, Ontario, Canada. Email: ncaciagli@barrick.com

3.
PhD Student, University of Tasmania – TMVC (The Mining Value Chain), Hobart, Tasmania, Australia. Email: shawn.hood@utas.edu.au

ABSTRACT

Through utilisation of multivariate data analysis combined with machine learning techniques it is possible to identify geochemical domains considering many, and potentially all available, elements in a dataset. Resulting geochemical domains can reflect mineralogy to a degree that might be difficult or impossible to produce using geological mapping, visual logging, or considering geochemical data in conventional ways. Machine learning-assisted domains can provide a robust link between mineralogy and metallurgical variables such as flotation or leachability, recovery or concentration of deleterious elements, as well as comminution properties. Unfortunately, there are very few published studies that examine geochemical domaining of mineral deposits. We address that issue here using datasets from mineral deposits around the Pacific Rim, including Cu-porphyry, orogenic Au, epithermal Au, Carlin-style mineralisation and granodiorite-hosted Cu-Ag-Au. A flexible, four-step workflow, adapted to local deposit geology and local-expert knowledge, broadly has the following steps: (1) imputing missing values (those less/greater than the detection limit); (2) using a log ratio to convert the numbers to real-number (Euclidean) space; (3) a principal components analysis to decrease the dimensionality of the dataset and minimise the effects of autocorrelating variables; and (4) unsupervised classification algorithms applied to processed data to produce groups that maximise similarity of samples with groups and variance between groups. In all deposits examined (both around the Pacific Rim and elsewhere globally), the data points within these geochemical clusters exist in spatially coincident groups; the affinity for clusters to be dominated by specific logged lithologies or mineralisation types suggests that the clusters have meaning in terms of geological process in three dimensions, and thus ultimately could be interpreted to be geometallurgical domains. We regard this kind of domaining as the starting point of ore-body characterisation for near-mine exploration, mine planning, geometallurgy, ore dispatch and waste segregation or closure planning.