第32回日本情報地質学会講演会 講演要旨集 027-028頁 2021年

Machine Learning Application to Delineating Metal-rich Veins

Vitor Ribeiro de Sá* and Katsuaki Koike*

*Graduate School of Engineering, Kyoto University, Katsura C1–2, Kyoto 615–8540, Japan.

E–mail: ribeirodesa.vitor.7c@kyoto-u.ac.jp; koike.katsuaki.5x@kyoto–u.ac.jp

Key words: Hierarchical clustering, Epithermal deposit, Au–Ag mineralization, Geologic model

1. Introduction

Geological/geostatistical domain definition plays a pivotal role in mineral resource modeling. Machine learning (ML) techniques, such as cluster analysis, has been gained notoriety in earth science for their versality to furnish tools to handle sparse information in wide areas. ML and field studies can act in concert to improve the understanding of geological phenomenon.

This study aims to apply hierarchical clustering analysis (HCA) to a real geochemical data set from a low-sulfidation gold deposit. Such type of deposits account for significant occurrences of Au and Ag, besides base metals (i.e. Cu–Pb–Zn) (Pirajno, 2009). HCA can yield categories that assemble data with similar characteristics, chemical-wise. As it turns out, veins and veinlets, where higher Au–Ag concentrations are hosted in this deposit, can be identified. Highlighting these features aids effective field survey and provide elements to carry out geostatistical models in order to delineate target areas for future exploration in the study area.

The next sections introduce the study area and set out the methodology and its perks to the current study case. Besides, the results are presented and discussed. Although the methodology is applied to a specific environment of this study, HCA can be widely used for a myriad of geological contexts.

2. Study Area and Dataset

The Sirawai deposit, Mindanao, Philippines was selected as a case study area. The target zone is set in a 200 m × 210 m hilly field, elevation 280–330 meters above sea level (masl), where 56 drill sites (black dots) collected samples to analyze Au, Ag, Cu, Pb and Zn grades (Fig. 1a). This field is marked for several subparallel sulfide-quartz veins of 300–500 m length, which their strike 20–30° NW and dip 60–70° SW with the 1.0–5.1 m width and the average 2.5 m can be observed. Such structures host the greatest Au and Ag grades in the area.

The proposed methodology uses Au and Ag grade measurements, given in g/t, in varied depths. Fig. 1b shows the location of 56 drill sites and the length of the drill holes, 46 vertical holes (5 – 50 m depth) and ten inclined holes (70 m length) at dip 60°. The samples are separated by 1 m interval. Both precious metals are correlated in this deposit (linear correlation coefficient ρ = 0.69). Au grades varies from nearly zero to 183.3 g/t (average 0.46). Ag presents wider variation, from 0.01 to 1141 g/t (mean 12.75 g/t). According to the highest values of log (Au), the main occurrences of Au–Ag mineralized bodies are situated in the middle study area.

