Clustering of down-hole physical properties measurement to characterize rock units at the Victoria Cu-Ni property

Omid Mahmoodi, and Richard Smith, Earth Sciences Department, Laurentian University, Sudbury, ON, Canada*

Summary

Magnetic susceptibility, density, and gamma ray counts are three physical properties measured within a hole drilled at Victoria property, were used to identify rock types. The fuzzy k-means clustering algorithm was used to divide the data into different clusters, each of which represented a rock group with similar physical properties characteristics. The number of clusters reveals the number of rock groups which are identifiable based on physical properties measurements. Characterization of physical properties of rocks within the hole will help to modify the geological model and plan further exploration activities.

Introduction

In recent years the number of exposed deposits has decreased significantly; consequently, exploration companies are transitioning from surface-based exploration to subsurface exploration. Therefore, geophysics becomes an important tool to explore below the surface in exploration activities (Smith et al., 2012; Williams, 2008). Geophysical data can be used to infer the physical properties of the rock mass, so they can sometimes be used to predict the rock type. Knowing the links between the physical properties and the geology is potentially useful for defining geological zones at depth (Perron et al., 2011). This approach is more likely to succeed when physical rock property contrasts exist between the different lithological units (Mwenifumbo and Mwenifumbo, 2012; Perron et al., 2011).

One approach is to use borehole geophysics to measure multiple physical properties down the hole (Killeen et al., 1997). Multi-probe in-hole measurements can be used to acquire a large amount of data rapidly; in addition, they are typically sensitive to a larger volume than the core (McDowell et al., 1998; Granek, 2011). The measured physical properties are numerical data that can be analyzed mathematically to extract patterns of variation of physical properties that could be related to different geological units. To reduce the subjective bias in interpretation and to quantitatively link the geological and geophysical data, pattern recognition techniques can be used. Classification methods, both supervised and unsupervised, are commonly used for classifying rock types and identifying alteration zones (Granek, 2011; Williams and Dipple, 2007).

This study seeks to determine if the physical properties can be used to distinguish different types of host rock at the Victoria Ni-Cu property operated by KGHM International. Gamma ray counts, magnetic susceptibility, and density measurements were acquired in hole FNX1182, down to 1700 meter depth. The fuzzy k-means clustering technique is used to analyze this data and group the rocks in terms of their physical properties.

Geological setting

The Victoria property is located at the west part of the south range of the Sudbury Igneous Complex (SIC) where the SIC intersects the Worthington Offset Dyke. The northern margin of the property is dominated by the norite and norite breccias. Cyclical repetitions of sedimentary sequences with felsic and mafic volcanic rocks and gabbroic intrusives comprise the footwall rocks. Quartz diorite offset dykes intrude the area and are faulted in a complex manner. There is Cu-Ni-PGE mineralization at Victoria, which is mainly hosted by quartz diorite and re-crystalized Sudbury breccia (Perron et al., 2011).

Physical properties

The gamma ray response, magnetic susceptibility, and density were measured down hole FNX1182 with a vertical resolution of 20 cm. Each tool measured the physical property at slightly different depth, so a weighted average of measurements were computed every 20 cm to obtain measurements at consistent depths. Because there is high frequency variation in the physical properties, the data were filtered using a low-pass filter to increase the signal-to-noise ratio. The average and standard deviation (std) of each physical property is summarized in Table 1 for each rock type identified by the geologist. Figure 2 also shows the physical properties logs and rock types identified by geologists.

The gamma ray measurements plotted on Figure 2 are highly variable in the hole. The quartz diorite (QD) unit is characterized by gamma-ray values that do not change significantly. However, the values are significantly less within the meta-gabbro (MTGB) and meta-basalt (MTBS). The peaks at approximate depths of 1280, 1380, and 1520 m reflect high gamma ray response of quartzite (QTZT) and meta-sediments (MTSD).

DOI http://dx.doi.org/10.1190/segam2014-0190.1

The magnetic susceptibility log shows anomalous behavior in the diabase (DIA) and sulphide (SULP) zones, where sharp peaks characterize DIA and a zone of strong and erratic variation characterizes SULP. The rest of the log does not show significant variation.

The density log characterizes the QD as relatively low density and low variability, while MTBS and MTGB show higher density. QTZT and MTSD show anomalously low values on the density log. A maximum peak in the log corresponds to the SULP zone at depths between 1500 to 1530 m depth.

Table 1: Mean and std of magnetic susceptibility, density and gamma measurements within FNX1182, (SUBX: Sudbury breccia, PYRT: pyroxenite; other codes explained in text or on Figure 2)

Rock type	Log MS (10^{-3}×SI)	Density (g/cm^3)	Gamma (API)
QD	1.173 ± 0.028	2.866 ± 0.066	90.1 ± 25.4
MTGB	1.191 ± 0.038	3.027 ± 0.078	58.9 ± 36.4
MTBS	1.170 ± 0.043	3.013 ± 0.114	62.5 ± 62.0
MTSD	1.141 ± 0.056	2.914 ± 0.189	155.5 ± 115.4
SUBX	1.169 ± 0.021	2.881 ± 0.139	146.1 ± 103.4
QTZT	1.146 ± 0.115	2.878 ± 0.257	200.1 ± 80.3
DIA	1.325 ± 0.296	2.980 ± 0.042	54.3 ± 16.3
SULP	1.218 ± 0.196	3.080 ± 0.227	84.7 ± 31.3
PYRT	1.197 ± 0.038	2.990 ± 0.096	60.7 ± 32.7
Total	1.177 ± 0.057	2.940 ± 0.125	83.5 ± 57.9

Fuzzy k-means clustering

Clustering is an unsupervised classification technique in which data are divided into clusters based on the variables measured at each data point (Lofts, 1993; Rabaute et al., 1997). The fuzzy clustering algorithm allows data points to belong to more than one cluster and the degree of membership of each datum to each cluster is defined as a membership value: the higher the membership value, the more strongly the datum belongs to the cluster (Zadeh, 1965 and Ruspini, 1969).

In fuzzy *k*-means clustering *n* data are divided into *k* clusters based on *v* variables. The fuzzy membership value of datum i^{th} datum, m_{ik}, $i=1,\ldots,n;\ k=1,\ldots,p$) is obtained through minimization of the objective function J:

$$J = \sum_{i=1}^{n}\sum_{k=1}^{p} m_{ik}^{\emptyset}\, d_{ik}^{2}$$

$$d_{ik}^{2} = \sum_{v=1}^{q}(x_{iv} - c_{kv})^{2}$$

where data are represented as vectors x_i which is the vector of n^{th} data. Each vector is composed of *v* elements and x_{iv} is the value of the v^{th} variable measured at the i^{th} datum. The fuzzy exponent is represented by ϕ. The term c represents the position of the centroid of the cluster. It is a vector with *v* elements so that c_{kv} is the average of the v^{th} variable of all the points weighted by their degree of membership to cluster *k*. The term d_{ik} represents the distance between the i^{th} data and the centroid of the k^{th} cluster (Bezdek, 1981; McBratney and DeGruijter, 1992). Initially fuzzy membership values of data are randomly assigned between 0 (lowest membership) to 1 (highest membership) and the total of membership values for each data are conditioned to be 1. The fuzzy membership values are recalculated to minimize the objective function. As a result, the centroid of clusters is relocated when the fuzzy membership values change. Therefore, a new vector of centroids are used in the objective function each iteration. This procedure continues until the objective function is minimized. So, the membership values for each datum and the position of centroids of clusters are the outputs of the algorithm. Finally, each datum belongs to the cluster for which it has the highest membership value (Bezdek, 1981; McBratney and DeGruijter, 1992; Rabaute et al., 2003).

Prior to running the algorithm, two parameters need to be assigned: the optimal number of clusters and the fuzzy exponent. The optimal number of clusters is determined based on prior knowledge and minimization of three functions, fuzzy performance index (FPI):

$$FPI = \frac{1-(p\times F-1)}{F-1}, \qquad F = \frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{p}(m_{ik}^{\emptyset})^{2}$$

modified partition entropy (MPE):

$$MPE = \frac{H}{logp}, \qquad H = -\frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{p} m_{ik} log(m_{ik})$$

where *H* represents the entropy function, and finally, the separation distance (S):

$$S = \frac{J}{n\,(d_{min})^{2}}$$

where d_{min} is the minimum distance between cluster centroids. Minimization of these parameters results in a maximization of compactness and separation of clusters (Roubens, 1982 and Xie and Beni, 1991).