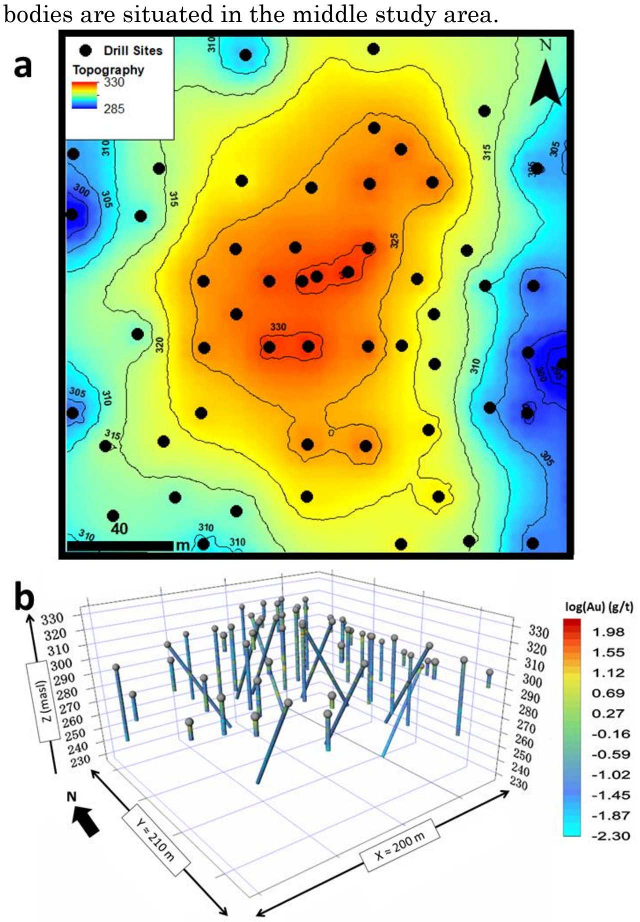

Figure 1. Dimension of study area and its spatial distribution of drill sites. (a) Detailed topography of the study area with 56 drill sites (black dots), and (b) 3D representation of the drill sites and holes lengths and log (Au) concentrations.

3. Methods

3.1. General structure of methodology

The proposed methodology assembles Au–Ag available data into relatively homogeneous groups or clusters. The members of a cluster are at once alike and at same time unlike members of other groups. To attain this goal, firstly, Au–Ag data are log-transformed, and their correlation is statistically verified. In sequence, HCA is applied to assign samples that are Rich, Intermediate and Poor in Au and Ag grades. The following subsection briefly describe the main HCA characteristics and the cluster number selection criterion.

3.2. Hierarchical Clustering Analysis (HCA)

HCA is an unsupervised machine learning technique which joins the most similar observations, then successively connects the next most similar observations to these. The similarities of all pairs of observations are calculated in a square matrix. Those pairs that share more similarities are merged, and the matrix is recomputed. Such procedures are run by averaging the similarities that the combined observations have with other observations. This process iterates until the similarity matrix is reduced to 2×2 (Davis, 2012). The "elbow" method is used to select the optimal number of clusters. Such method consists in graphically observing the magnitude of inertia considering number of clusters. Inertia is the sum of squared distances of samples to their closest cluster centers.

4. Results and Discussion

The point after which the inertia start decreasing in a linear fashion is deemed the elbow point (Fig. 2a). Thus, for the given data, we conclude that the optimal number of clusters is three: Poor (purple), Intermediate (yellow) and Rich (green) (Fig. 2b).

Poor cluster is impoverished in both precious metals and represents roughly 31% of the samples. Gold grades range from 0.001 to 0.2 g/t (mean 0.02 g/t) while silver values vary from 0.01 to 5.0 g/t (average 1.85 g/t). Cluster considered as Intermediate occurrence of these metals has minimum grades approximately 0.1 g/t and its maximum grade is 0.8 g/t, mean 0.11. Intermediate grades of Ag ranges from 0.9 to 74 g/t, mean 8.07 g/t. These samples are the most representative among the data, approximately 57%. The most enriched samples in Au and Ag are encompassed by cluster Rich and accounts for nearly 11% of the total data. Au minimum grade is 0.15 g/t and reaches 183 g/t, averaging 3.54 g/t. Ag ranges from 5.9 to 1141 g/t, mean 67.65 g/t.

Cluster Rich is more conspicuous in shallow parts over the area. However, in the central region of the study area its occurrence takes place in deeper parts, which suggest the nested geological structures (i.e. veins and veinlets). Impoverished samples (cluster Poor) are relegated to deeper portions and more predominant in peripheral zones of the study area. These extreme clusters have no direct contact in the entire area. Cluster Intermediate plays as a buffer zone between Rich and Poor clusters.

We interpret the cluster Rich as possible evidence of the sulfide quartz veins in central area and deeper zones. When this cluster takes place in shallow zones and adjacent regions of the study area, it can be soil enriched by weathering. Cluster Intermediate is regarded as hydrothermal alteration halos when it is situated near Au–Ag rich veins whereas regions or zones slightly affected by hydrothermal activity, otherwise. On the other hand, Cluster Poor stands for the regions or zones where hydrothermal did not take place or its influence was weak.