The fuzzy exponent determines the degree of fuzziness which represents the compactness and separation of final clusters. This value can be assigned between 1 and infinity so that larger value results in higher fuzziness in the final clustering. Due to the nature of physical properties data and their contrast in different rock types, 2 was chosen as a

DOI http://dx.doi.org/10.1190/segam2014-0190.1

proper value for the fuzzy exponent (Bezdek, 1981; DeGruijter and McBratney, 1988; Rabaute et al., 2003).

Data Clustering

Data were analyzed statistically to remove any unrealistic or missing data. The clustering technique requires the data to be normalized and standardized between 0 and 1. Therefore, the logarithm (to the base 10) of the magnetic susceptibility and the gamma-ray data were used in the analyses. The program FUZME2.1 (Minasny and McBratney, 2000) was used to classify the data. The program requires the number of clusters and the value of the fuzzy exponent for fuzzy analysis. The optimal number of cluster was determined to be 3 as all three mathematical functions discussed above show minimum peak for three clusters (Figure 1).

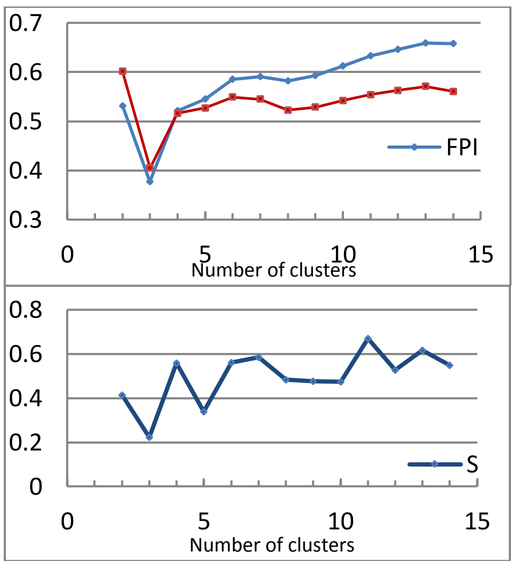

Figure 1- FPI, MPE, and S plots were used to determine the optimal number of clusters

Results

Data vectors at each depth are classified into the three clusters based on their fuzzy memberships of the three clusters. The result of clustering is shown in Figure 2. Physical properties logs, clusters, and rock types defined by geological logging are also shown in the plot. The percentage of contribution of rock types in different clusters are represented in Table 2. Clusters are described by their centroids and standard deviation. The centroid represents the weighted average of physical properties of the cluster and the standard deviation is a measure of the scatter of the data points in the cluster. Clusters are statistically characterized in the Table 3.

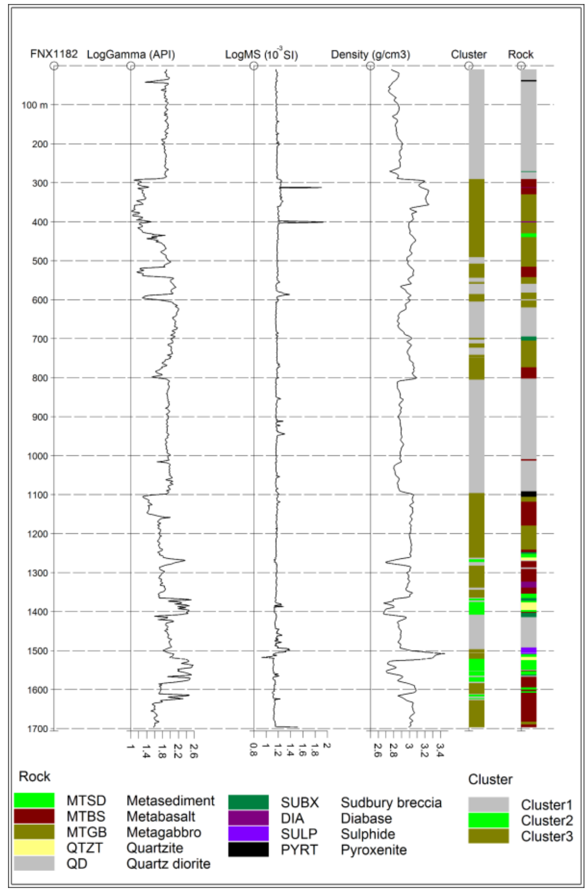

Figure 2- Gamma ray, magnetic susceptibility, density logs, rock groups and clusters within FNX1182 are shown in last two columns.

Table 2-Percentage of each rock type classified in three clusters

	Cluster1	Cluster2	Cluster3	Total Length (m)
QD	98	1	1	756
MTGB	19	0	81	375
MTBS	10	7	83	356
MTSD	5	44	51	88
SUBX	43	32	24	38
QTZT	13	75	13	34
DIA	0	0	100	19.5
SULP	29	0	71	15
PYRT	33	0	67	13
UM	100	0	0	3.57

Table 3- Average and standard deviation of physical properties in each cluster

Centroids	Log MS (10^{-3}×SI)	Density (g/cm^3)	Gamma (API)
Cluster1	1.174 ± 0.030	2.874 ± 0.073	91.3 ± 24.8
Cluster2	1.142 ± 0.037	2.784 ± 0.122	253.1 ± 75.9
Cluster3	1.185 ± 0.079	3.045 ± 0.095	48.3 ± 28.0
Total	1.177 ± 0.057	2.940 ± 0.125	83.5 ± 57.9

DOI http://dx.doi.org/10.1190/segam2014-0190.1

QD and UM were grouped together in Cluster 1 which represents relatively medium MS, density, and gamma ray response. UM represents high MS and low gamma which matches cluster 3, but the density of UM and QD are close together, which results in UM being grouped in cluster 1. The SUBX, PYRT, and SULP rock groups are partly classified into cluster 1.

Low MS, low density, and high gamma characterize cluster 2 in which QTZT is a dominant rock group, although MTSD and SUBX are present as a smaller population. The mixed membership of these rock types probably reveals the heterogeneity of MTSD and SUBX.

Cluster 3 comprises DIA, MTBS, MTGB, SULP, and PYRT. To lesser degree, MTSD and SUBX are partly grouped in cluster 3. High MS, high density, and low gamma ray are characteristics of cluster 3. MTGB and MTBS, the two most populous rock groups in cluster 3, are similar in density, gamma, and MS.

From the perspective of homogeneity, the UM, QD, MTGB, MTBS, DIA, QTZT, and to a lesser degree SULP and PYRT are homogenous rock groups which are dominantly (more than ~70%) grouped in one cluster. On the other hand, SUBX (present in three clusters) and MTSD (present in two clusters) are the two most heterogeneous rock groups in the hole. Histograms of density measurements within QD and SUBX units in Figure 3 demonstrate their homogeneity and heterogeneity, respectively.

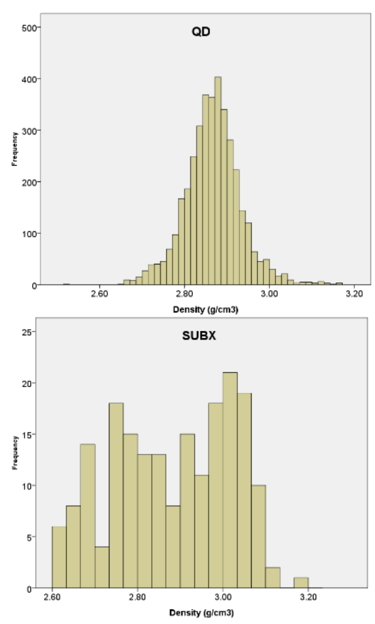

Figure 3- Histograms of density measurements within QD (top) and SUBX (bottom).

Conclusion

Comparison of rock type and clustering results illustrates that the fuzzy k-means clustering is a reliable algorithm to analyze physical properties measurements to identify rock types. A higher contrast in the physical properties of rocks enhances the certainty of the results. Clustering of physical properties provides a simplified geological model in which units with distinct physical properties are represented. At Victoria, three clusters describe the physical behavior of rock groups within the hole FNX1182. Such a simplified model can be used to modify the current geological model and plan further physical properties studies in Victoria property or other sites with a similar geological setting.

In some cases, physical properties manifest subtle difference in one rock type, which can result in a modified and more precise understanding of that rock type than provided by geologists. The observed heterogeneity of magnetic susceptibility and density in MTSD can be related to the magnetite content in the rock. Of the three physical properties used in this research, density and gamma ray measurements provided the most valuable information in clustering. We feel that the magnetic susceptibility can be used more effectively to detect sulfides, mineralization or alteration zones.