5. Conclusion

The HCA application is demonstrated to be effective for classifying the samples of precious metals (gold and silver) grades. In addition, this methodology aids to gain insight into vein continuities in subsurface. HCA can contribute to (a) yield categorical data and (b) support the construction of geological models using geostatistical methods, e.g. sequential indicator and pluri-Gaussian simulations as de Sá *et al.* (2021).

Acknowledgement: We sincerely thank Dr. Hiroshi Takahashi of Itochu Mineral Resources Development Corp. for providing the geological and metal analysis data of the Sirawai mine.

References

Davis, J. C. (2012) Statistics and data analysis in geology. Wiley, New York, 646 p.

de Sá, V. R., Koike, K., Goto, T.-N. *et al.* (2021) 3D geostatistical modeling of metal contents and lithofacies for mineralization mechanism determination of a seafloor hydrothermal deposit in the middle Okinawa Trough, Izena Hole. *Ore Geology Reviews*, https://doi.org/10.1016/j.oregeorev.2021.104194

Pirajno, F. (2009). Hydrothermal processes and mineral systems, Springer, Berlin, 1250 p.

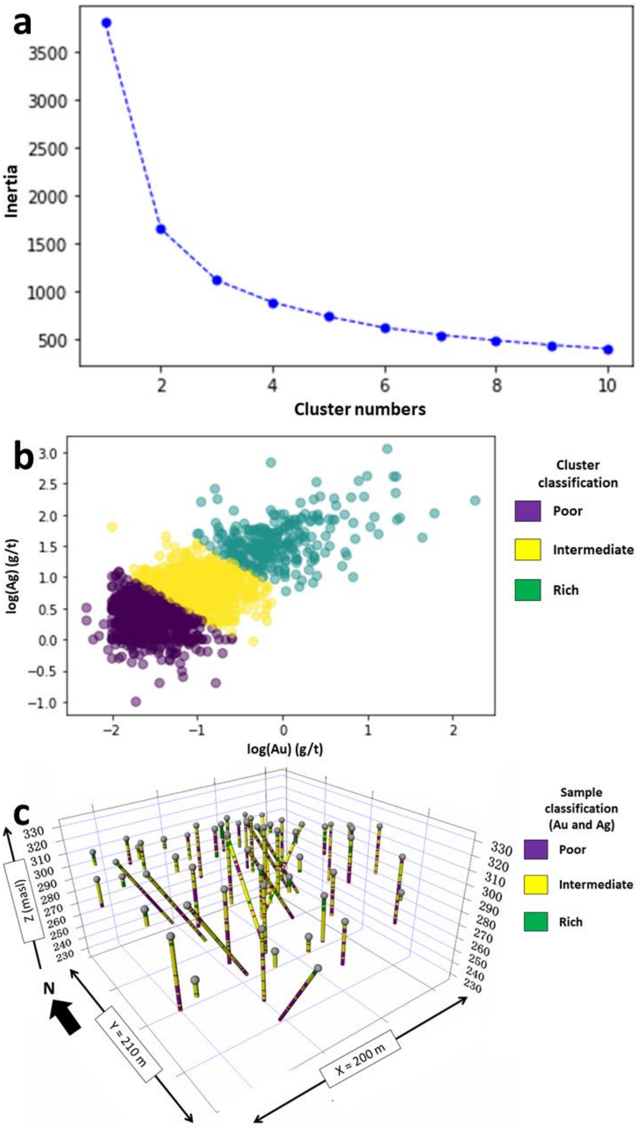

Figure 2. (a) Plot of inertia gains of hierarchical clustering results, (b) Biplot of log(Au) and log(Ag) showing cluster features, and (c) 3D representation of the cluster spatial distribution.