In the clustering process, those rock types with homogenous physical properties and more common rock types were dominantly grouped in unique clusters (e.g. QD); whereas, rock types with heterogeneous physical properties were divided into different clusters (e.g. MTSD). Different rock types with similar physical properties were clustered together (e.g. MTBS and MTGB). Not populous rock types usually are grouped with dominant rock types in a cluster which does not represent its physical properties because cluster characteristics are strongly controlled by more populous rock types.

The clustering produces a data dependent classification of the lithology. This provides geologists with another dataset that they can compare with their classification. By observing discrepancies and anomalies between the geological and statistical classification the attention may be drawn to specific locations, perhaps warranting more detailed study. This will help in understanding the geology and looking for mineral deposits.

Acknowledgments

Thanks to NSERC, Vale, Sudbury Integrated Nickel Operations A Glencore Company, KGHM International, Wallbridge Mining, and Centre for Excellence in Mining Innovation for providing financial and technical support.

http://dx.doi.org/10.1190/segam2014-0190.1

EDITED REFERENCES

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2014 SEG Technical Program Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

Bezdek, J. C., 1981, Pattern recognition with fuzzy objective function algorithm: Springer US.

Granek, J., 2011, Computing Geologically Consistent Models from Geophysical Data: University of British Columbia.

Killeen, P., C. Mwenifumbo, B. Elliott, and C. Chung, 1997, Improving Exploration Efficiency by Predicting Geological Drill Core Logs With Geophysical Logs: Fourth Decennial International Conference on Mineral Exploration.

Lofts, J. C., 1993, Integrated geochemical-geophysical studies of sedimentary reservior rock: PhD thesis, University of Leicester.

McBratney, A. B., and J. J. DeGruijter, 1992, A continuum approach to soil classification by modified fuzzy k-means with extragrades: Soil Science, **43**, no. 1, 159–175, doi: 10.1111/j.1365-2389.1992.tb00127.x.

McDowell, G. M., A. King, R. E. Lewis, E. A. Clayton, and J. A. Grau, 1998, In-situ nickel assay by prompt gamma neutron activation wireline logging: SEG Technical Program Expanded Abstracts, 772-775.

Minasny, B., and A. B. McBratney, 2000, FuzME version 2.1: Center for Precision Agriculture, University of Sydney, Australia. Retrieved from http://www.usyd.edu.au/su/agric/acpa

Mwenifumbo, C., and A. Mwenifumbo, 2012, Borehole Geophysical Logging in the Flin Flon Mining Camp: Geological survey of Canada.

Perron, G., P. Fullagar, G. Pears, N. Phillips, C. Williston, V. Gerrie, and J. Everest, 2011, 3D Litho-Prediction Model from In-Situ Physical Rock Property Logging and Constrained Potential Fields Data Inversion: 31st GOCAD meeting.

Rabaute, A., L. Briqueu, P. K. Harvey, and H. Mercadier, 1997, Characterisation of oceanic sediments and basalts from Hole 833B, New Hebrides island arc (Vanuatu) from downhole measurements: Scientific Drilling, **6**, 91–101.

Rabaute, A., B. Yven, W. Chelini, and M. Zamora, 2003, Subsurface geophysics of the Phlegrean Fields: New insights from downhole measurements: Geophysical Research, 108.

Roubens, M., 1982, Fuzzy clustering algorithms and their cluster validity: European Journal of Operational Research, **10**, no. 3, 294–301, doi: 10.1016/0377-2217(82)90228-4.

Ruspini, E. H., 1969, A new approach to clustering: Information and Control, **15**, no. 1, 22–32, doi: 10.1016/S0019-9958(69)90591-9.

Smith, R., M. Shore, and D. Rainsford., 2012, How to make better use of physical properties in mineral exploration: The Leading Edge, **31**, 330-337.

Williams, N., 2008. Geologically-constrained UBC–GIF gravity and magnetic inversions with examples from the Agnew-Wiluna greenstone belt, Western Australia: University of British Coloumbia.

Williams, N., and G. Dipple, 2007, Mapping Subsurface Alteration Using Gravity and Magnetic Inversion Models: Fifth Decennial International Conference on Mineral Exploration.

Xie, X. L., and G. Beni, 1991, A validity measure for fuzzy clustering. Pattern Analysis and Machine Intelligence: IEEE Transactions, **13**, 841–847.

Zadeh, L. A., 1965, Fuzzy sets: Information and Control, **8**, no. 3, 338–353, doi: 10.1016/S0019-9958(65)90241-X.